



DNA methylation as a prognostic marker in acute lymphoblastic leukemia

Magnus Borssén



Department of Medical Biosciences, Pathology
Umeå University 2016

DNA methylation as a prognostic marker in acute lymphoblastic leukemia.

Magnus Borssén



Department of Medical Biosciences, Pathology
Umeå University
Umeå 2016

Responsible publisher under Swedish law: the Dean of the Medical Faculty
This work is protected by the Swedish Copyright Legislation (Act 1960:729)
ISBN: 978-91-7601-583-4
ISSN: 0346-6612
New Series No:1857
Cover by: Vilhelm and Hanna Borssén, Leukemic blasts, inspired by the cartoon "*Il
était une fois... la Vie*"
Elektronisk version tillgänglig på <http://umu.diva-portal.org/>
Tryck/Printed by: Print & Media, Umeå University
Umeå, Sverige 2016

You can't even begin to understand biology, you can't understand life, unless you understand what it's all there for, how it arose - and that means evolution.
- Richard Dawkins

Table of Contents

Table of Contents	i
Abstract	iii
Populärvetenskaplig sammanfattning	iv
DNA metylering i akut lymfatisk leukemi	iv
Original papers	vii
Abbreviations	1
Background	2
Cancer	2
Hematopoiesis	2
Acute Lymphoblastic Leukemia	3
<i>Epidemiology and etiology</i>	3
<i>Symptoms, diagnosis and treatment- a short overview</i>	4
Molecular basis of BCP-ALL	6
<i>ETV6-RUNX1</i>	6
<i>Hyperdiploid</i>	7
<i>TCF3-PBX1</i>	7
<i>KMT2A(11q23)-rearranged</i>	7
<i>BCR-ABL1</i>	8
<i>Intra chromosomal amplification of chromosome 21</i>	8
<i>Other reoccurring genetic aberrations</i>	8
Molecular basis of T-ALL	9
<i>CDKN2 locus and NOTCH1</i>	9
<i>T-cell receptor and other translocations</i>	9
<i>Molecular subgroups in T-ALL</i>	10
<i>Early immature T-cell precursor phenotype</i>	11
Relapsed acute lymphoblastic leukemia	11
Epigenetics and DNA methylation	12
<i>The Epigenome</i>	12
<i>DNA Methylation and demethylation</i>	14
<i>DNA methylation in the genome</i>	16
<i>DNA methylation in the hematopoietic system</i>	18
<i>Epigenetics in leukemia</i>	19
Telomere length as a prognostic marker in leukemia	21
Aims	22
General aim	22
Specific aims	22
<i>Paper I</i>	22
<i>Paper II</i>	22
<i>Paper III</i>	22
<i>Paper IV</i>	22

Materials and methods	23
Patients	23
DNA methylation	24
<i>Methy-Light</i>	25
<i>High Resolution Melting curve analysis</i>	26
<i>DNA Methylation Arrays</i>	26
Defining CIMP	27
<i>Paper II</i>	27
<i>Paper III</i>	27
<i>Paper IV</i>	28
Gene expression	28
Telomere length	28
Results and Discussion	30
<i>Telomere length and hTERT promoter methylation in ALL (Paper I)</i>	30
<i>CIMP profile as a prognostic marker in T-ALL (Paper II&III)</i>	32
<i>Integrated gene expression and DNA methylation (Paper II)</i>	35
<i>Enrichment analysis of CIMP profile</i>	35
<i>ETP and DNA methylation (Paper II&III)</i>	36
<i>CIMP as a prognostic marker in relapsed BCP-ALL (paper IV)</i>	36
General discussion	38
DNA methylation as a prognostic marker in ALL	38
Driver or passenger?	39
Conclusions	42
Acknowledgements	43
References	45

Abstract

Acute lymphoblastic leukemia (ALL) is the most common childhood malignancy. Most ALL cases originate from immature B-cells (BCP-ALL) and are characterized by reoccurring structural genetic aberrations. These aberrations hold information of the pathogenesis of ALL and are used for risk stratification in treatment. Despite increased knowledge of genetic aberrations in pediatric T-cell ALL (T-ALL), no reliable molecular genetic markers exist for identifying patients with higher risk of relapse. The lack of molecular prognostic markers is also evident in patients with relapsed ALL. During the last decades, aberrant epigenetic mechanisms including DNA methylation have emerged as important components in cancer development. Telomere maintenance is another important factor in malignant transformation and is crucial for long-term cell survival. Like DNA methylation, telomere length maintenance has also been implicated to reflect outcomes for patients with leukemia.

In this thesis, the prognostic relevance of DNA methylation and telomere length was investigated in pediatric ALL at diagnosis and relapse. The telomere length (TL) was significantly shorter in diagnostic ALL samples compared to normal bone marrow samples collected at cessation of therapy, reflecting the proliferation associated telomere length shortening. Prognostic relevance of TL was shown in low-risk BCP-ALL patients where longer telomeres at diagnosis were associated with higher risk of relapse.

Genome-wide methylation characterization by arrays in diagnostic T-ALL samples identified two distinct methylation subgroups denoted CIMP+ (CpG Island Methylator Phenotype high) and CIMP- (low). CIMP- T-ALL patients had significantly worse outcome compared to CIMP+ cases. These results were confirmed in a Nordic cohort treated according to the current NOPHO-ALL2008 protocol. By combining minimal residual disease (MRD) status at treatment day 29 and CIMP status at diagnosis we could further separate T-ALL patients into risk groups.

Likewise, the CIMP profile could separate relapsed BCP-ALL patients into risk groups, where the CIMP- cases had a significantly worse outcome compared to CIMP+ cases. From these data we conclude that DNA methylation subgrouping is a promising prognostic marker in T-ALL, as well as in relapsed BCP-ALL two groups where reliable prognostic markers are currently missing. By elucidating the biology behind the different CIMP profiles, the pathogenesis of ALL will be further understood and may contribute to new treatment strategies.

Populärvetenskaplig sammanfattning

DNA metylering i akut lymfatisk leukemi

Cancer är en okontrollerad tillväxt av genetiskt förändrade celler som inte respekterar kroppens spelregler. Risken att drabbas av cancer ökar med ålder men även barn kan drabbas av cancer. Till skillnad från de flesta cancertyper är akut lymfatisk leukemi (ALL) vanligare hos barn än vuxna. Vid ALL sker en okontrollerad tillväxt av en specifik typ av blodkroppar, så kallade lymfocyter. Det finns två typer av lymfocyter, B och T lymfocyter som båda är viktiga för människans immunförsvar. Varför ca 100 barn drabbas av leukemi varje år i Sverige är inte känt, däremot vet vi att specifika genetiska förändringar gör att cellens tillväxtkontrollmaskineri sätts ur spel och gör att leukemicellerna tränger undan produktionen av friska blodkroppar. Genom studier av leukemicellernas olika genetiska förändringar (förändringar i arvsmassan) har man lärt sig att vissa förändringar är kopplade till en mer svårbehandlad sjukdom och dessa barn får en mer intensiv behandling. Detta gäller framför allt de barn som drabbas av B-cells leukemi. Hos barn med T-cells leukemi finns inte samma möjlighet, eftersom de återkommande förändringar man identifierat inte påverkar svaret på behandling i samma utsträckning som i B-cells leukemier. Även hos barn med återfall i sin sjukdom behövs bättre markörer för hur intensivt sjukdomen behöver behandlas.

Under de senaste decennierna har ett nytt forskningsfält vuxit fram, epigenetik. Epigenetik beskriver hur DNAt är packat och strukturerat i cellen, vilket har betydelse för arvsmassans stabilitet och påverkar hur gener uttrycks. Studier har visat att den epigenetiska regleringen ofta är urspårad i cancerceller på ett sjukdomsspecifikt sätt. Epigenetiska förändringar kan antingen ske på de proteiner (histon) som DNA strängen är upplindad på eller på själva byggstenarna (nukleotiderna) i DNA kedjan. Den mest studerade epigenetiska förändringen är DNA metylering, vilket är när en metylgrupp binds till cytosin, en av de fyra nukleotider som bygger upp DNA.

En annan och viktig mekanism för cancercellers förmåga till överlevnad är att motverka kritiskt korta telomerer. Telomerer är ändstrukturerna på cellens kromosomer. Telomererna förkortas varje gång en cell delar sig och fungerar som en biologisk klocka för cellen. När telomererna blir kritiskt korta slutar cellen att dela på sig. För att kompensera förkortningen i celler som kräver konstant delning (exempelvis lymfocyter och stamceller) kan ett proteinkomplex, telomeras, förlänga telomererna. Att kunna bibehålla en adekvat telomerlängd är essentiellt för cancerceller som ständigt delar på sig.

Målet med denna avhandling var att kartlägga DNA metylering och mäta telomerlängd i akut lymfatisk leukemi och undersöka dessa faktorer prognostiska betydelse. Fokus har legat på att hitta nya markörer som avspeglar prognos, framför allt i grupper där det idag saknas bra markörer.

Fyra delarbeten ingår i avhandlingen, i det första arbetet undersöktes telomerlängd och metylering av genen *hTERT*, som är en av de viktigaste komponenterna i telomeraskomplexet. I den första studien drog vi slutsatsen att patienter med långa telomerer hade en något ökad risk att få återfall i sin sjukdom. Dessutom såg vi att frekvensen av metylering av *hTERT* genen skilde sig mellan olika undergrupper av leukemier. Resultaten från den första studien gjorde att vi ville göra en utökad studie av metyleringsmönstret i T-cells leukemier (T-ALL). T-ALL är en undergrupp av leukemier som historiskt sett haft en sämre prognos och där det har saknats bra markörer som kan förutsäga hur intensivt en patient behöver behandlas, vilket gjort att man betraktat alla T-ALL som högriskpatienter. Med hjälp av så kallade arrayer analyserades tusentals metyleringsförändringar samtidigt i ett och samma prov. I både delarbete II och III studerades T-ALL patienter. Vi visade att hos patienter med T-cells leukemi finns två grupper med skilda metyleringsmönster; en grupp med många DNA-metyleringsförändringar (en grupp som i avhandlingen kallas CIMP-hög) och en grupp med få metyleringsförändringar (CIMP-låg).

I delarbete II analyserades patienter som behandlats för T-ALL mellan 1992-2008. Vi såg att patienter med få metyleringsförändringar svarade mycket sämre på behandling medan patienter med många metyleringsförändringar i regel svarade mycket bra på behandling.

I delarbete III verifierade vi våra resultat på patienter med T-ALL som behandlats med det nuvarande behandlingsprotokollet vilket använts sedan 2008. Det nuvarande behandlingsprotokollet skiljer sig en del jämfört med det tidigare. Framför allt övervakar man numer behandlingssvaret genom att mäta hur mycket leukemiceller som finns kvar efter att den första behandlingsomgången givits. Patienter som har kvar mer än 0,1% leukemiceller ges då den mest intensiva behandlingen och betraktas som högrisk patienter. I delarbete III såg vi att de patienter som hade kvarvarande sjukdom efter första behandlingsomgången och dessutom var CIMP-låga hade en signifikant ökad risk för återfall. CIMP-hög metyleringsprofil var kopplat till ett mycket bra terapivar trots kvarvarande sjukdom efter första behandlingsomgången.

Slutligen analyserades data från 601 patienter med B-cells leukemi. Vi genomförde samma typ av metyleringsanalys som på T-ALL patienter. Resultaten visade att av de 137 patienter som fick återfall i sin sjukdom

överlevde 2/3 av alla som hade en CIMP-hög metyleringsprofil medan bara 1/3 av alla med CIMP-låg profil överlevde.

Den sammantagna slutsatsen från avhandlingen är att vi utifrån metyleringsmönstret kan förutsäga svaret på behandling i två grupper som tidigare saknat tillförlitliga biologiska markörer, det vill säga patienter med T-cells leukemi och patienter med B-cells leukemi som får återfall efter första behandlingen. Fynden är viktiga ur flera aspekter, dels att vi identifierar grupper som i regel svarar bra på behandling, då dessa patienter löper en risk att överbehandlas. Den andra viktiga poängen att identifiera de patienter som inte svarar på nuvarande behandling och därmed skulle kunna vara kandidater för nya behandlingsstrategier. Att vidare kartlägga vilka mekanismer som ligger bakom skillnaden i metyleringsmönstret kan även ge nya infallsvinklar på hur mer skräddarsydda behandlingar skulle kunna se ut.

Original papers

This thesis is based on the following papers, which will be referred to by the corresponding Roman numbers (I-IV)

- I. **Borssén M**, Cullman I, Norén-Nyström U, Sundström C, Porwit A, Forestier E and Roos G. hTERT promoter methylation and telomere length in childhood acute lymphoblastic leukemia: associations with immunophenotype and cytogenetic subgroup. *Exp Hematol.* 2011; 39:1144-1151.
- II. **Borssén M**, Palmqvist L, Karrman K, Abrahamsson J, Behrendtz M, Heldrup J, Forestier E, Roos G and Degerman S. Promoter DNA methylation pattern identifies prognostic subgroups in childhood T-cell acute lymphoblastic leukemia. *PLoS One.* 2013; 6;8(6)
- III. **Borssén M***, Haider Z*, Landfors M, Norén-Nyström U, Schmiegelow K, Åsberg AE, Kanerva J, Madsen HO, Marquart H, Heyman M, Hultdin M, Roos G, Forestier E, Degerman S. DNA Methylation Adds Prognostic Value to Minimal Residual Disease Status in Pediatric T-Cell Acute Lymphoblastic Leukemia. *Pediatr Blood Cancer.* 2016; 63:1185-1192 *Contributed equally
- IV. **Borssén M**, Nordlund J, Haider Z, Landfors M, Larsson P, NOPHO collaborators, Forestier E, Heyman M, Hultdin M, Lönnerholm G, Syvänen AC, and Degerman S. DNA methylation holds prognostic information in relapsed precursor B-cell acute lymphoblastic leukemia. Manuscript.

Abbreviations

5mC	5-Methylcytosine
ALL	Acute Lymphoblastic Leukemia
BCP-ALL	B-cell precursor ALL
cBM	Combined extra medullary and bone marrow relapse
CD	Cluster of differentiation
CGI	CpG island
Ch	Chromosome
CIMP	CpG island methylator phenotype
CNS	Central nervous system
CpG	Cytosine phosphate guanine
CR	Complete remission
DMG	Differently methylated gene
DNMT	DNA methyl transferase
EFS	Event free survival
EM	Isolated extra medullary
ETP	Early T-cell precursor phenotype
HeH	High hyperdiploid
HR	High risk
HSC	Hematopoietic stem cell
iBM	Isolated bone marrow
IR	Intermediate risk
KDM2a-r	Lysine demethylase 2A rearranged
MRD	Minimal residual disease
NOPHO	Nordic society of pediatric hematology and oncology
OS	Overall survival
PCR	Polymerase chain reaction
PRC	Polycomb repressive complex
RTL	Relative telomere length
SR	Standard risk
T-ALL	T-cell acute lymphoblastic leukemia
TET	Ten eleven translocation

Background

Cancer

Cancer is the result of accumulating lesions at the cellular level leading to transformation of a benign to a malignant cell. Malignant cells harbor certain hallmarks i.e. resisting cell death, sustaining proliferative signaling, evading growth suppressors, immortal replication, generating angiogenesis and the ability to invade and form metastasis [1]. These phenotypic changes often have genotypic explanations. Amplification of oncogenes sustains proliferative signaling, deletions of tumor suppressors evade growth arrest and mutations in DNA repair machinery components can cause genomic instability [1,2]. Different tumors often acquire specific spectra of genetic abnormalities, which is most evident in hematological malignancies. In leukemia, information about these structural genetic aberrations can be used as biomarkers for prognosis and enable stratification of patients into different treatment intensities [3].

Transcription of genes does not occur on naked DNA but at the chromatin level. Chromatin is formed by histones and DNA, and can undergo chemical modifications via, for example, addition of methyl groups that may influence gene expression without affecting the DNA sequence. Configuration of chromatin is also essential to maintain genomic stability, so just like genetic lesions epigenetic changes in cancer cells can drive expression of oncogenes, silence tumor suppressors and cause genomic instability[4].

Hematopoiesis

The hematopoietic system starts to develop early during embryogenesis in the yolk sac and specific tissues in the aorta, after which the process is relocated to the fetal liver. Finally, during the perinatal period hematopoiesis is translocated to the bone marrow where it persists throughout life [5].

Hematopoiesis has a hierarchical structure, at the top of which lie hematopoietic stem cells (HSC). HSCs have the ability of both self-renewal and differentiation into more committed precursor cells. The number of HSC in the bone marrow is very low, the majority of which are in a quiescent state [6]. More committed precursor cells have made lineage decision into either common lymphoid precursor (CLP) cells or common myeloid precursor (CMP) cells. Further lineage commitment of CMP results in production of myeloid cells like erythrocytes, megakaryocytes, neutrophils and monocytes.

Lymphocytes arise from CLP, which can differentiate, into B-cells or T-cells as well as NK-cells [7].

This classical view of a strict bifurcation between myeloid and lymphoid lineages is probably not accurate since there are progenitors with the ability to generate both T-cells, B-cells and granulocytes [7].

Acute Lymphoblastic Leukemia

Epidemiology and etiology

The first known case of childhood acute lymphoblastic leukemia (ALL) was described in 1850 in a 9 years-old girl. Today we know that ALL is the most common form of childhood cancer, accounting for almost 1/3 of all cancers in children. The incidence of malignancies among children in Sweden is 11,4 cases /100 000 children/year equivalent to about 250-300 cases each year, of which 75-80 are ALL [8]. The incidence has been stable over the last four decades[9] .

In a global perspective there are regional differences in ALL incidence. ALL is most common among Europeans and Caucasians of the United States of America (USA), and lowest among Indians and African Americans of the USA, indicating ethnically linked susceptibilities. Unlike most other types of cancer, ALL has a higher incidence among children than adults, with an annual incidence peak in 2-5 years-old children [10]. The etiology of childhood ALL is largely unknown. Some syndromes are associated with an increased risk of childhood-ALL. The most studied is Down's syndrome (DS), comprising 2% of all patients with ALL, in comparison with 0,01% in the general population. ALL in DS is associated with worse outcome compared to patients without DS [11]. Other rare genetic conditions with increased risk of developing ALL are Bloom's syndrome, Nijmegen breakage syndrome and ataxia telangiectasia [12].

Genome wide association studies have identified a handful of single nucleotide polymorphisms (SNPs) associated with increased risk for ALL. Identified SNPs are located in loci linked to *ARID5B* (AT-rich interaction domain), *IKZF1* (IKAROS family zinc finger 1), *CEBPE* (CCAAT/enhancer binding protein epsilon), *GATA3* (GATA binding protein 3) and *CDKN2A-CDKN2B* (Cyclin dependent kinase 2 a and b) [13-15]. Of interest is that most of these genes are frequently deleted or mutated in ALL, and in the case of *CDKN2A* the risk allele is often the preserved copy in leukemic cells with loss of heterozygosity (LOH) of *CDKN2A*[16].

Symptoms, diagnosis and treatment- a short overview

ALL is an uncontrolled clonal expansion of malignant lymphocytes and without proper treatment the disease is fatal. Expansion of the malignant clone in the bone marrow eventually leads to repression of normal hematopoiesis. Hence, patients with ALL present signs including (i) a dysfunctional hematopoietic system like fever caused by the leukemia itself, (ii) prolonged or repetitive infections due to repressed immune function, (iii) paleness and fatigue reflecting repressed erythropoiesis, or (iv) even bruising as a sign of thrombocytopenia. Extra medullary involvement is common, blasts can infiltrate the central nervous system (CNS), spleen, liver, bone, testis and lymph nodes. The extra medullary spread can give rise to symptoms like limping, lymph node enlargement or more rarely compression of vessels or airways.

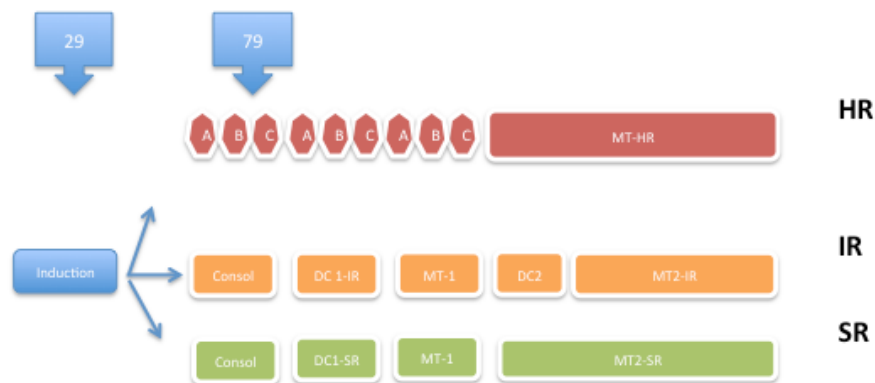
When leukemia is suspected, a bone marrow biopsy/aspiration is performed. Samples are examined morphologically and classified according to the French-American-British (FAB) criteria [17]. Still, as the FAB criteria are of limited clinical significance, immunophenotypic classification and molecular genetics remain cornerstones in diagnosing and risk stratification of ALL [3,18].

When it comes to treatment, there are few counterparts in medical history with such an exceptional development than the treatment of childhood leukemia. This might seem odd given that very few new drugs have been introduced over the last 40-50 years. Vincristine, prednisone, methotrexate and mercaptopurine were already in use in the 1960's but with limited success. Instead real success was achieved by combining drugs together with increased knowledge about the importance of supportive care. Further, careful evaluation of treatment protocols and better insight in the biology of ALL has led to the ability to risk stratify patients into different treatment intensities[18].

In the current Nordic protocol NOPHO ALL 2008 (Nordic Society of Pediatric Hematology and Oncology) the following three risk groups are defined: (i) standard risk (SR), (ii) intermediate risk (IR), and (iii) high risk (HR) [19]. Risk group stratification is important from several aspects. First,

to give tougher treatment to patients with high risk of disease relapse, and second not to over treat patients with a more favorable outcome. In NOPHO ALL 2008 and most other modern protocols, stratification is based on a combination of biological features of the leukemia i.e. immunophenotype and genetic aberrations, and response to treatment [18]. Treatment response is determined by detection of residual leukemic cells (Minimal Residual Disease =MRD) at certain time points during the treatment. MRD is analyzed by qPCR of leukemia-specific immunoglobulin (Ig) or T-cell receptor (TCR) –rearrangements or by detection of leukemia specific immunophenotypic patterns by flow cytometry [20,21]. If the bone marrow contains more than 0,1% leukemic cells at the end of induction (at treatment day 29) the patient is translocated to a higher risk group. Final risk stratification is determined after MRD analysis at day 79 [19].

All patients receive a cortisone based induction treatment, after which treatment is based on the risk groups to which the patient is allocated. SR and IR treatment have similar structure focused on antimetabolite treatment, but the IR arm includes extra cycles of antracyclines and alkylating agents and includes more CNS directed therapy. High risk patients are treated with an intensive block-based regime, and is very different from SR and IR protocols. The high risk arm includes antracyclines and alkylating drugs and includes topoisomerase inhibitors. The total treatment time for all risk groups is about 2,5 years [19].



Figur 1.

Schematic overview of the NOPHO ALL 2008 treatment protocol. Treatment intensities are color-coded. Green represent standard risk (SR), yellow intermediate risk (IR) and red represent high risk (HR). Bars represent different treatment phases. Consol: Consolidation, DC1: Delayed intensification I & Consolidation II, MTI: Maintenance I, DC2: Delayed intensification II & Consolidation III (only IR), MT2: Maintenance II. Septagons marked A, B and C represent block treatment in HR. Blue arrowed squares represent time-points (treatment day) for bone marrow examination for risk stratifications.

Molecular basis of BCP-ALL

Acute leukemia of B-cell origin account for 85-90% of all ALL cases in childhood. The leukemic cells usually have a B-cell precursor phenotype and are characterized by recurrent structural genetic aberrations or chromosomal modal number abnormalities. These cytogenetic aberrations are almost mutually exclusive and of great importance not only from a biological perspective but also hold important clinical information. All modern treatment protocol utilizes this information for risk stratification [3,22].

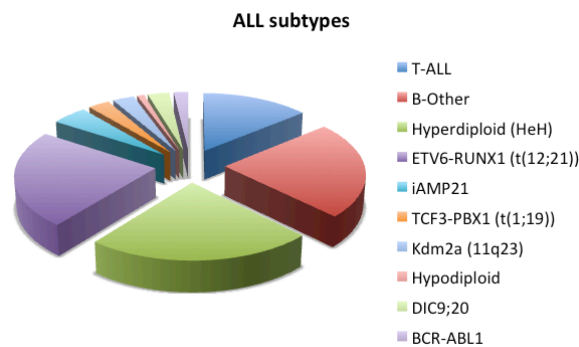


Figure 2.

Pie chart representing distribution of cytogenetic and immunophenotypic subtypes in pediatric ALL. Data retrieved from a representative cohort of 764 children diagnosed with ALL during 1996-2008 in the Nordic countries (Nordlund et al Genome Biol 2013)

ETV6-RUNX1

Translocation between chromosomes 12p13 and 21q22 is detected in about 25% of all cases and by far the most common translocation in BCP-ALL, resulting in the fusion of genes *ETV6* (ETS variant 6) and *RUNX1* (Runt related transcription factor 1). *ETV6-RUNX1* has to be considered a weak oncogene based on a number of observations. Transgenic mouse models of *ETV6-RUNX1* fail to develop full blown leukemia but generate maturation arrest in subpopulations of B-cells, which require additional oncogenic events to develop into an overt leukemia [23-25]. The same is observed in human samples, and additional genetic lesions are always detected in *ETV6-RUNX1* samples, such as deletion of *PAX5* (Paired box 5) and *RAS* (Rat sarcoma) mutations or deletion of the remaining *ETV6* copy. The cooperating oncogene that is involved in malignant transformation does not appear to influence the generally excellent outcome in *ETV6-RUNX1* BCP-ALL [26-28]. Occurrence of *ETV6-RUNX1* has a distinct age peak around 2-

5 years of age but is a rare entity in adults, which is believed to reflect a prenatal origin of ETV6-RUNX1 [5].

Hyperdiploid

In 25% of BCP-ALL diagnostic samples a gain of 5-21 extra chromosomes (total 51-67 chromosomes) are observed, a karyotype denoted hyperdiploid. The non-random gain of chromosomes most frequently involves chromosomes 21, 6, X, 14, 4, 18 and 17 [22,29]. The underlying process is unknown but whole genome studies suggest that the increased modal number is the driving event, probably through dosage skewing of gene expression [30]. Hyperdiploid ALL cases lack signs of chromosomal instability otherwise seen in hyperdiploid solid tumors, and the most common mutations are in receptor tyrosine kinase pathways (RTK), like RAS. Hyperdiploidy is most prevalent in young children with a peak around 2-5 years of age and the prognosis is often favorable especially in younger children [31,32].

TCF3-PBX1

TCF3-PBX1 is the product of a fusion between *TCF3* (transcription factor 3) on chromosome 1q23 and *PBX1* (PBX homeobox 1) on chromosome 19p13.3, initially identified as a group of children with inferior outcome. TCF3-PBX1 translocation can be either a balanced (25%) or unbalanced (75%), in both cases the fusion transcript result in activation of PBX1 controlled genes. Intensified treatment in this group of patients has improved outcome significantly [33].

TCF3 is occasionally involved in another fusion together with the *HLF* gene, and the outcome in this group of patients is extremely poor [34].

KMT2A(11q23)-rearranged

Lysine-specific methyl transferase 2a (KMT2A) also known as MLL (Mixed Lineage Leukemia) is a promiscuous translocation partner, over 80 different partners have been identified [35]. KMT2A rearrangement (KMT2A-r) is frequently seen in infant leukemia, a distinct entity with an aggressive disease course linked to high white blood cell count at diagnosis and high risk of relapse and death. Infant leukemia with KMT2A-r also exhibits very few reoccurring genetic abnormalities. About 80% of infant-ALLs are KMT2A-r whereas only about 3% of ALL in children over 1 year of age are KMT2A-r [36]. Although KMT2A-r do not have the same profound impact on outcome in non-infant BCP-ALL it is still regarded as a high risk leukemia

[36]. The molecular pathology of KMT2A is at least in part regulated through epigenetic changes and will be discussed later in this thesis.

BCR-ABL1

Breakpoint cluster region - ABL protooncogene 1 (BCR-ABL1) fusion, also known as Philadelphia chromosome (Ph+), is the archetype of leukemic translocations first discovered in 1960 [37]. BCR-ABL1 is detected in about 3-5% of all pediatric BCP-ALL. The fusion results in constitutively active tyrosine kinase ABL1, driving the oncogenic process [38]. BCR-ABL1 is associated with older age at diagnosis and is more common in adults than in children [3]. The development of the BCR-ABL1 specific tyrosine kinase inhibitor (TKI) Glivec® has had a positive effect on the outcome in this group of patients who had a dismal prognosis in the pre-TKI era. However, Ph+ patients are still considered as a high risk subgroup [39,40].

Intra chromosomal amplification of chromosome 21

Intra chromosomal amplification of chromosome 21 (iAMP21) was first noted as an amplification of *RUNX1*, but was later shown to be the result of repeated breakage-fusion-bridges cycles [41]. Initially iAMP21 was associated with high risk of relapse, but intensified treatment has improved the prognosis [42]. Median age among iAMP21 patients is 9 years and has never been reported in adults over 25 years of age [41].

Other reoccurring genetic aberrations

About 30% of all BCP-ALL cases are not assigned to any specific cytogenetic group, and they are often referred to as B-other. During the last ten years new molecular techniques have enabled discovery of recurrent submicroscopic lesions. Genes like *IKZF1* (IKAROS family zinc finger 1), *CRLF2* (cytokine receptor-like factor 2), *PAX5* and *ERG* (ETS transcription factor) are frequently mutated or deleted, and they may hold prognostic information. However, results from different studies are conflicting [43]. The most promising marker is probably *IKZF1* especially among B-other cases where it is associated with negative effect on outcome [44-46].

Gene expression array analysis have also identified a molecular subgroup named BCR-ABL1-like. This group lacks the BCR-ABL1 translocations but has a similar gene expression signature as Ph+ cases. BCR-ABL1-like cases constitute a significant part of the B-other subgroup and are associated with worse prognosis. Different gene expression signatures to define this entity

exists but they are only partly overlapping, and currently not of clinical use [47].

Molecular basis of T-ALL

T-cell phenotype in ALL has an uneven distribution both regarding age at diagnosis and sex. Among younger children 10% of ALL patients have a T-phenotype, compared to around 30% among patients aged 18-45. Irrespective of age there is a male predominance [19,48]. Several genetic alterations cooperate in the malignant transformation of developing T-cells, resulting in differentiation block, loss of cell cycle control and activation of cell proliferation.

CDKN2 locus and NOTCH1

The most frequent genetic alterations in T-ALL are deletions involving the CDKN2 loci and activating mutations of *NOTCH1* signaling, both occurring in more than 50% of all cases [49,50].

The CDKN2 locus is located on chromosome (ch) 9p21 and contain the gene encoding p14, which stabilize p53 and hence participate in DNA damage control. This locus also contains genes that encode p15 and p16 both involved in cell cycle control [51,52].

The *NOTCH1* gene is also located on chromosome 9 but at (9q34) and encodes a trans membrane receptor. Activation of the receptor initiates an intercellular signaling pathway that leads to activation of oncogenes like MYC (*v-myc* avian myelocytomatosis viral oncogene) and HES1 (*hes* family bHLH transcription factor 1) [53]. *NOTCH1* signaling in T-ALL can also be achieved through inactivation of FBXW7 (F-box and WD repeat domain containing 7), a negative regulator of *NOTCH1*. Hence loss of FBXW7 also leads to activation of *NOTCH1* signaling [52].

Several studies have investigated the clinical impact of *NOTCH1* activation showing no coherent results but it seems clear that *NOTCH1* activation is not a negative prognostic factor [50,54].

T-cell receptor and other translocations

Another common oncogenic event in T-ALL is translocations involving a T-cell receptor (TCR) and an oncogene. The oncogene in question is usually placed under the control of either TCRB (7q34-35) or TCRD (14q11) leading to aberrant expression of the oncogene. TLX1 (T cell leukemia homeobox 1),

TAL1 (TAL bHLH transcription factor 1), TLX3 (T cell leukemia homeobox 3) and LMO2 (LIM domain only 2) are the most frequent oncogenic transcription factors involved in rearrangement and are present in 5-25% of T-ALL cases and usually involve TCRD [55]. TAL1 expression can also be deregulated through non-TCR translocations. A deletion near the TAL1 gene in ch1p32 positions TAL1 under the control of the STIL (SCL/TAL1 interrupting locus) promoter [52].

KMT2A-r is sometimes detected in T-ALL, its impact on prognosis is not fully elucidated but might reflect a higher risk of treatment failure[56].

Molecular subgroups in T-ALL

In a seminal paper by Ferrando et al it was shown that T-ALL could be divided into molecular subgroups based on gene expression profiles. These profiles were based on expression of oncogenes known to be involved in translocations as described above. They showed that some of the T-ALL samples lacked structural genetic rearrangements in the known locis but still overexpressed genes like TAL/LMO, TLX1, TLX3 or HOXA (Homeo box A cluster) [57]. The different subgroups reflect specific blocks during the course of T-cell development, where TAL1 or LMO over expression reflect a more mature cortical phenotype. High TLX expression is linked to an early cortical immunophenotype and MEFC2 and LYL1 (Lymphoblastic leukemia associated haematopoiesis regulator 1) overexpression are linked to early stages of T-cell development [57,58](Figure 3).

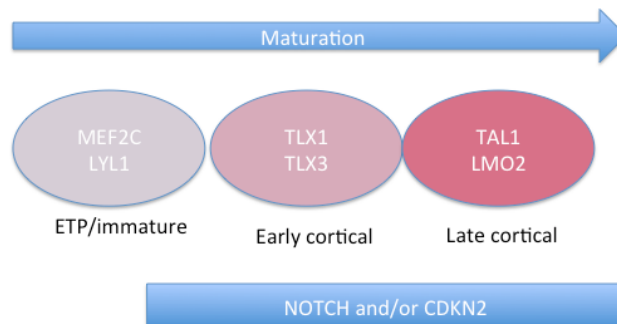


Figure 3.

Molecular subgroups in T-ALL, overexpression of key transcription factor involved in T-ALL development associate to block in specific maturation stages. MEF2C and LYL1 driven leukemia are associated with an immature state, TLX overexpression is linked to an intermediate early cortical phenotype. TAL1 and LMO driven T-ALL associate with the most mature form of T-ALL. NOTCH1 activating mutations and CDKN2 deletions are more frequent in the more mature phenotypes.

Early immature T-cell precursor phenotype

Especially the immature phenotype has gained considerable attention during the last years. In 2009, Coustan-Smith et al identified a subgroup of T-ALL patients with an early immature T-cell precursor (ETP) phenotype (CD4⁺, CD8⁻ and CD1a⁻ and expressed myeloid and stem cell markers). The ETP cells also had a gene expression pattern similar to early thymic progenitor cells and retained their ability to differentiate into a myeloid lineage [59]. In the first studies ETP patients had a very poor outcome making ETP-phenotype a promising stratifying marker in T-ALL treatment. Later studies have however failed to confirm the prognostic value of ETP [60,61].

Relapsed acute lymphoblastic leukemia

Among children treated according the NOPHO ALL 2000 protocol 17% had a relapse as their primary event, an incidence at the level of “common” solid tumors like Wilms tumor or neuroblastoma [62]. An important issue at relapse is to establish whether the new leukemia is a relapse or a second leukemia. A relapse should exhibit the same genetic aberrations and immunophenotype as the initial leukemia.

Time from initial diagnosis to relapse together with relapse localization are the most powerful prognostic factors when it comes to relapse [63]. As defined by the Berliner-Frankfurt-Munster study group (BFM) very early relapses are from diagnosis until 18 months after. Early relapses are from 18 months after diagnosis until 6 months after cessation of therapy, the rest are classified as late relapses. The earlier relapse the worse prognosis [64]. Regarding relapse site, isolated bone marrow relapse (iBM) has the worst outcome, followed by combined extramedullary and bone marrow relapse (cBM), best outcome is seen in patients with isolated extramedullary relapse (EM). T- or B-cell immunophenotype also influence relapse outcome with worse prognosis in relapsed T-ALL [62,63]. Cytogenetic background has an impact on relapse treatment outcome where BCR-ABL1 is associated with inferior survival rate, at least for the pre TKI-era [65]. In a recent Nordic study of relapsed ALL patients between 1992 and 2012 were analyzed, OS ranged from 8% in T-ALL patients with very early relapse to 82% for for BCP-ALL with late EM relapses [62].

Epigenetics and DNA methylation

The Epigenome

DNA is organized at several levels in the cell and the basic organizing structure is the nucleosome, formed by two sets of the four core histones H2A, H2B, H3 and H4. One hundred and forty-seven DNA nucleotides are wrapped around each histone octamer and together they form the nucleosome. The different histone residues can undergo posttranslational modifications like methylation (one, two or tri methylation) and acetylation. The effect of a specific modification of DNA or histones is based on several factors including location in the genome (i.e. in promoter region of genes or in gene bodies), co-localization of other histone modifications and spatial organization of the chromatin in the cell [66]. Table 1 provides an overview of different modifications of histones and DNA.

Genomic area	Active state	Bivalent state	Inactive state
Promoter	H3K4 me ² /me ³ H4 acetylation	H3K4 me ² /me ³ H3K27 me ³	H3K27 me ³ H3K9 me ³ DNA methylation
Gene body	H3K79me ² H3K36me ³ DNA methylation		H3K9me ² or ³
Enhancer Regions	H3K4 me ¹		H3K9 me ² /me ³ DNA methylation

Table 1.

Overview of chromatin signatures in relation to gene regulation. Me² =di methylation Me³ =tri methylation

As seen in the table 1, bivalent genes have markers of both active and inactive transcription. Genes with bivalent markers are often involved in differentiation processes. In primitive cells like embryonic stem cells (ESC) or hematopoietic stem cells (HSC), repression of bivalent genes are mediated through Polycomb repressive complex 1 and 2 (PRC1 and PRC2)[67]. Enhancer of Zeste Homologue 2 (EZH2) is the catalytic component of PRC2 responsible for maintaining H3K27me3. EZH2 forms a complex with Suppressor of Zeste 12 (SUZ12) and Embryonic Ectoderm Development (EED), both essential for the enzymatic activity of the complex [66,68]. PRC2 can interact with PRC1 thereby enabling ubiquitination of histone H2 and further repress transcription and promote compacting chromatin. When the PRC complexes are depleted from lineage specific genes they can become actively transcribed and differentiation can proceed [66]. PRC occupancy of gene promoters also appears to protect from DNA methylation [69].

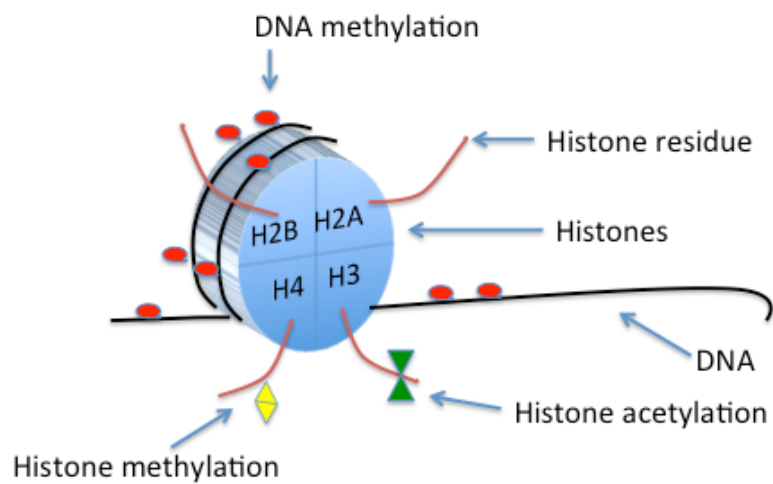


Figure 4.
Schematic picture over nucleosome composition. Red circles represent DNA cytosine methylation, yellow triangles represent histone methylation and green triangles represent histone acetylation.

DNA Methylation and demethylation

DNA methylation generally refers to addition of a methyl group to the 5th carbon atom in the cytosine base. Cytosines are almost exclusively methylated when followed by guanine, a unit referred to as CpG site. CpG methylation is an enzymatic process mediated by a group of methyl transferases. Five DNA methyl transferases (DNMT) have been identified, of which only the three DNMTs (DNMT 1, 3a and 3b) with enzymatic activity will be discussed here. DNMT1 is maintenance transferase, passing on the methylation pattern from mother to daughter cell. DNMT1 is recruited to the replication fork and copies an already existing methylation pattern by methylation of hemi-methylated DNA. DNMT3a and b are *de novo* transferases and can add new methyl groups to non-methylated CpG-sites. They are consequently independent of the template pattern (figure 5) [70,71].

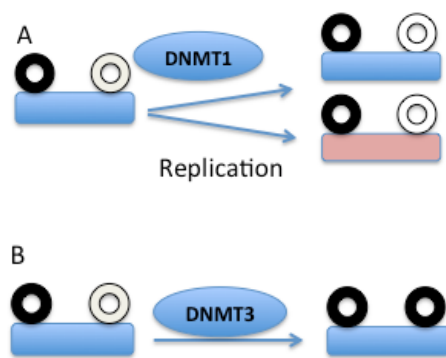


Figure 5.

- A. DNMT1 copy an already existing DNA methylation pattern when DNA is replicated, daughter cell (red) will inherit the methylation pattern from the mother cell (blue). White circles represent unmethylated CpG and black circles methylated CpG
- B. Both DNMT3A and DNMT3B add methyl groups in non-methylated cytosine nucleotides, creating new methylation patterns

Despite extensive research, the precise mechanism by which DNMTs are directed to specific targets for *de novo* methylation is unknown. However, it is a highly complex process influenced by DNA sequence recognition by DNMTs, the surrounding chromatin structure, interaction with transcription factors, guidance of long non-coding RNA and other factors [72,73]. The original model with strict grouping between *de novo* and maintenance transferases are now being questioned as it has been shown that different DNMTs cooperate in both *de novo* methylation and maintaining an already established methylation pattern [71,74].

DNA methylation is both a stable and a dynamic process. Dynamics is seen during fertilization of the egg when both the oocyte and the spermatocyte methylation pattern are erased except for some repeat sequences and some imprinted regions [70,75]. This highly dynamic process is only observed during initial stages of early development, thereafter smaller but significant changes take place during physiological processes like differentiation and aging [76-78].

Demethylation of CpG sites has attracted much attention over the last decade. Demethylation can come about by different mechanisms. Passive demethylation is simply when DNA replication takes place in the absence of maintenance methylation systems i.e. DNMT1 (figure 6a).

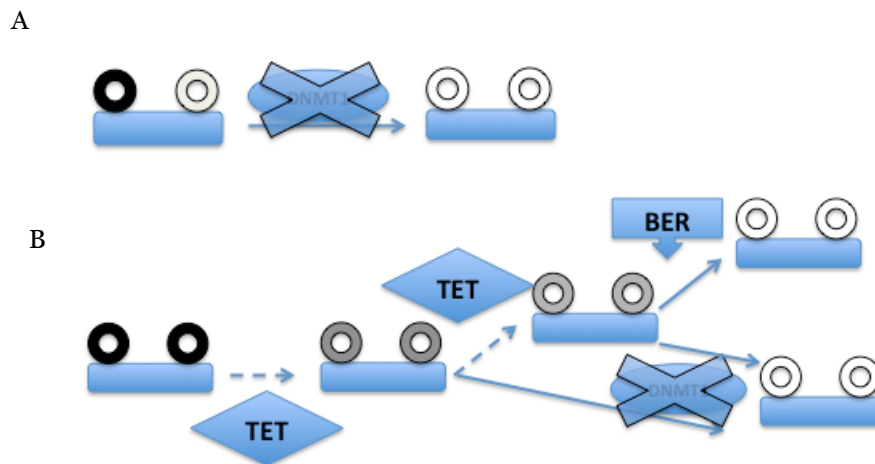


Figure 6.

A. DNA replication in the absence of DNA methyltransferases resulting in loss of the existing DNA methylation pattern on newly synthesized DNA strands, a process known as passive DNA de-methylation.

B. Enzymatic serial oxidation of 5mC mediated by TET. De-methylation can occur through passive replication dependent demethylation (DNMT1 is unable to recognize oxidized 5mC) or active demethylation through BER. Black circles represent 5mC, grey circles represent oxidized 5mC, BER base excision repair .

Active demethylation is a more complex process mediated through a family of enzymes named ten eleven translocated (TET), first discovered as a translocation partner of KMT2A [79]. There are 3 members of the TET

family TET1-3: they can convert 5-methylcytosine (5mC) through successive oxidation to (i) 5-hydroxymethylcytosine (5hmC), (ii) 5-formylcytosine (5fC), and (iii) 5-carboxylcytosine (5caC). During this process all three forms of oxidized cytosines can be passively demethylated during replication, since oxidized forms of 5mC are not recognized by DNMT1. Furthermore, 5fC and 5caC can be removed by base excision repair (BER) to generate a new unmethylated cytosine (figure 6b) [80,81]. But as always in biology, this seemingly straightforward process is complicated since it has become clear that the intermediate products exert biological functions of their own [82].

DNA methylation in the genome

CpG dinucleotides have a lower than expected occurrence in the genome, this is because of spontaneous deamination of 5mC to thymidine (T). C to T transition is also the most common mutation in cancer cells [83]. CpGs are unevenly distributed over the genome. The majority (70-80%) of CpG sites are scattered over the genome, enriched over repetitive regions and constitutively hypermethylated. In contrast 60-70% of all annotated genes have a much higher than expected CpG content in their promoter regions, clustered in so-called CpG islands (CGI)[84]. CGIs are enriched in proximity to transcription start sites (TSS) and usually unmethylated [83,85,86]. These differences indicate that both methylation state and location of CpG sites are of importance.

The widespread hypermethylated CpG-sites throughout the genome are important for genomic stability. For example, mutations in the catalytic domain of DNMT3b as seen in ICF-syndrome (immunodeficiency, centromeric region instability and facial anomalies syndrome) lead to loss of 5mC in centromeric repeats. ICF-cells display signs of chromosomal instability like anaphase bridges [86].

Most research however has been focused on DNA methylation at promoters as regulator of transcription. CGI rich promoters have different architecture compared to promoters with low CpG content. CGI rich promoters usually lack TATA-boxes and are enriched for general transcription factors like SP1 (Sp 1 transcription factor), and are nucleosome depleted around transcription start site. RNA polymerase II (RNAPolII) are bound to almost all promoter CGIs in embryonic stem cells even in non transcribed genes, indicating that unmethylated CGIs are transcriptionally permissive [83,85]. Hypermethylation of CGIs in promoter regions are usually linked to long term silencing, for example in x-chromosome inactivation and imprinting (a phenomenon where parent of origin specific expression is controlled through

inactivation of the other allele)[87]. Promoter hypermethylation is probably a late event in gene regulation, locking down an already silenced gene [88]. Considering that only 10% of CGI promoters are methylated in healthy differentiated cells, and most cells only express about 50% of all genes, there is not a clear cut relationship between methylation and expression [89].

Further down the gene, in the gene body, an even more complex pattern exists between DNA methylation and gene expression. Gene bodies have a relatively low CpG content and most are hypermethylated. Cytosine methylation is not associated with transcriptional silencing but instead with active transcription [90]. There are also CGIs within gene bodies, usually in an unmethylated state. Occasionally, when CGIs of gene bodies become hypermethylated they coincide with several features of non-active chromatin like H3K9me3. Despite this, gene transcription still occurs. The fact that promoter methylation is associated with gene repression and that DNA methylation of gene bodies does not show the same inverse correlation indicate that 5Cm can block initiation of transcription but has little effect on transcriptional elongation [83].

Gene transcription can also be regulated by enhancer elements. They are located outside promoter region at various distances up to mega bases from the promoter and act as regulators of transcription. Gene regulations by enhancers are usually mediated through recruitment of transcription factors at enhancer elements that then interact with the promoter region. Enhancers are important to control tissue specific expression and chromatin signature of enhancers is correlated to its activity [91]. The exact role of CpG methylation at enhancers is still not known. Some studies however indicate that DNA methylation can modulate transcription factors binding at enhancers and thereby affect enhancer activity [92].

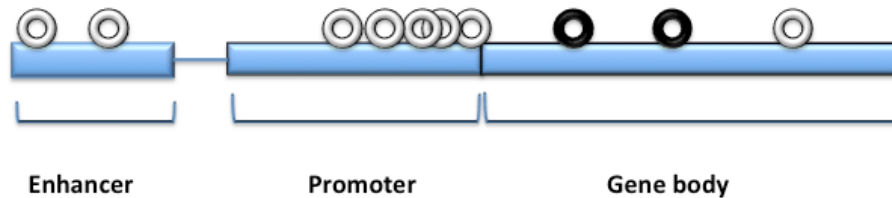


Figure 7.

Graphic overview over CpG distribution in a gene. Circles represent CpG sites, unfilled circles represent unmethylated CpG sites and filled circles represent methylated CpG

DNA methylation in the hematopoietic system

The stepwise differentiation from multipotent hematopoietic stem cells to mature hematopoietic cells from is a highly complex process, which at least in part is mediated by DNA methylation. DNMT1 deficient mice show skewing of lineage commitment, favoring myeloid differentiation [93]. Mice deficient for *de novo* methyltransferases DNMT3a and b on the other hand exhibit differentiation block and an increase in HSC renewal [94]. Manipulation of TET family members also disrupts normal hematopoiesis, but on the contrary to DNMT1, TET1 loss favors B-cell development [95]. In humans, DNMT3b mutation syndrome ICF includes immunodeficiency based on lymphopoietic defects mainly in the B-cell compartment [96].

In normal B-cell development, dynamic DNA methylation changes take place as maturation proceeds from an uncommitted progenitor to a mature B-cell. Whole genome bisulfite sequencing of B-cells in different maturation stages have revealed distinct differences in their methylation pattern. Initially, between HSC and preB-cell stage some demethylation was observed, mostly in enhancer regions clearly enriched for binding sites of B-cell specific transcription factors. Later in development, as the B-cell moves along the maturation axis from a naïve B-cell to an antigen experienced germinal centered B-cell and the a mature B-cell, dramatic demethylation of primarily repeat structures and heterochromatin regions are observed. Significant hypermethylation is a rather late step in B-cell development and is mostly seen in plasma cells and predominantly affect PRC silenced genes [78]. The DNA methylation pattern established in the previous differentiation step is retained and passed on through differentiation, pointing towards a developmental epigenetic memory.

In T-cell development, the initial step from progenitor cells with both myeloid and lymphoid potential to a T-cell committed progenitor (CD34⁺/CD1a⁺ to CD34⁺/CD1a⁻) is associated with *de novo* methylation. This gain in methylation is predominantly localized to CGIs but has little effect on gene expression. During later developmental stages demethylation is the dominating feature. Demethylation is also more strongly correlated to gene expression. Of interest is that TCR maturation appears to be controlled in part by methylation. Important TCR signaling genes undergo concomitant demethylation and expressional up regulation. As in B-cell development DNA methylation changes accumulated throughout development [97].

These data clearly indicate that DNA methylation pattern changes with differentiation of lymphocytes. DNA methylation changes also coincide with other physiological process like aging. Aging is associated with hypermethylation of PRC repressed genes and global hypomethylation in hematopoietic cells [98].

Epigenetics in leukemia

As mentioned earlier, DNA methylation by itself can cause mutations through spontaneous 5mC deamination. Vice versa, mutations can also cause epigenetic changes. Sequencing studies has shown that mutations in epigenetic regulators are common in hematopoietic malignancies. DNMT1a and TET2 are frequently mutated in adult AML and T-cell lymphomas and both genotypes display distinct DNA methylation patterns [99-101]. Mutations in DNMT3A have also been detected in about 20% of adult T-ALL cases, and associate with immature phenotype and poor prognosis [102,103].

Regarding pediatric ALL, the spectrum of mutations in epigenetic modifiers differ between different subgroups. CREB binding protein (CREBBP) has histone acetylation activity and is an important regulator in glucocorticoid response. *CREBBP* mutations are enriched in relapsed BCP-ALL especially among hyperdiploid cases [104]. Mutations in epigenetic regulators are even more common in T-ALL. All essential components of PRC2 are known to be mutated in T-ALL, and most frequently in ETP-ALL [105]. Mutations in the PRC2 complex are usually inactivating mutations, placing the PRC2 complex as a tumor suppressor but the exact mechanism through which PRC2 inactivation contributes to T-ALL development is still not known. Surprisingly, considering the tumor suppressive role of PRC2, inactivating mutations in the *KDM6A* (Lysine demethylase 6A) gene has been detected in about 5% of all T-ALL patients. *KDM6A* is a specific histone H3K27 demethylase and thereby exert the exact opposite function of PRC2. How

this paradox can be explained is not clear, but a possible explanation is that they exert their effect in different genomic regions [106,107].

Deregulation of chromatin modifiers can also occur through translocations. KMT2A activate transcription through its H3K4 methyltransferase activity and can also enhance transcription via histone H4 acetylation. In addition, KMT2A can also repress transcription as it interacts with PRC complexes[67]. In the translocated setting, the repressive function of KMT2A is lost and activation is further enhanced through interaction with DOT1L, responsible for H3K79 methylation [108]. KMT2A-r also protect Homeobox A9 (HOXA9) from *de novo* methylation and repression. HOXA9 is an important oncogene of KMT2A-r leukemia [35,109]. However, KMT2A-r does not only affect methylation status of HOXA9, but also influence global methylation pattern, a phenomenon KMT2A-r cases share with most other ALL subgroups. Several studies have examined the methylation landscape in ALL. A common denominator in all studies is that the genetic background of the leukemia is reflected in the methylation pattern [110-114]. Nordlund et al recently developed a cytogenetic/immunophenotypic classifier based on methylation pattern. Based on the methylation pattern of 246 CpG sites it was possible to accurately classify leukemic samples to specific cytogenetic subgroups [111].

Most studies examining methylation pattern in ALL use array-based approaches and since arrays usually are CGI and promoter focused most studies are focused on these regions. In general, *de novo* methylation in ALL is located in CGIs both within and outside promoters and DNase hypersensitive regions (like nucleosome depleted regions around TSS). Regions hypermethylated in ALL are also enriched for sites with bivalent chromatin markers, and therefore enriched for PRC binding sites [110,111]. Demethylation of repetitive sequences genome wide is a hallmark of cancer cells and linked to genomic instability as described above, but these repetitive regions are poorly represented in methylation arrays. One study has used a whole genome bisulfite sequencing (WBS) approach to scrutinize the genomic distribution of DNA methylation in two patients, one with hyperdiploid ALL and one with ETV6-RUNX1. Demethylation was observed at repeat structures and in blocks associated with Lamina associated domains in the nuclear periphery. A similar trend is seen in array data from larger materials. Even though demethylation is seen in ALL (especially in hyperdiploid cases), global demethylation is less widespread in ALL compared to solid tumors [110].

Some studies have investigated methylation patterns in relapsed ALL. BCP-ALL appear to accumulate additional hypermethylations in the same

locations as at diagnosis i.e. CGI and bivalent marked promoters [111,115]. In relapsed T-ALL hypomethylation of gene-associated CpGs appear more common than hypermethylation [116].

Telomere length as a prognostic marker in leukemia

In addition to genetic aberrations, telomere length has also been implicated as a prognostic marker in hematological malignancies [117]. The telomeres are tandem (TTAGGG)_n DNA repeats at the end of chromosomes which are associated with specific telomere binding proteins. These telomeric structures have several essential functions. For example, they protect the chromosome ends from being recognized as chromosome breaks, and serve as a buffer of non-coding DNA. The DNA polymerase is unable to replicate the very distal part of the lagging strand and therefore some bases are lost each cell division. So without functional telomeres, there would be severe chromosomal instability and loss of genetic information in each cell cycle. Because critically short telomeres induces senescence, an irreversible growth arrested state, most somatic cells have limited replicative capacity [118]. Stem cells and cells with high replicative needs (i.e. activated lymphocytes) express the enzyme telomerase. Telomerase is a reverse transcriptase enzyme complex, including the catalytic subunit- telomerase reverse transcriptase (hTERT) and the telomerase RNA template (hTR) subunit, that bind specifically to telomeres and elongate the chromosome ends by adding telomeric repeats. hTERT is also expressed in 90% of all cancer cells and is a hallmark of cancer and also a tempting target for cancer therapy [119]. In this thesis, telomere length and methylation of the hTERT promoter were analyzed in relation to cytogenetic aberrations and prognosis in ALL.

Aims

General aim

During the last decade epigenetics have developed into an important field of research and its importance in health and disease has become apparent. The general aim of this thesis was to find new epigenetic prognostic markers in childhood acute lymphoblastic leukemia. Our focus has been on subgroups of ALL patients where we today lack reliable stratifying markers.

Specific aims

Paper I

The objective in paper I was to investigate the biological and clinical significance of telomere length and promoter methylation of the human telomerase reverse transcriptase gene in childhood acute lymphoblastic leukemia.

Paper II

In paper II we wanted to characterize the DNA methylation pattern in diagnostic samples from pediatric T-ALL patients, using genome-wide promoter focused methylation arrays to find new prognostic markers for treatment outcome.

Paper III

In paper III we aimed to validate the prognostic value of the CIMP profile identified in Paper II in the most recent Nordic treatment protocol for T-ALL. Specifically, we wanted to investigate whether CIMP could add prognostic value to MRD.

Paper IV

The aim of paper IV was to investigate the prognostic value of CIMP in a large cohort of BCP-ALL patients.

Materials and methods

Patients

In paper I, II and III mononuclear cells were separated from diagnostic bone marrow aspirates or peripheral blood at the time of diagnosis. All patients were between 0 to 17,9 years of age when diagnosed with ALL. Diagnosis was based on morphology, immunophenotyping, and cytogenetic analysis. The following structural cytogenetic aberrations were established by G-band karyotyping, fluorescence in situ hybridization, and/or PCR : ETV6/RUNX1, iAMP21, TCF3/PBX1, BCR/ABL, KDM2a and dic9;20.

Paper I included 169 patients diagnosed with ALL between 1988 and 2006 at the Swedish regional pediatric oncologic centers in Umeå, Uppsala or Stockholm, Sweden. Remission samples were obtained at cessation of therapy from 40 of these 169 patients. Only patients treated according to NOPHO ALL protocols 1992 or 2000 were included in survival analysis. BCP-ALL patients with no high-risk features according to protocol were defined as low-risk in this study (n = 102). Median follow-up time in first complete remission was 58 months (range, 4–242 months). This study was approved by the ethical committee at Umeå University.

Paper II. Between January 1, 1992 and June 30, 2008, 75 infants, children, and adolescents <18 years were diagnosed with T-ALL at the Swedish regional pediatric oncologic centers in Lund, Göteborg, Linköping and Umeå. 46 patients were treated according to the NOPHO ALL 1992 protocol and 29 patients according to the NOPHO ALL 2000 protocol. DNA from diagnostic samples where available from 43 out of 75 patients and thus analyzed by methylation arrays. The remaining 32 patients served as a control group to evaluate the representability of the patients available for DNA methylation analysis. This study was approved by the ethical committee at Umeå University.

Paper III. Between July 2008 and March 2013, 113 children (age <18 years) were diagnosed in the Nordic countries with T-ALL and treated according to the common NOPHO ALL 2008 protocol. Sixty-five diagnostic bone marrow/peripheral blood samples were available in the NOPHO leukemia biobank in Uppsala, Sweden, and were analyzed for methylation status. MRD was monitored by PCR and/or flow cytometry. PCR analysis of clonal gene rearrangements was recommended for MRD quantification in T-cell

ALL and such MRD data were used when available (n = 41). However, if no PCR-based MRD was available, flow cytometric quantification of MRD [14] was used (n = 20). Four cases lacked both PCR and flow MRD data and were excluded from the survival analyses that included MRD. The regional and/or national ethics committees approved the study, and the patients and/or their guardians provided informed consent in accordance with the Declaration of Helsinki.

Paper IV included 601 pediatric patients aged 1-18 years diagnosed with B-cell precursor ALL (BCP-ALL) between years 1996 and 2008 in the Nordic countries and treated according to the common NOPHO ALL 1992 and 2000 protocols as described earlier. Clinical follow up data was extracted from the NOPHO leukemia registry in June 2016 and the mean follow up time for patients was 115 months (range 0-221). The regional and/or national ethics committees approved the study, and the patients and/or their guardians provided informed consent in accordance with the Declaration of Helsinki.

DNA methylation

Several different techniques are available for DNA methylation analyzes. Most techniques rely on bisulfite treatment of DNA since no PCR primer or probes can discriminate between methylated and unmethylated cytosine (C) residues. During bisulfite treatment methylated C remains C whereas unmethylated cytosines are converted into uracil (U). During the next PCR step U is converted to thymidine (T).

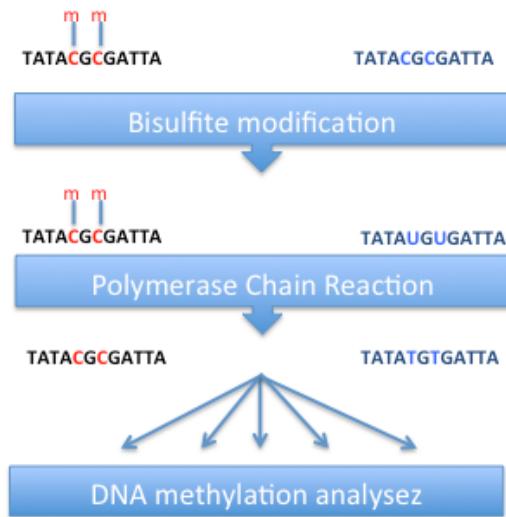


Figure 8. Illustration of DNA bisulfite treatment. Methylated cytosines are red and unmethylated cytosines are blue. During bisulfite treatment unmethylated cytosines are converted into uracil. Through PCR uracil is replaced by thymidine.

In our studies DNA was bisulfite treated with the EZ DNA methylation kit (Zymo Research) according to manufacturer's recommendations.

Methy-Light

In paper I diagnostic DNA samples from ALL patients, were bisulfite-treated and analyzed for methylation status of the hTERT promoter using quantitative MethyLight, a fluorescence-based real-time PCR method [120]. MssI (New England Biolab)-treated (generating genome wide methylation) leukocyte DNA, was used as reference sample and Alu repetitive sequences as control for DNA content in methylation-independent PCR reactions [121]. The analyzed hTERT promoter region corresponds to positions 10996 to 11111 of hTERT sequence AF128893 (Gene Bank). To determine if a sample was considered methylated or not, the following formula was used:

$$Met(\%) = 100 \cdot \frac{Met_{sample}/Control_{sample}}{Met_{MssI-ref}/Control_{MssI-ref}} \quad (1)$$

where Met (%) is the percent methylated of the sample, Met_{sample} the methylated reaction of the sample, $Met_{MssI-ref}$ the methylated reaction of the MssI-treated reference sample, $Control_{sample}$ the control reaction of the sample and $Control_{MssI-ref}$ the control reaction of the MssI-treated reference sample. Samples were considered positive for methylation if percent methylated reference >10 .

High Resolution Melting curve analysis

High-resolution melting (HRM) assays were designed for a selection of six genes in the CIMP panel, representing CpG sites with distinct differences in methylation levels between CIMP subgroups. HRM was conducted by melting PCR products from 60 to 90°C, rising by 0.1° each step. A standard curve was prepared by mixing 100% methylated DNA (MssI treated) in different ratios with DNA from mitogen (wheat germ agglutinin) stimulated primary lymphoblast T-cell (theoretically 0% methylated mononuclear cells). The methylation level of each gene region covered in the HRM assay was estimated in relation to the standard curve, and a mean methylation level (%) of the six-gene HRM panel was calculated.

DNA Methylation Arrays

In paper I, genome-wide promoter methylation profiling was performed using the Illumina Infinium HumanMeth27K BeadArray (Illumina, San Diego, CA, USA). These arrays generate data for 27578 CpG nucleotides, corresponding to 14473 individual gene promoter regions. CpG nucleotides are preferentially located within CpG islands as defined by Takai and Jones relaxed criteria [84]. The methylation assay was performed according to the manufacturers instructions. Briefly, bisulfite converted DNA was fragmented and amplified with supplied reagents. Processed DNA was hybridized to the arrays. Each CpG site on the array was represented by two site-specific probes, one designed for the methylated allele and another for the unmethylated allele. Single-base extension of the probes incorporates labeled nucleotides, which subsequently was stained with a fluorescence reagent. The methylation level for a locus was determined by calculating the ratio (β value) of the fluorescent signals from the methylated (M) vs. unmethylated (U) sites where:

$$\beta = \text{Max}(M,o) / (\text{Max}(M,o) + \text{Max}(U,o) + 100) \quad (2)$$

ranging in theory from 0, corresponding to completely unmethylated, to 1, representing fully methylated DNA.

In paper III and IV diagnostic T-ALL and BCP-ALL samples were analyzed by the HumMeth450K methylation array (Illumina) an upgraded version of the 27k array, covering 485,577 CpG sites. Bisulfite conversion and array analysis including preprocessing and normalization was performed. Normalization is necessary for the 450K array since two different chemistries are used. All CpG sites on the 27K array as described above, were

represented by two bead types (methylated and unmethylated), both beads utilize the same color channel for detection. For the 450k array a new type of chemistry was adopted in which a single bead type is used but two different color channels are applied for discrimination between methylated and unmethylated CpG sites. The fluorescence intensities were extracted using the Methylation Module (1.9.0) in the Genome Studio software (V2011.1). CpG probes that align to multiple loci in the genome or were located less than 3 bp from a known SNP were excluded. The methylation level (β value) of each CpG site ranged from 0 (no methylation) to 1 (complete methylation) as described above. The quality of each individual array was evaluated with built-in controls. Two replicate samples were included to assess inter-assay reproducibility ($R^2 = 0.97-0.99$).

Defining CIMP

Paper II

In paper II the first CIMP profile was generated based on HumMeth27k array data. All dataprocessing was performed in BeadStudio (Illumina inc.), no normalization was performed (as recommended by Illumina technical support). Initially, probes with failed detection in one or more arrays were excluded from further analyzes and to avoid sex bias, all CpGs located on the X or Y chromosomes were excluded. Analysis was subsequently restricted to the remaining 26436 CpG sites. In order to identify the most variably methylated probes within the T-ALL group a standard deviation ≥ 0.30 of the β -value across the samples was used as cut-off. Euclidian distances were used for cluster generation.

Paper III

In paper III samples were analyzed by HumMeth 450K arrays, and CIMP classified according to the above described CIMP panel with slight modifications (normalized, filtered for CpG probes that align to multiple genomic locations or within 3 bp from a SNP, and like in paper II CpG sites located on X and Y chromosomes were excluded). A total of 1,293 CpG sites of the original 1347 CpG sites in the CIMP panel were represented on the HumMeth450K arrays and remained after filtration. In order to standardize CIMP classification independent of clustering, a cut-off level for “percentage of methylated CpGs within the CIMP panel” was defined for the CIMP subgroups. Diagnostic T-ALL samples with $>40\%$ methylated CpG sites (each CpG site was considered methylated if the beta value was >0.4) were classified as CIMP positive, whereas samples with $\leq 40\%$ methylated sites

were denoted CIMP negative. The cut-off was set to reflect the previously identified clusters to be the most discriminating with respect to prognosis.

Paper IV

HumMeth450K array data from 601 diagnostic BCP samples previously analysed by Nordlund et al [111](Nordlund et al Genome biology 2013) were downloaded from Gene Expression Omnibus (GEO) and preprocessed. Samples were classified according to our previously defined CIMP panel. The CIMP subgroups were defined based on percentage of methylated CpG sites (methylated CpG site defined as a β -value > 0.4) within the 1293 CpG site CIMP panel. Since approximately 1/3 of the CpG sites in the CIMP panel seemed to be methylated only in T-ALL samples the CIMP percentage separating CIMP low and CIMP high was empirically adjusted for BCP-ALL. Samples with $\leq 25\%$ methylated CpG sites were classified as CIMP low, and samples $> 25\%$ methylated CpG sites were classified as CIMP high.

Gene expression

Total RNA was isolated from 17 frozen bone marrow samples and RNA was quality checked as described in paper II. RNA from each sample was amplified and biotinylated according manufacturers instructions. Biotinylated cRNA were hybridized to human HT12 Illumina Beadchip gene expression arrays (Illumina inc.) according to the manufacturer's protocol. The arrays were scanned using the Illumina BeadArray Reader (Illumina). The Illumina GenomeStudio software was used for data analysis and normalization by the quantile algorithm. Genes with signals below background level were excluded from the analysis, and differentially expressed genes (2-fold) between CIMP subgroups were identified by fold change calculations.

Telomere length

Telomere length was measured by qPCR, as previously described [122]. Human beta-globin (HBG) was used as the single-copy gene. Briefly, the ratio of telomere repeat copy number to single-copy gene number was determined. Relative telomere length (RTL) were generated by dividing

sample values with the value of a reference cell line DNA (CCRF-CEM), included in all runs and loaded in the same amount as the unknown samples.

Statistical analysis

Methylation data analysis was carried out in the R environment (v2.15.0) or GenomeStudio as described in the respective study. The Statistical Package for the Social Sciences (SPSS Inc., Chicago, IL) software was used for the statistical analyses. Chi-square and the Mann-Whitney U tests were used to determine whether differences among subgroups existed for discrete and continuous variables, respectively. Estimates of event free survival (EFS), overall survival (OS), and cumulative incidence of relapse (CIR) are given at 5 years in paper I, II, IV and at 3 years in paper III and were calculated using the Kaplan-Meier method and subgroups compared using the log rank test. The significance limit for two-sided p-values was set to <0.05 in all tests.

Results and Discussion

Telomere length and hTERT promoter methylation in ALL (Paper I)

In the first study we investigated two aspects of telomere biology in childhood-ALL, telomere length and promoter methylation status of the catalytic component (hTERT) of the telomerase complex. Previous studies from our lab have shown that telomere length has prognostic relevance in adult leukemia, e.g. chronic lymphocytic leukemia [122,123]. In paper I, telomeres length was analyzed by Tel-PCR in 169 diagnostic ALL bone marrow samples and 40 morphologically normal bone marrow samples after cessation of therapy. Patients were diagnosed in Umeå, Uppsala or Stockholm between 1988 and 2006.

Telomeres were generally shorter in leukemic cells than in post treatment bone marrow cells, a feature ALL share with most other hematological malignancies [117]. T-cell leukemic samples had a tendency towards longer telomeres than BCP-ALL at diagnosis but the difference did not reach significance. Among BCP-samples, BCR-ABL1 cases had the shortest telomeres and had significantly shorter telomeres than both hyperdiploid and ETV6-RUNX1 leukemias ($p=0,001$ and $p=0,002$).

To relate telomere length to outcome, BCP-ALL patients were divided into two risk groups, high and low risk based on treatment intensity. All patients included in survival analysis were treated according to the NOPHO ALL 1992 and 2000 protocols. These protocols had similar backbones and patients treated by these protocols have comparable outcomes [32]. BCP-ALL patients with no high-risk features according to the protocol were defined as low-risk in this study, the rest were considered high risk. Among low risk patients, a telomere length over the mean value for the entire cohort was associated with inferior outcome (Figure 9).

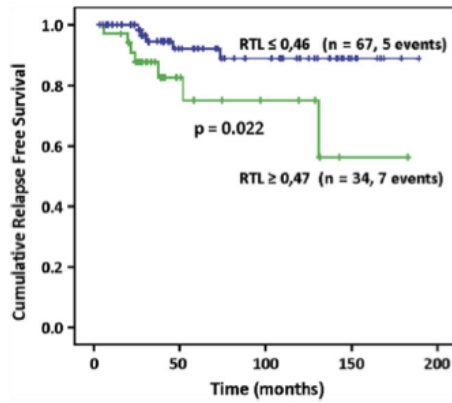


Figure 9.

Relapse free survival in low risk BCP patients in relation to telomere length. Patients with longer telomeres (above mean telomere length (RTL=0,47) for the entire cohort) had inferior outcome (Borssen et al. Exp Hematol 2011).

Among high risk patients no difference in outcome was observed with regard to telomere length. The reason for this discrepancy could be explained by the short telomeres in BCR-ABL1, a subgroup with at the time very poor prognosis. Even though there are no other studies on telomere length in BCR-ABL1 positive ALL, several studies have analyzed telomere length in chronic myeloid leukemia (CML). Collectively, these studies have shown that CML cells have short telomeres and pronounced telomere shortening is observed in transition from chronic phase to blast crisis [124,125]. Short telomeres might even predict for inferior response to TKI treatment in CML [126].

The second aim of paper I was to characterize the methylation status of the core promoter of hTERT (147 samples were available for analyzes). In post treatment samples the analyzed region appeared completely unmethylated in line with other studies [127]. The frequency of hTERT hypermethylation differed between different ALL subgroups. In T-ALL and BCP-ALL patients positive for ETV6-RUNX1 translocation, the majority of samples were hypermethylated (65 and 63%, respectively), whereas all other subgroups had no or very few hypermethylated cases.

The impact of DNA methylation on hTERT expression is complex and appear to be cell type specific [128]. We found higher levels of hTERT mRNA in BCP-ALL samples with unmethylated hTERT promoter when a small

subgroup with available RNA was analyzed (n=18). Methylation status had no obvious impact on outcome in BCP-ALL. But we made an interesting observation in the T-ALL patients in this study. Among T-ALL patients who had an unmethylated promoter more than half of the patients relapsed in contrast to patients with hypermethylated hTERT promoter among which only two out of twelve patients relapsed.

CIMP profile as a prognostic marker in T-ALL (Paper II&III)

Based on our observation in paper I we decided to investigate if DNA methylation could serve as a potential prognostic marker in T-ALL. DNA methylation as a prognostic marker has been widely studied in several types of malignancies. In 1999 Toyota et al. described a subset of colorectal cancer samples with distinct histology and hypermethylation of certain CpG sites, they also introduced a term to describe this epitype, CpG island methylator phenotype (CIMP). Later it was shown that colorectal CIMP tumors were associated with B-Raf proto-oncogene (BRAF) mutations and micro satellite instability (MSI) [120,129]. The term CIMP has later been used in other malignancies including leukemia. Although there is no universal definition of CIMP, it rather refers to a group of samples with a set of hypermethylated CpG sites [130].

In paper II we used the 27K Illumina methylation array, which interrogates over 27 000 individual CpG sites, mostly located in promoter regions of 14473 genes. We analyzed a cohort of Swedish T-ALL patients treated according the NOPHO 1992 and NOPHO 2000 protocols. In both protocols T-ALL was classified as HR. Diagnostic samples from 43 patients were available for analysis. Considering the small sample size, it was important to investigate if our available samples were representative regarding clinical outcome. We compared our cohort to the remaining patients treated according to the same protocol during the same time period at the same pediatric oncology centers (Umeå, Linköping, Gothenburg and Lund) where no diagnostic material was available for methylation analysis (n=32). No difference in outcome, sex, age or WBC at diagnosis were found between the two groups.

When analyzing data from the methylation arrays we first removed all probes detecting CpG sites on the X and Y chromosome to avoid sex bias, an important step considering that DNA methylation is involved in X-chromosome inactivation in females [131]. CpG sites with failed detection were also removed. After these steps 26436 CpG sites remained for analysis.

In order to find probes that were differently methylated within the T-ALL group, CpG sites where the β -value had a standard deviation $\geq 0,3$ over all analyzed samples were selected. From these criteria we were left with a CIMP-panel of 1347 CpG sites associated to 1038 genes.

Hierarchical clustering using Euclidian distance separated T-ALL samples into three groups. The hypermethylated and the intermediately methylated groups were co-analyzed as CIMP high (n=21). The third group had a methylation pattern similar to normal T-cells and normal bone marrow over the 1347 genes. This group was referred to as CIMP low (n=22).

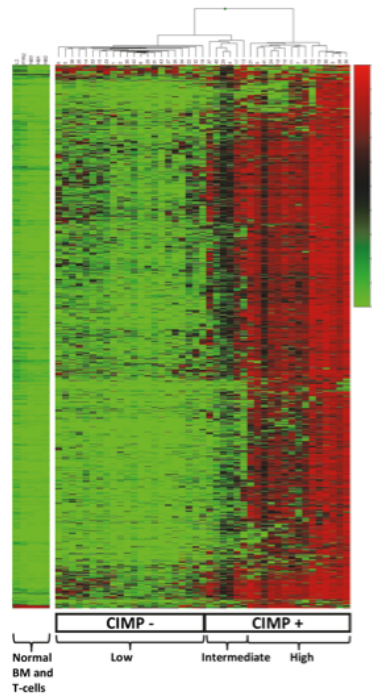


Figure 10.

Heatmap over the 1347 most variable CpG sites in T-ALL (standard deviation ≥ 0.30). High methylation levels are shown in red and low levels in green. The T-ALL methylation subtypes are marked in the figure as low, intermediate and high, and with the merged classification as CIMP- and CIMP+. The methylation status of three normal bone marrow samples and two stimulated primary T-cell lines are shown next to the hierarchical cluster (Borssén et al PLoS One 2012).

Among CIMP high patients, three out of 21 patients relapsed (14%) compared to 12 out of 22 CIMP low patients (55%). Both OS (86% vs. 45%, $p = 0.02$) and EFS (86% vs. 36%, $p = 0.001$) were significantly better in the CIMP high group. Since there was a lack of reliable prognostic/stratifying markers in T-ALL these results were very encouraging. However, patients in paper II were treated with older protocols, and probably had poorer outcome than T-ALL patients treated with more contemporary protocols. Furthermore, no minimal residual disease data were available for these patients and therefore we were unable to conclude the current prognostic value of CIMP in T-ALL.

In paper III the aim was to verify our results in an independent cohort and to investigate the prognostic value of CIMP in the current NOPHO ALL 2008 protocol. Diagnostic bone marrow samples or peripheral blood were retrieved from the NOPHO biobank in Uppsala. Samples from 65 out of 113 patients diagnosed with T-ALL between 2008 and 2013 in the Nordic countries were available for analysis (57,5%).

Outcome for patients included for analysis of DNA methylation did not differ to those not included. A 3-year OS was 79% for both groups.

DNA methylation was analyzed using the 450K Illumina methylation arrays, an updated version of the 27k array covering 485 577 CpG sites. CIMP-classification was based on the same panel as in paper II with some modifications. A few of the 1347 CpG were not represented on the 450K array and hence not possible to include in the CIMP panel. CpG sites located within 3bp from SNPs were also removed since it was at this point known that SNP variations can influence CpG methylation status. With these modifications 1293 CpG sites remained in the CIMP panel. To get a more stringent and reproducible CIMP classification, that did not rely on cluster analysis, we defined specific criteria to classify CIMP status. Cut off for CIMP high was set to more than 40% methylated (methylated defined as $\beta > 0,4$) CpG sites in the CIMP panel, and patients with $\leq 40\%$ were classified as CIMP low.

Twenty-five patients were classified as CIMP low and 40 as CIMP high. In general, OS_{3y} appear to have improved with the NOPHO ALL 2008 protocol, now reaching 79%, but it's important to remember that follow up time still was short. Despite the generally improved survival, CIMP status still predicted outcome. CIMP low patients had cumulative incidence of relapse (CIR) of 29% compared to only 6% among CIMP high patients.

In the NOPHO ALL 2008 protocol, T-ALL is primarily stratified based on MRD at day 29. If MRD is below 0,1% at day 29 patients are stratified for the IR arm, patients are otherwise allocated for HR treatment [19]. Despite the difference in survival between CIMP groups, the fraction of patients with MRD over 0,1% did not differ between groups. No relapses occurred among patients with MRD below 0,1% and all but two patients received IR treatment. By combining MRD status at the end of induction (day 29) and CIMP status we were able to separate MRD-based high risk patients into two groups with different clinical outcome. Among the 33 patients with MRD over 0,1% the distribution between CIMP high/CIMP low were 18/15. The

cumulative incidence of relapse after 3 years was 50% among CIMP low cases compared to 12% among CIMP high patients. The OS was 83% vs 45% respectively. It appears as the combination of MRD and CIMP data has synergistic effect on outcome prediction.

Integrated gene expression and DNA methylation (Paper II)

In paper II we also examined the correlation between promoter methylation and gene expression. We performed gene expression array analysis (Illumina HT-12) on 17 T-ALL samples. A significant (Spearman correlation $Rho = -0.260$, $p, 0.001$) inverse correlation was observed between promoter methylation and gene expression.

By comparing the CIMP profile with a list of differently expressed genes ($n=405$) defined as 2-fold up/down regulated between the CIMP groups, 39 genes were both differently methylated and differently expressed. The majority of these 39 genes showed a negative correlation between methylation and gene expression, i.e. high promoter methylation was associated with lower expression. One interesting finding was that TAL1 gene expression was negatively correlated ($r^2 = -0.611$) with TAL1 promoter methylation level. Only 47% of the differently methylated genes (DMGs) were expressed above background level in the gene expression array, and therefore evaluation of a possible influence of promoter methylation on gene expression was not feasible for all genes.

Enrichment analysis of CIMP profile

To identify potential genes and cellular processes that were connected to the CIMP profile we used two bioinformatic tools, GeneGO Metacore software and DAVID Functional Annotation Tool.

Our analysis of DMGs indicated an overrepresentation of genes associated with PRC 1 and 2 complexes. To further investigate this we compared previously published lists of PRC target genes identified in human embryonic stem cells and human embryonic fibroblasts [132,133] with our CIMP profile. A high proportion of genes (62%) in our list were identified as polycomb target genes, which were much higher than expected, indicating a preferential methylation of polycomb target genes. Even though the 27K array is slightly enriched for PRC target genes, there was a highly significant overrepresentation of polycomb target genes among the differently methylated CpG sites in T-ALL that could not be explained by the small array bias.

Focused analysis of cell processes and metabolic pathways revealed a large number of genes involved in ATP metabolism and cAMP signaling pathways to be differently methylated within the T-ALL group. The same pathways are also frequently hypermethylated in BCP-ALL, as shown in a study focused on HeH, ETV6-RUNX1 and B-other subtypes [134].

ETP and DNA methylation (Paper II&III)

In paper II no immunophenotype data was available to analyze whether ETP cases were overrepresented in the CIMP low group. ETP-ALL not only has a distinct immunophenotype but can also be identified by a distinct gene expression signature. We compared expression data between CIMP low and CIMP high T-ALL from 17 of our patients in paper II where RNA was available [135]. The ETP expression signature did not discriminate CIMP high from CIMP low individuals.

In paper III we were able to classify samples as immature or non-immature phenotype based on available immunophenotype data and ETP was defined according to Coustan-Smith et al [59]. No enrichment of ETP cases was found among CIMP low patients.

CIMP as a prognostic marker in relapsed BCP-ALL (paper IV)

Nordlund et al have previously analyzed a large cohort of BCP-ALL patients with the 450k Illumina methylation array. By applying the CIMP profile on the BCP-ALL data set we were able to analyze the prognostic relevance of the CIMP panel in BCP-ALL. We had to modify our definition of CIMP to some extent because of the number of differently methylated CpG sites was generally lower in BCP-ALL compared to T-ALL [111,112]. We defined two groups based on number of methylated CpG sites within our CIMP panel (methylated defined as a β -value >0.4) of 1293 CpG sites. Samples with $<25\%$ methylated CpG sites were denoted CIMP low, if more than 25% of CpG sites were methylated the sample was classified as CIMP high.

A total of 601 diagnostic BCP-samples were analyzed for CIMP status, infants (age under 12 months) were not included since they constitute a very distinct entity with poor outcome. 175 patients were classified as CIMP low and 426 as CIMP high. Different cytogenetic subgroups were not evenly distributed between CIMP-groups. CIMP low were enriched for BCR-ABL1 and TCF3-PBX1 positive cases and CIMP high were enriched for ETV6-RUNX1 cases. The over all survival was significantly better among CIMP high patients although no difference was found concerning cumulative incidence of relapse (CIR). Considering the skewed distribution of

cytogenetic aberrations between CIMP-groups a difference in outcome is not surprising. But the fact that CIMP predicted OS better than CIR indicate that CIMP could hold information about response to relapse treatment.

Among the 601 analyzed samples 137 patients experienced relapse as their primary event, 42 were CIMP low, and 95 CIMP high. The most important prognostic factor at relapse is time in complete remission, the earlier relapse the worse outcome. Patients with very early relapse defined as within 18 months of diagnosis had an OS of only 26% compared to 46% if relapse occurred between 18 months after diagnosis to six months post treatment, late relapse had the best outcome with 77% survival rate. By grouping patients by both time to relapse and CIMP status six groups were generated. CIMP status had no impact on late relapses but CIMP low cases had significantly worse outcome among both very early and early relapses (Figure 11).

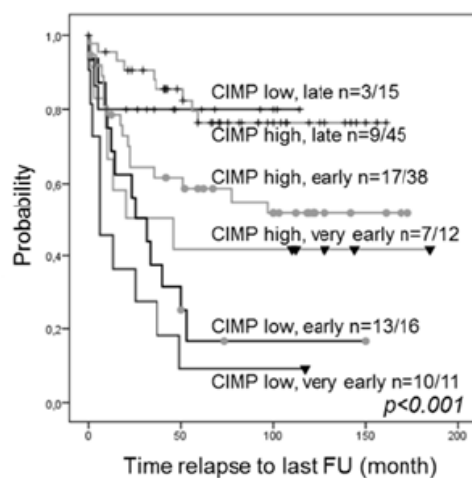


Figure 11. 137 relapsed BCP-ALL samples with combined CIMP status at diagnosis and time to relapse. Time to relapse was classified as: very early (<18 months from primary diagnosis), early (≥ 18 months from diagnosis to ≥ 6 months post treatment), and late (>6 months post treatment). Follow up time (month) from relapse to last follow up.

Even though genetic aberrations are not taken into account at relapse (as stratifying factor) some studies have shown a significant impact of cytogenetic background at relapse as well [62,136]. Because of the skewed distribution of cytogenetic subgroups between CIMP groups we analyzed risk groups separately. CIMP low had inferior outcome among HR patients ($p=0,016$) but not for non-HR patients where 45 of 64 CIMP high patients remained in complete remission compared to 12 out of 23 among CIMP low patients ($p=0,06$). The prognostic relevance of CIMP in predicting survival after relapse could not be explained by frequency of stem cell transplanted (SCT), since the proportion of SCT-treated individuals after relapse were similar in both CIMP groups.

General discussion

Modern treatment protocols have made tremendous progress in treatment of childhood leukemia and we can today offer curative treatment for most of our young patients. Success have at large been to offer increasingly tougher treatments. In fact, anyone who has ever visited a pediatric oncology ward knows that such treatments do not come without a cost. All children suffer from side effects from the treatment, one of which includes death. The aim of these studies was to find new potential prognostic markers to push the ability to pinpoint which patients need the toughest treatment but also as important to find subgroups that do not need intensified treatment.

DNA methylation as a prognostic marker really gained momentum in the late 90s when CIMP was discovered in colon cancer [129]. The concept has translated to most tumor types including leukemia [137].

Earlier studies mostly used quite small sets of genes, often inspired by the CIMP-panel in colon cancer [138-141]. The introduction of array-based techniques paved the way for more detailed analysis and the use of cluster analysis to define groups with distinct methylation patterns [142-145].

DNA methylation as a prognostic marker in ALL

Several studies have characterized DNA methylation pattern in childhood ALL [111,113,146,147]. Regarding T-ALL few studies have investigated the relation between treatment response and DNA methylation. We clearly show in two independent cohorts that T-ALL patients can be divided into two distinct DNA methylation subgroups and that a CIMP low profile is clearly associated with inferior outcome. This is in contrast to earlier publications. Kraszewska et al examined the methylation pattern of a gene panel in 61 T-ALL patients and found no prognostic significance of CIMP [148]. Some, but few, of the analyzed genes were overlapping with our profile, which could be one explanation why their results are different from ours. CCAAT/Enhancer Binding Protein Alpha (CEBPA) was included in both studies. In the Polish study CEBPA was never hypermethylated whereas four CpG sites localized to the CEBPA promoter were included in our CIMP profile and hypermethylated in most CIMP high cases. Terriou et al have also investigated the methylation status of CEBPA in T-ALL. They found a similar frequency of hypermethylation as we did, but did not specify any prognostic significance of DNA methylation [149]. From this we conclude that selection of regions and sites for CIMP profiling is of crucial importance.

Numerous studies have examined the prognostic value of DNA methylation in other hematological and solid malignancies. Regarding ALL, Sandoval et al. investigated DNA methylation in a small group of BCP-ALL patients and employed a similar approach as we did. Similar to our results CIMP high patients had better prognosis than CIMP low patients. In contrast to our study on BCP-ALL, their profile was associated with higher risk of relapse [150]. Other studies have failed to find methylation patterns predictive of relapse [113]. A general finding from several groups is that the genetic background i.e. cytogenetic subgroup and immunophenotype of the leukemia has a profound impact on DNA methylation pattern. This phenomenon is important to take into account when examining DNA methylation as a prognostic marker in ALL since both cytogenetics and immunophenotype influence outcome. When we applied our T-ALL generated CIMP profile on BCP-samples, an enrichment of high-risk cytogenetic subgroups was seen in CIMP low, which explain the differences in survival when all patients are analyzed and not only relapse cases. Thereby limiting the utility of CIMP classification at diagnosis in BCP-ALL. Little is known about DNA methylation and relapse. In BCP-ALL there appear to be a general gain in methylation from diagnosis to relapse but there are large individual variations [115]. However, it should be noted that we base our findings on diagnostic samples and not on samples at the time of relapse. The fact that CIMP at first diagnosis reflect response to the second treatment is unexpected and the underlying mechanism is not known.

Collectively, our data on both T-ALL and BCP-ALL, CIMP low phenotype seem to reflect a more aggressive leukemia with enrichment of other biological risk factors i.e. high WBC at diagnosis (in T-ALL), high-risk cytogenetics and very early relapse.

Driver or passenger?

As discussed earlier, DNA methylation can mediate silencing of tumor suppressors and thereby act as a driver in tumor development [4]. However numerous studies have shown that the correlation between gene expression and DNA methylation is rather weak and most methylated genes (up to 90%) are already silenced when they become methylated [151]. The lack of a clear-cut correlation between gene expression and promoter methylation has paved the way for alternative explanations regarding DNA methylation changes in cancer.

Still some genes appear to be regulated at least in part by *de novo* methylation. CDKN2A is a classical tumor suppressor and as described above frequently mutated in both T-ALL and BCP-ALL [152,153]. Methylation of CDKN2A could serve as an alternative to deletion when it comes to leukemia pathogenesis, and there are animal models that support this hypothesis [154]. We found another interesting example in TAL1, even though not included in the CIMP panel, TAL1 showed a strong negative correlation to DNA methylation [155]. Since TAL1 is an oncogene and not a tumor suppressor, lack of methylation at the TAL1 promoter would facilitate TAL1 expression and promote oncogenesis. Other studies have shown that TAL1 overexpression (TAL1 overexpression as a result of STIL-TAL1 translocation) is associated to high WBC at diagnosis, in analogy to CIMP low T-ALL. However, TAL1 overexpression does not appear to be a negative prognostic marker [156].

A second mechanism for deregulated DNA methylation has been well established in AML. Mutations in genes involved in DNA methylation, like DNMT3A, cause a distinct methylation phenotype with pronounced hypomethylation and is associated with inferior outcome [157](Ying et al. In a mouse model system, enforced expression of DNMT3B delayed malignant transformation, and upregulation of DNMT3B target genes are associated with inferior outcome in AML patients [158]. Mutations in genes involved in DNA methylation metabolism is uncommon in both pediatric T-ALL and BCP-ALL, but mutations in PRC related genes are prevalent especially in ETP-ALL [106]. Since the CIMP profile had no connection to maturation state in T-ALL and we noted a similar methylation pattern in both BCP-ALL and T-ALL it is unlikely that specific mutations in epigenetic regulators are responsible for the different CIMP phenotypes.

During the last years it has become apparent that normal cells accumulate DNA methylation changes during aging and proliferation. It is even possible to predict age rather accurately in peripheral blood lymphocytes (PBL) based on methylation pattern [76,77]. This ageing phenotype is accelerated in cancer cells, cancer cells are predicted older than its normal counterparts. In AML for example cancer cells were predicted 10-20 years older than the chronological age of the patient [159]. Arguing that DNA methylation changes seen in leukemia is initiated before transformation and is accelerated by the increased proliferation associated with cancer development.

It is possible that CIMP high is in part a consequence of longer proliferative history. An observation supporting this is the overrepresentation of ETV6-RUNX1, positive cases among CIMP high samples in BCP-ALL. ETV6-

RUNX1 probably has a prenatal origin and therefore a longer proliferative history, supporting the relation between proliferative age and DNA methylation pattern [5]. The fact that long telomeres in BCP-ALL was associated with inferior outcome also support a model where a more aggressive and treatment resistant leukemia undergoes fewer rounds of replication, and thereby less telomere attrition.

Conclusions

Based on our findings we can draw the following conclusions

Telomeres were shorter in morphologically normal bone marrow compared to diagnostic samples from ALL patients. Long telomeres at diagnosis was linked to inferior outcome among non high-risk ALL patients.

Among T-ALL patients, two distinct methylation phenotypes exist. Patients with a CIMP low methylation profile more often have relapses and shorter overall survival. In combination with MRD, high risk patients can be stratified based on CIMP phenotype.

In BCP patients with relapse, CIMP low phenotype is associated with inferior outcome among high risk patients.

In summary we have developed a defined panel of genes that adds prognostic value in two groups of patients where we today lack reliable biological markers for treatment response. It will be very important to further investigate the biological processes that generate these distinct CIMP profiles.

Acknowledgements

Over the last 10 years (or more) numerous people have helped me finish this book.

Not all PhD-students have had TWO supervisors! **Sofie Degerman**, my second and current supervisor. You have had a major contribution to this work. You are really enthusiastic and inspiring. Your organizing skills have been crucial - from data analyzes and paper writing, to guiding me through all the paperwork for this book. From the bottom of my heart, THANK YOU! **Göran Roos**, my first supervisor. We started at the ends, unfortunately we never really figured out the never ending story of CCRF and 1301. Thank you for always encouraging me to test new ideas, I really learned a lot of methods. Thank you for pursuing the T-ALL project, the foundation of my thesis.

Erik Forestier, my co-supervisor and Nestor of childhood leukemia. Your never-ending “flow” of knowledge has been crucial for this thesis, and you have shared that knowledge generously. It has been a pleasure to work with you and thank you for all the stories about childhood oncology back in the days!

Nicol Keith, my co supervisor and one of the smartest and most inspiring people I have ever come across. Thank you for your hospitality in Glasgow and all input in my projects.

All co-authors, of both published and unpublished papers. Specially all NOPHO collaborators, **Mats Hayman** for providing data from the NOPHO registry and helpful input, **Kjeld Schmiegelow** for lots of input on the T-ALL/MRD paper. **Ulrika Norén-Nyström** for teaching me the basics of survival analysis in the hTERT paper. **Ann-Christine Syvänen**, **Gudmar Lönnnerholm** and **Jessica Nordlund** in Uppsala for sharing data and fruitful discussions.

The former ROOS-lab, in the beginning we were many, now only a few remain. **Ulrika**, thank you for all the good times and laughter, for being a good friend and a lifesaver when it comes to SPSS. **Emma**, always good ideas and insightful comments and a lot of laughs. **Zahra**, for good collaborations, **Mattias**, the computer wiz and a crucial component of the team. **Pär Larsson**, always in a good mood and always open to new ideas, also I really enjoy learning more about tularemia. **Magnus**, for good ideas

and lots of laughs. Former members, **Pawel**, long time friend and former member of the group, I really miss our food-orgies, we have to have one soon. All other former members, **Kattis, Ravi, Linda, Helene, Pia, Ellinor, Inger, and Ulla-Britt** thank you.

To all my colleagues at the **Department of Medical Biosciences** for creating a dynamic and inspiring working place. A special thanks to the foundation of the department, **Terry, Karin, Claes, Åsa and Carina** for fantastic support and excellent service!

I am also grateful to my employer Västerbottens läns landsting, for giving me the opportunity to pursuit my research.

Så till mitt andra jobb, Barnkliniken, Norrlands universitetssjukhus.

Ett stort tack till alla fantastiska kollegor på Barnkliniken, stöttande bakjourer, trevliga primärjurskollegor. Alla fantastiska sköterskor, paramedicinare och sekreterare som gör livet på barnkliniken lite bättre, hatten av för er! Ett särskilt tack till mina fantastiska kollegor på Barn 3! **Ulrika** (igen) och **Frans**, båda på sitt eget sätt ansvariga för att jag slog in på den barnonkologiska banan. **Caroline**, min mentor och stöttepelare, **Mattias, Camilla** och **Ulf**, för allt stöd.

Per-Erik, inte bara barnonkolog utan också chef, tack för att du skapat möjlighet för mig att forska.

Som tur är har jag även ett liv utanför jobbet som jag har ynnesten att dela med fantastiska vänner, många som jag känt i väldigt många år, ni är många och ingen är glömd. Ett lite extra tack till familjerna **Adamsson, Danielsson-Ringvall, Holmlund och Boily**. Och givetvis pojkarna i värmen.

Vilka föräldrar man har kan vara det som påverkar en mest av allt i livet, ni vet arv och miljö. Jag hade tur och begåvades med två enastående kloka, snälla och omtänksamma föräldrar, **Bengt** och **Eva** TACK. Ni är fortfarande en helt oundgänglig del av min vardag! Mina syskon **Johan** och **Åsa**, mina fina syskonbarn, tack för sköna dagar i stugan, vinter och sommar.

Min andra familj, klanen Danielsson. **Britt-Marie**, tack för allt stöd vi får, för din omtänksamhet. **Roland**, för all mat, alla skratt, tack. **Familjen Grüssner-Danielsson**, för er gästfrihet, i Stockholm, på Åland, i USA. Det är alltid lika roligt när vi ses. **Magnus 1**, för att du urholkar min ekonomi genom att höja nivån för vad som är ett ok vin. **Sofia** med familj, grattis till examen, du gjorde det!

Så till min alldeles egna helt underbara, älskade familj. Utan tvekan det bästa jag har. Mina två underbara barn, **Vilhelm** och **Hanna**, dagens bästa är alltid ni två. **Åsa**, min älskade fru och vardagshjälte! Tack för att du finns vid min sida och hjälper mig att bli en bättre människa. I owe you big time.

References

1. Hanahan D, Weinberg RA. The Hallmarks of Cancer. *Cell* 2000;100(1):57-70.
2. Hanahan D, Weinberg Robert A. Hallmarks of Cancer: The Next Generation. *Cell* 2011;144(5):646-674.
3. Pui C-HH, Carroll WL, Meshinchi S, et al. Biology, risk stratification, and therapy of pediatric acute leukemias: an update. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2011;29(5):551-565.
4. Esteller M. Epigenetics in cancer. *The New England journal of medicine* 2008;358(11):1148-1159.
5. Marshall GM, Carter DR, Cheung BB, et al. The prenatal origins of cancer. *Nature Reviews Cancer* 2014;14(4):277-289.
6. Höfer T, Busch K, Klapproth K, et al. Fate Mapping and Quantitation of Hematopoiesis In Vivo. *Annual Review of Immunology* 2016;34(1):449-478.
7. Jensen CT, Strid T, Sigvardsson M. Exploring the multifaceted nature of the common lymphoid progenitor compartment. *Current Opinion in Immunology* 2016;39:121-126.
8. Gatta G, Botta L, Rossi S, et al. Childhood cancer survival in Europe 1999-2007: results of EUROCARE-5--a population-based study. *The Lancet Oncology* 2014;15(1):35-47.
9. Hjalgrim LL, Rostgaard K, Schmiegelow K, et al. Age- and sex-specific incidence of childhood leukemia by immunophenotype in the Nordic countries. *Journal of the National Cancer Institute* 2003;95(20):1539-1544.
10. Forestier E, Schmiegelow K, of, et al. The incidence peaks of the childhood acute leukemias reflect specific cytogenetic aberrations. *Journal of pediatric hematology/oncology* 2006;28(8):486-495.
11. Zeller B, Gustafsson G, Forestier E, et al. Acute leukaemia in children with Down syndrome: a population-based Nordic study. *British journal of haematology* 2005;128(6):797-804.

12. Schütte P, Möricke A, Zimmermann M, et al. Preexisting conditions in pediatric ALL patients: Spectrum, frequency and clinical impact. *European journal of medical genetics* 2016;59(3):143-151.
13. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nature genetics* 2009;41(9):1006-1010.
14. Treviño LR, Yang W, French D, et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. *Nature genetics* 2009;41(9):1001-1005.
15. Perez-Andreu V, Roberts KG, Harvey RC, et al. Inherited GATA3 variants are associated with Ph-like childhood acute lymphoblastic leukemia and risk of relapse. *Nature genetics* 2013;45(12):1494-1498.
16. Xu H, Zhang H, Yang W, et al. Inherited coding variants at the CDKN2A locus influence susceptibility to acute lymphoblastic leukaemia in children. *Nature communications* 2015;6:7553.
17. Bennett JM, Catovsky D, Daniel MT, et al. The morphological classification of acute lymphoblastic leukaemia: concordance among observers and clinical correlations. *British journal of haematology* 1981;47(4):553-561.
18. Pui C-HH, Yang JJ, Hunger SP, et al. Childhood Acute Lymphoblastic Leukemia: Progress Through Collaboration. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2015;33(27):2938-2948.
19. Toft N, Birgens H, Abrahamsson J, et al. Risk group assignment differs for children and adults 1-45 yr with acute lymphoblastic leukemia treated by the NOPHO ALL-2008 protocol. *European journal of haematology* 2013;90(5):404-412.
20. van der Velden VH, Panzer-Grümayer ER, Cazzaniga G, et al. Optimization of PCR-based minimal residual disease diagnostics for childhood acute lymphoblastic leukemia in a multi-center setting. *Leukemia* 2007;21(4):706-713.
21. Björklund E, Matinlauri I, Tierens A, et al. Quality control of flow cytometry data analysis for evaluation of minimal residual disease in bone marrow from acute leukemia patients during treatment. *Journal of pediatric hematology/oncology* 2009;31(6):406-415.

22. Mullighan CG. Molecular genetics of B-precursor acute lymphoblastic leukemia. *Journal of Clinical Investigation* 2012;122(10):3407-3415.
23. Schindler JW, Van Buren D, Foudi A, et al. TEL-AML1 corrupts hematopoietic stem cells to persist in the bone marrow and initiate leukemia. *Cell stem cell* 2009;5(1):43-53.
24. Tsuzuki S, Seto M. TEL (ETV6)-AML1 (RUNX1) initiates self-renewing fetal pro-B cells in association with a transcriptional program shared with embryonic stem cells in mice. *Stem cells (Dayton, Ohio)* 2013;31(2):236-247.
25. Tsuzuki S, Seto M, Greaves M, et al. Modeling first-hit functions of the t(12;21) TEL-AML1 translocation in mice. *Proceedings of the National Academy of Sciences of the United States of America* 2004;101(22):8443-8448.
26. Kuiper RP, Schoenmakers EF, van Reijmersdal SV, et al. High-resolution genomic profiling of childhood ALL reveals novel recurrent genetic lesions affecting pathways involved in lymphocyte differentiation and cell cycle progression. *Leukemia* 2007;21(6):1258-1266.
27. Case M, Matheson E, Minto L, et al. Mutation of genes affecting the RAS pathway is common in childhood acute lymphoblastic leukemia. *Cancer research* 2008;68(16):6803-6809.
28. Parker H, An Q, Barber K, et al. The complex genomic profile of ETV6-RUNX1 positive acute lymphoblastic leukemia highlights a recurrent deletion of TBL1XR1. *Genes, chromosomes & cancer* 2008;47(12):1118-1125.
29. Paulsson K, Forestier E, Lilljebjörn H, et al. Genetic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences of the United States of America* 2010;107(50):21719-21724.
30. Paulsson K, Lilljebjörn H, Biloglav A, et al. The genomic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. *Nature genetics* 2015;47(6):672-676.
31. Paulsson K, Forestier E, Andersen MK, et al. High modal number and triple trisomies are highly correlated favorable factors in childhood B-cell precursor high hyperdiploid acute lymphoblastic

- leukemia treated according to the NOPHO ALL 1992/2000 protocols. *Haematologica* 2013;98(9):1424-1432.
32. Schmiegelow K, Forestier E, Hellebostad M, et al. Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. *Leukemia* 2010;24(2):345-354.
 33. Andersen MK, Autio K, Barbany G, et al. Paediatric B-cell precursor acute lymphoblastic leukaemia with t(1;19)(q23;p13): clinical and cytogenetic characteristics of 47 cases from the Nordic countries treated according to NOPHO protocols. *British journal of haematology* 2011;155(2):235-243.
 34. Fischer U, Forster M, Rinaldi A, et al. Genomics and drug profiling of fatal TCF3-HLF-positive acute lymphoblastic leukemia identifies recurrent mutation patterns and therapeutic options. *Nature genetics* 2015;47(9):1020-1029.
 35. Rao RC, Dou Y. Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nature reviews Cancer* 2015;15(6):334-346.
 36. Andersson AK, Ma J, Wang J, et al. The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nature genetics* 2015;47(4):330-337.
 37. Nowell PC. Discovery of the Philadelphia chromosome: a personal perspective. *The Journal of clinical investigation* 2007;117(8):2033-2035.
 38. Notta F, Mullighan CG, Wang JC, et al. Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. *Nature* 2011;469(7330):362-367.
 39. Schultz KR, Bowman WP, Aledo A, et al. Improved early event-free survival with imatinib in Philadelphia chromosome-positive acute lymphoblastic leukemia: a children's oncology group study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2009;27(31):5175-5181.
 40. Schultz KR, Carroll A, Heerema NA, et al. Long-term follow-up of imatinib in pediatric Philadelphia chromosome-positive acute lymphoblastic leukemia: Children's Oncology Group study AALL0031. *Leukemia* 2014;28(7):1467-1471.

41. Li Y, Schwab C, Ryan SL, et al. Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia. *Nature* 2014;508(7494):98-102.
42. Harrison CJ, Moorman AV, Schwab C, et al. An international study of intrachromosomal amplification of chromosome 21 (iAMP21): cytogenetic characterization and outcome. *Leukemia* 2014;28(5):1015-1021.
43. Moorman AV. New and emerging prognostic and predictive genetic biomarkers in B-cell precursor acute lymphoblastic leukemia. *Haematologica* 2016;101(4):407-416.
44. Olsson L, Castor A, Behrendtz M, et al. Deletions of IKZF1 and SPRED1 are associated with poor prognosis in a population-based series of pediatric B-cell precursor acute lymphoblastic leukemia diagnosed between 1992 and 2011. *Leukemia* 2014;28(2):302-310.
45. Olsson L, Ivanov Öfverholm I, Norén-Nyström U, et al. The clinical impact of IKZF1 deletions in paediatric B-cell precursor acute lymphoblastic leukaemia is independent of minimal residual disease stratification in Nordic Society for Paediatric Haematology and Oncology treatment protocols used between 1992 and 2013. *British journal of haematology* 2015;170(6):847-858.
46. Boer JM, van der Veer A, Rizopoulos D, et al. Prognostic value of rare IKZF1 deletion in childhood B-cell precursor acute lymphoblastic leukemia: an international collaborative study. *Leukemia* 2016;30(1):32-38.
47. Boer JM, Marchante JRR, Evans WE, et al. BCR-ABL1-like cases in pediatric acute lymphoblastic leukemia: a comparison between DCOG/Erasmus MC and COG/St. Jude signatures. *Haematologica* 2015;100(9):7.
48. Goldberg JM, Silverman LB, Levy DE, et al. Childhood T-cell acute lymphoblastic leukemia: the Dana-Farber Cancer Institute acute lymphoblastic leukemia consortium experience. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2003;21(19):3616-3622.
49. Krieger D, Moericke A, Oschlies I, et al. Frequency and clinical relevance of DNA microsatellite alterations of the CDKN2A/B, ATM and p53 gene loci: a comparison between pediatric precursor T-cell lymphoblastic lymphoma and T-cell lymphoblastic leukemia. *Haematologica* 2010;95(1):158-162.

50. 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* 2010;24(12):2014-2022.
51. Kim WY, Sharpless NE. The regulation of INK4/ARF in cancer and aging. *Cell* 2006;127(2):265-275.
52. Belver L, Ferrando A. The genetics and mechanisms of T cell acute lymphoblastic leukaemia. *Nature reviews Cancer* 2016;16(8):494-507.
53. Margolin AA, Palomero T, Sumazin P, et al. ChIP-on-chip significance analysis reveals large-scale binding and regulation by human transcription factor oncogenes. *Proceedings of the National Academy of Sciences of the United States of America* 2009;106(1):244-249.
54. Fogelstrand L, Staffas A, Wasslavik C, et al. Prognostic implications of mutations in NOTCH1 and FBXW7 in childhood T-ALL treated according to the NOPHO ALL-1992 and ALL-2000 protocols. *Pediatric blood & cancer* 2014;61(3):424-430.
55. Larmonie NS, Dik WA, Meijerink JPP, et al. 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* 2013;98(8):1173-1184.
56. Matlawska-Wasowska K, Kang H, Devidas M, et al. MLL rearrangements impact outcome in HOXA-deregulated T-lineage acute lymphoblastic leukemia: a Children's Oncology Group Study. *Leukemia* 2016;30(9):1909-1912.
57. Ferrando AA, Neuberg DS, Staunton J, et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. *Cancer cell* 2002;1(1):75-87.
58. 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.

59. Coustan-Smith E, Mullighan CG, Onciu M, et al. Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. *The Lancet Oncology* 2009;10(2):147-156.
60. Patrick K, Wade R, Goulden N, et al. Outcome for children and young people with Early T-cell precursor acute lymphoblastic leukaemia treated on a contemporary protocol, UKALL 2003. *British journal of haematology* 2014;166(3):421-424.
61. Zuurbier L, Gutierrez A, Mullighan CG, et al. Immature MEF2C-dysregulated T-cell leukemia patients have an early T-cell precursor acute lymphoblastic leukemia gene signature and typically have non-rearranged T-cell receptors. *Haematologica* 2014;99(1):94-102.
62. Oskarsson T, Söderhäll S, Arvidson J, et al. Relapsed childhood acute lymphoblastic leukemia in the Nordic countries: prognostic factors, treatment and outcome. *Haematologica* 2016;101(1):68-76.
63. Bhojwani D, Pui C-HH. Relapsed childhood acute lymphoblastic leukaemia. *The Lancet Oncology* 2013;14(6):17.
64. Tallen G, Ratei R, Mann G, et al. Long-term outcome in children with relapsed acute lymphoblastic leukemia after time-point and site-of-relapse stratification and intensified short-course multidrug chemotherapy: results of trial ALL-REZ BFM 90. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010;28(14):2339-2347.
65. Beyermann B, Adams HP, Henze G. Philadelphia chromosome in relapsed childhood acute lymphoblastic leukemia: a matched-pair analysis. Berlin-Frankfurt-Münster Study Group. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 1997;15(6):2231-2237.
66. Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. *Nature reviews Genetics* 2011;12(1):7-18.
67. Radulović V, de Haan G, Klauke K. Polycomb-group proteins in hematopoietic stem cell regulation and hematopoietic neoplasms. *Leukemia* 2013;27(3):523-533.
68. Cao R, Wang L, Wang H, et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science (New York, NY)* 2002;298(5595):1039-1043.

69. Boulard M, Edwards JR, Bestor TH. FBXL10 protects Polycomb-bound genes from hypermethylation. *Nature genetics* 2015;47(5):479-485.
70. Bergman Y, Cedar H. DNA methylation dynamics in health and disease. *Nature Structural & Molecular Biology* 2013;20(3).
71. Jeltsch A, Jurkowska RZ. New concepts in DNA methylation. *Trends in biochemical sciences* 2014;39(7):310-318.
72. Di Ruscio A, Ebralidze AK, Benoukraf T, et al. DNMT1-interacting RNAs block gene-specific DNA methylation. *Nature* 2013;503(7476):371-376.
73. Fang J, Cheng J, Wang J, et al. Hemi-methylated DNA opens a closed conformation of UHRF1 to facilitate its histone recognition. *Nature communications* 2016;7:11197.
74. Elliott EN, Sheaffer KL, Kaestner KH. The 'de novo' DNA methyltransferase Dnmt3b compensates the Dnmt1-deficient intestinal epithelium. *eLife* 2016;5.
75. Smith ZD, Chan MM, Mikkelsen TS, et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* 2012;484(7394):339-344.
76. Hannum G, Guinney J, Zhao L, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular cell* 2013;49(2):359-367.
77. Horvath S. DNA methylation age of human tissues and cell types. *Genome biology* 2013;14(10).
78. Kulis M, Merkel A, Heath S, et al. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nature genetics* 2015;47(7):746-756.
79. Ono R, Taki T, Taketani T, et al. LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer research* 2002;62(14):4075-4080.
80. Ko M, An J, Rao A. DNA methylation and hydroxymethylation in hematologic differentiation and transformation. *Current Opinion in Cell Biology* 2015;37:91-101.

81. Rasmussen KD, Helin K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes & development* 2016;30(7):733-750.
82. Kafer GR, Li X, Horii T, et al. 5-Hydroxymethylcytosine Marks Sites of DNA Damage and Promotes Genome Stability. *Cell reports* 2016;14(6):1283-1292.
83. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature reviews Genetics* 2012;13(7):484-492.
84. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proceedings of the National Academy of Sciences of the United States of America* 2002;99(6):3740-3745.
85. Deaton AMM, Bird A. CpG islands and the regulation of transcription. *Genes & development* 2011;25(10):1010-1022.
86. Jeong M, Goodell MA. New answers to old questions from genome-wide maps of DNA methylation in hematopoietic cells. *Experimental hematology* 2014;42(8):609-617.
87. Barlow DP, Bartolomei MS. Genomic imprinting in mammals. *Cold Spring Harbor perspectives in biology* 2014;6(2).
88. Lock LF, Takagi N, Martin GR. Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. *Cell* 1987;48(1):39-46.
89. Illingworth RS, Bird AP. CpG islands--'a rough guide'. *FEBS letters* 2009;583(11):1713-1720.
90. Hellman A, Chess A. Gene body-specific methylation on the active X chromosome. *Science (New York, NY)* 2007;315(5815):1141-1143.
91. Schmidl C, Klug M, Boeld TJ, et al. Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. *Genome research* 2009;19(7):1165-1174.
92. Wiench M, John S, Baek S, et al. DNA methylation status predicts cell type-specific enhancer activity. *The EMBO journal* 2011;30(15):3028-3039.

93. Bröske A-MM, Vockentanz L, Kharazi S, et al. DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nature genetics* 2009;41(11):1207-1215.
94. Challen GA, Sun D, Mayle A, et al. Dnmt3a and Dnmt3b have overlapping and distinct functions in hematopoietic stem cells. *Cell stem cell* 2014;15(3):350-364.
95. Cimmino L, Dawlaty MM, Ndiaye-Lobry D, et al. TET1 is a tumor suppressor of hematopoietic malignancy. *Nature immunology* 2015;16(6):653-662.
96. Sterlin D, Velasco G, Moshous D, et al. Genetic, Cellular and Clinical Features of ICF Syndrome: a French National Survey. *Journal of clinical immunology* 2016;36(2):149-159.
97. Rodriguez RM, Suarez-Alvarez B, Mosén-Ansorena D, et al. Regulation of the transcriptional program by DNA methylation during human $\alpha\beta$ T-cell development. *Nucleic acids research* 2015;43(2):760-774.
98. Beerman I, Bock C, Garrison BS, et al. Proliferation-dependent alterations of the DNA methylation landscape underlie hematopoietic stem cell aging. *Cell stem cell* 2013;12(4):413-425.
99. Couronné L, Bastard C, Bernard OA. TET2 and DNMT3A mutations in human T-cell lymphoma. *The New England journal of medicine* 2012;366(1):95-96.
100. Feinberg AP, Koldobskiy MA, Göndör A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nature reviews Genetics* 2016;17(5):284-299.
101. Ley TJ, Ding L, Walter MJ, et al. DNMT3A mutations in acute myeloid leukemia. *The New England journal of medicine* 2010;363(25):2424-2433.
102. Grossmann V, Haferlach C, Weissmann S, et al. The molecular profile of adult T-cell acute lymphoblastic leukemia: mutations in RUNX1 and DNMT3A are associated with poor prognosis in T-ALL. *Genes, chromosomes & cancer* 2013;52(4):410-422.
103. Neumann M, Heesch S, Schlee C, et al. Whole-exome sequencing in adult ETP-ALL reveals a high rate of DNMT3A mutations. *Blood* 2013;121(23):4749-4752.

104. Mullighan CG, Zhang J, Kasper LH, et al. CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature* 2011;471(7337):235-239.
105. Zhang J, Ding L, Holmfeldt L, et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* 2012;481(7380):157-163.
106. Peirs S, Van der Meulen J, Van de Walle I, et al. Epigenetics in T-cell acute lymphoblastic leukemia. *Immunological reviews* 2015;263(1):50-67.
107. Van der Meulen J, Sanghvi V, Mavrakis K, et al. The H3K27me3 demethylase UTX is a gender-specific tumor suppressor in T-cell acute lymphoblastic leukemia. *Blood* 2015;125(1):13-21.
108. Sanjuan-Pla A, Bueno C, Prieto C, et al. Revisiting the biology of infant t(4;11)/MLL-AF4+ B-cell acute lymphoblastic leukemia. *Blood* 2015;126(25):2676-2685.
109. Cierpicki T, Risner LE, Grembecka J, et al. Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. *Nature structural & molecular biology* 2010;17(1):62-68.
110. Lee S-TT, Muench MO, Fomin ME, et al. Epigenetic remodeling in B-cell acute lymphoblastic leukemia occurs in two tracks and employs embryonic stem cell-like signatures. *Nucleic acids research* 2015;43(5):2590-2602.
111. Nordlund J, Bäcklin CL, Wahlberg P, et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. *Genome biology* 2013;14(9).
112. Nordlund J, Bäcklin CL, Zachariadis V, et al. DNA methylation-based subtype prediction for pediatric acute lymphoblastic leukemia. *Clinical epigenetics* 2015;7(1):11.
113. Gabriel AS, Lafta FM, Schwalbe EC, et al. Epigenetic landscape correlates with genetic subtype but does not predict outcome in childhood acute lymphoblastic leukemia. *Epigenetics* 2015;10(8):717-726.
114. Milani L, Lundmark A, Kiialainen A, et al. DNA methylation for subtype classification and prediction of treatment outcome in

- patients with childhood acute lymphoblastic leukemia. *Blood* 2010;115(6):1214-1225.
115. Hogan LE, Meyer JA, Yang J, et al. Integrated genomic analysis of relapsed childhood acute lymphoblastic leukemia reveals therapeutic strategies. *Blood* 2011;118(19):5218-5226.
 116. Kunz JB, Rausch T, Bandapalli OR, et al. Pediatric T-cell lymphoblastic leukemia evolves into relapse by clonal selection, acquisition of mutations and promoter hypomethylation. *Haematologica* 2015;100(11):1442-1450.
 117. Svenson U, Roos G. Telomere length as a biological marker in malignancy. *Biochimica et biophysica acta* 2009;1792(4):317-323.
 118. Jafri MA, Ansari SA, Alqahtani MH, et al. Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. *Genome medicine* 2016;8(1):69.
 119. Shay JW. Role of Telomeres and Telomerase in Aging and Cancer. *Cancer discovery* 2016;6(6):584-593.
 120. Weisenberger DJ, Siegmund KD, Campan M, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nature genetics* 2006;38(7):787-793.
 121. Weisenberger DJ, Campan M, Long TI, et al. Analysis of repetitive element DNA methylation by MethyLight. *Nucleic acids research* 2005;33(21):6823-6836.
 122. Grabowski P, Hultdin M, Karlsson K, et al. Telomere length as a prognostic parameter in chronic lymphocytic leukemia with special reference to VH gene mutation status. *Blood* 2005;105(12):4807-4812.
 123. Hultdin M, Rosenquist R, Thunberg U, et al. Association between telomere length and V(H) gene mutation status in chronic lymphocytic leukaemia: clinical and biological implications. *British journal of cancer* 2003;88(4):593-598.
 124. Brümmendorf TH, Holyoake TL, Rufer N, et al. Prognostic implications of differences in telomere length between normal and malignant cells from patients with chronic myeloid leukemia measured by flow cytometry. *Blood* 2000;95(6):1883-1890.

125. Drummond M, Lennard A, Brümmendorf T, et al. Telomere shortening correlates with prognostic score at diagnosis and proceeds rapidly during progression of chronic myeloid leukemia. *Leukemia & lymphoma* 2004;45(9):1775-1781.
126. Wenn K, Tomala L, Wilop S, et al. Telomere length at diagnosis of chronic phase chronic myeloid leukemia (CML-CP) identifies a subgroup with favourable prognostic parameters and molecular response according to the ELN criteria after 12 months of treatment with nilotinib. *Leukemia* 2015;29(12):2402-2404.
127. Zinn RL, Pruitt K, Eguchi S, et al. hTERT is expressed in cancer cell lines despite promoter DNA methylation by preservation of unmethylated DNA and active chromatin around the transcription start site. *Cancer research* 2007;67(1):194-201.
128. Pettigrew KA, Armstrong RN, Colyer HAA, et al. Differential TERT promoter methylation and response to 5-aza-2'-deoxycytidine in acute myeloid leukemia cell lines: TERT expression, telomerase activity, telomere length, and cell death. *Genes, chromosomes & cancer* 2012;51(8):768-780.
129. Toyota M, Ahuja N, Ohe-Toyota M, et al. CpG island methylator phenotype in colorectal cancer. *Proceedings of the National Academy of Sciences of the United States of America* 1999;96(15):8681-8686.
130. Hughes LA, Melotte V, de Schrijver J, et al. The CpG island methylator phenotype: what's in a name? *Cancer research* 2013;73(19):5858-5868.
131. Gendrel A-VV, Apedaile A, Coker H, et al. Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. *Developmental cell* 2012;23(2):265-279.
132. Bracken AP, Dietrich N, Pasini D, et al. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes & development* 2006;20(9):1123-1136.
133. Lee TI, Jenner RG, Boyer LA, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 2006;125(2):301-313.

134. Chatterton Z, Morenos L, Mechinaud F, et al. Epigenetic deregulation in pediatric acute lymphoblastic leukemia. *Epigenetics* 2014;9(3):459-467.
135. Gutierrez A, Dahlberg SE, Neuberg DS, et al. Absence of biallelic TCRgamma deletion predicts early treatment failure in pediatric T-cell acute lymphoblastic leukemia. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010;28(24):3816-3823.
136. Irving JA, Enshaei A, Parker CA, et al. Integration of genetic and clinical risk factors improves prognostication in relapsed childhood B-cell precursor acute lymphoblastic leukemia. *Blood* 2016;128(7):911-922.
137. Toyota M, Kopecky KJ, Toyota MO, et al. Methylation profiling in acute myeloid leukemia. *Blood* 2001;97(9):2823-2829.
138. Garcia-Manero G, Bueso-Ramos C, Daniel J, et al. DNA methylation patterns at relapse in adult acute lymphocytic leukemia. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2002;8(6):1897-1903.
139. Garcia-Manero G, Daniel J, Smith TL, et al. DNA methylation of multiple promoter-associated CpG islands in adult acute lymphocytic leukemia. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2002;8(7):2217-2224.
140. Shen L, Kondo Y, Issa J-PP, et al. Lack of p21(CIP1) DNA methylation in acute lymphocytic leukemia. *Blood* 2002;100(9):3432.
141. Ueki T, Toyota M, Skinner H, et al. Identification and characterization of differentially methylated CpG islands in pancreatic carcinoma. *Cancer research* 2001;61(23):8540-8546.
142. Asmar F, Punj V, Christensen J, et al. Genome-wide profiling identifies a DNA methylation signature that associates with TET2 mutations in diffuse large B-cell lymphoma. *Haematologica* 2013;98(12):1912-1920.
143. Hinoue T, Weisenberger DJ, Lange CP, et al. Genome-scale analysis of aberrant DNA methylation in colorectal cancer. *Genome research* 2012;22(2):271-282.

144. Noushmehr H, Weisenberger DJ, Diefes K, et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer cell* 2010;17(5):510-522.
145. O'Riain C, O'Shea DM, Yang Y, et al. Array-based DNA methylation profiling in follicular lymphoma. *Leukemia* 2009;23(10):1858-1866.
146. Figueroa ME, Chen S-CC, Andersson AK, et al. Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia. *The Journal of clinical investigation* 2013;123(7):3099-3111.
147. Paulsson K, An Q, Moorman AV, et al. Methylation of tumour suppressor gene promoters in the presence and absence of transcriptional silencing in high hyperdiploid acute lymphoblastic leukaemia. *British journal of haematology* 2009;144(6):838-847.
148. Kraszewska MD, Dawidowska M, Larmonie NS, et al. DNA methylation pattern is altered in childhood T-cell acute lymphoblastic leukemia patients as compared with normal thymic subsets: insights into CpG island methylator phenotype in T-ALL. *Leukemia* 2012;26(2):367-371.
149. Terriou L, Ben Abdelali R, Roumier C, et al. C/EBPA methylation is common in T-ALL but not in Mo AML. *Blood* 2009;113(8):1864.
150. Sandoval J, Heyn H, Méndez-González J, et al. Genome-wide DNA methylation profiling predicts relapse in childhood B-cell acute lymphoblastic leukaemia. *British journal of haematology* 2013;160(3):406-409.
151. Keshet I, Schlesinger Y, Farkash S, et al. Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nature genetics* 2006;38(2):149-153.
152. Krentz S, Hof J, Mendioroz A, et al. Prognostic value of genetic alterations in children with first bone marrow relapse of childhood B-cell precursor acute lymphoblastic leukemia. *Leukemia* 2013;27(2):295-304.
153. Moorman AV, Enshaei A, Schwab C, et al. A novel integrated cytogenetic and genomic classification refines risk stratification in pediatric acute lymphoblastic leukemia. *Blood* 2014;124(9):1434-1444.

154. Volanakis EJ, Boothby MR, Sherr CJ. Epigenetic regulation of the Ink4a-Arf (Cdkn2a) tumor suppressor locus in the initiation and progression of Notch1-driven T cell acute lymphoblastic leukemia. *Experimental hematology* 2013;41(4):377-386.
155. Borssén M, Palmqvist L, Karrman K, et al. Promoter DNA methylation pattern identifies prognostic subgroups in childhood T-cell acute lymphoblastic leukemia. *PloS one* 2013;8(6).
156. D'Angiò M, Valsecchi MG, Testi AM, et al. Clinical features and outcome of SIL/TAL1-positive T-cell acute lymphoblastic leukemia in children and adolescents: a 10-year experience of the AIEOP group. *Haematologica* 2015;100(1):3.
157. Qu Y, Lennartsson A, Gaidzik VI, et al. Differential methylation in CN-AML preferentially targets non-CGI regions and is dictated by DNMT3A mutational status and associated with predominant hypomethylation of HOX genes. *Epigenetics* 2014;9(8):1108-1119.
158. Schulze I, Rohde C, Scheller-Wendorff M, et al. Increased DNA methylation of Dnmt3b targets impairs leukemogenesis. *Blood* 2016;127(12):1575-1586.
159. Lin Q, Wagner W. Epigenetic Aging Signatures Are Coherently Modified in Cancer. *PLoS genetics* 2015;11(6).

hTERT promoter methylation and telomere length in childhood acute lymphoblastic leukemia—associations with immunophenotype and cytogenetic subgroup

Magnus Borssén^a, Inger Cullman^a, Ulrika Norén-Nyström^b, Christer Sundström^c, Anna Porwit^d, Erik Forestier^b, and Göran Roos^a

^aDepartment of Medical Biosciences; ^bClinical Sciences, Pediatrics, Umeå University, Umeå, Sweden; ^cDepartment of Genetics and Pathology, Uppsala University, Uppsala, Sweden; ^dDepartment of Pathology, Karolinska University Hospital, Stockholm, Sweden

(Received 6 March 2011; revised 17 August 2011; accepted 22 August 2011)

Telomere maintenance, important for long-term cell survival and malignant transformation, is directed by a multitude of factors, including epigenetic mechanisms, and has been implicated in outcomes for patients with leukemia. In the present study, the objective was to investigate the biological and clinical significance of telomere length and promoter methylation of the human telomerase reverse transcriptase gene in childhood acute lymphoblastic leukemia. A cohort of 169 childhood acute lymphoblastic leukemias was investigated for telomere length, human telomerase reverse transcriptase gene promoter methylation status, genomic aberrations, immunophenotype, and clinical outcomes. Methylation of the core promoter of the human telomerase reverse transcriptase (*hTERT*) gene was demonstrated in 24% of diagnostic samples, with a significant difference between B-cell precursor (n = 130) and T-cell acute lymphoblastic leukemia (ALL) (n = 17) cases (18% and 72%, respectively; $p < 0.001$). No remission sample demonstrated *hTERT* promoter methylation (n = 40). Within the B-cell precursor group, t(12;21)(p13;q22) [ETV6/RUNX1] cases (n = 19) showed a much higher frequency of *hTERT* methylation than high-hyperdiploid (51–61 chromosomes) ALL (n = 44) (63% and 7%, respectively; $p < 0.001$). *hTERT* messenger RNA levels were negatively associated with methylation status and, in the t(12;21) group, methylated cases had shorter telomeres ($p = 0.017$). In low-risk B-cell precursor patients (n = 101), long telomeres indicated a worse prognosis. The collected data from the present study indicate that the telomere biology in childhood ALL has clinical implications and reflects molecular differences between diverse ALL subgroups. © 2011 ISEH - Society for Hematology and Stem Cells. Published by Elsevier Inc.

Telomere maintenance is important for long-term cell survival. Upregulation of telomerase activity, leading to preserved telomeres, is a hallmark of cancer cells. The actual telomere length of a cell depends on a balance between various positive and negative forces. Telomerase-directed synthesis of new telomeric repeats is the main positive factor, although an alternative lengthening process through homologous recombination is rather common in specific tumor types. Each cell cycle round is associated with telomere attrition due to the end replication problem and environmental causes, e.g., oxidative stress, can induce enhanced telomere loss [1,2]. Accessibility to the telomere end is an important step for telomerase action on the telomere governed by a multitude of factors, including

telomere-binding proteins and epigenetic modifications at the telomere ends [3,4]. Alteration of the promoter chromatin landscape is one main factor controlling the transcriptional activity of *hTERT* and telomerase RNA (*hTR*) genes. One example is chromatin relaxation occurring at hypoxia and leading to increased *hTERT* and *hTR* gene expression [4]. One of the best studied chromatin modifications is methylation of cytosines followed by guanine, i.e., CpG dinucleotides. Especially, clusters of CpGs, called CpG islands, have been of interest. The global methylation pattern generally decreases during cancer progression, whereas methylation of CpG islands at tumor suppressor genes usually results in gene silencing [5–7].

Many hematopoietic malignancies, including myelodysplastic syndrome and acute leukemia, demonstrate promoter hypermethylation of tumor suppressor and cell cycle control genes [6]. Methylation “hot spots” have been demonstrated

Offprint request to: Göran Roos, Ph.D., Department of Medical Biosciences, Umeå University, S-90187, Umeå, Sweden; E-mail: goran.roos@medbio.umu.se

Table 1. Clinical characteristics of 169 children with ALL subdivided by genetic subgroup

Group	Diploid	51–61 chr	<45 chr	t(12;21)	t(9;22)	t(1;19)	MLL/11q23	dic(9;20)	iAMP21	Other	No result*	T cell	Total
Male	10	25	0	12	7	2	1	1	1	12	4	17	92
Female	10	28	1	8	1	2	0	2	0	19	2	4	77
n (%)	20 (12)	53 (31)	1 (0.6)	20 (12)	8 (4.7)	4 (2.4)	1 (0.6)	3 (1.8)	1 (0.6)	31 (18)	6 (3.6)	21 (12)	169
WBC × 10 ⁹ /L, median (range)	10 (2–1400)	6.0 (1–124)	17	5.7 (1–140)	52 (2–172)	6.5 (2–16)	193	46 (8–88)	2.0	11 (0–744)	28 (4–74)	30 (0–588)	
Age (y), median (range)	8.5 (0–17)	4.0 (1–15)	12	4.5 (1–14)	10 (5–13)	6.5 (4–18)	0	2.0 (1–14)	15	5.0 (0–16)	7.0 (5–15)	9.0 (1–16)	
Risk group LR/HR	16/4	45/8	0/1	19/1	0/8	4/0	0/1	1/2	1/0	20/11	5/1	0/21	111/58
Events, CR1/relapse	14/6	45/8	0/1	20/0	4/4	3/1	1/0	3/0	1/0	22/9	4/2	15/6	132/37

Chr = chromosomes; LR/HR = low risk/high risk; WBC = white blood cell count.

*No metaphases found.

using microarray analysis of ALL [8] and methylation patterns seem to give prognostic information for both B- and T-ALL [9,10]. In a separate analysis restricted to t(12;21)(p13;q22) [ETV6/RUNX1]-positive ALL, the presence of CpG island methylator phenotype was associated with a poor prognosis [10].

Regarding the *hTERT* gene, conflicting data exist for promoter methylation in relation to gene expression and telomerase activity [11–14]. Recent findings have clarified that the methylation status of the *hTERT* core promoter close to the transcription start site is important, showing an inverse relationship between methylation and gene expression, making DNA methylation a potentially important regulatory mechanism [15,16]. *hTERT* promoter methylation has also been associated with lower telomerase activity in chronic lymphocytic leukemia [17], whereas no data has been published for ALL. Regarding telomere length in malignant hematopoietic disorders, several reports have demonstrated short telomeres to be connected with progressive disease and worse prognosis [18].

In the present study, we have analyzed a series of childhood ALL cases and demonstrate that *hTERT* gene methylation was coupled to gene expression, telomere length, immunophenotype, and genotype. Furthermore, telomere length was a prognostic indicator in low-risk B-cell precursor (BCP) ALL.

Materials and methods

Patients and samples

Mononuclear cells were separated from diagnostic bone marrow aspirates obtained from 169 children, 0 to 17.9 years of age, diagnosed with ALL from 1988 through 2006 at the childhood oncology centers in Umeå, Uppsala, and Stockholm, Sweden. Remission samples were obtained at cessation of therapy from 40 of these 169 patients. Patients with Down syndrome were excluded. The leukemia diagnosis was in each case based on morphology, immunophenotyping, and cytogenetic analysis. BCP immunophenotype was found in 148 and precursor T-ALL in 21 cases. Our study included the following structural cytogenetic aberrations: t(12;21)(p13;q22)[ETV6/RUNX1];iAMP21[RUNX1x3-7];t(1;19)(q23;p13)[TCF3/PBX1]; t(9;22)(q34;q11) [BCR/ABL]; der(11)(q23)[MLL]; and dic(9;20)(p13;q11) established by G-band karyotyping, fluorescence in situ hybridization, and/or reverse transcription polymerase chain reaction (PCR). Clinical data from the study is presented in Table 1. All patients available for survival analysis were treated according to Nordic Society of Pediatric Hematology and Oncology (NOPHO) ALL protocols 1992 or 2000 [19]. The NOPHO ALL-1992 and NOPHO ALL-2000 protocols had similar backbones with minor changes in the latter protocol. BCP-ALL patients with no high-risk features according to protocol were defined as low-risk in this study (n = 102). Median follow-up time in first complete remission was 58 months (range, 4–242 months). This study was approved by the ethical committee at Umeå University (Dnr 05-102M).

DNA histogram analysis

Propidium iodide–stained leukemic samples and control cells (chicken and trout erythrocytes) were analyzed by flow cytometry

using standard techniques. DNA histograms were evaluated regarding ploidy according to the RFIT model (Becton Dickinson Immunocytometry Systems, San Jose, CA, USA) in order to identify high-hyperdiploid (HeH) cases.

Telomere length determination by quantitative real-time PCR

Genomic DNA was extracted using conventional methods at the respective center. Relative telomere length (RTL) was determined using real-time PCR as described elsewhere [20]. In the present study, modifications to the original protocol included new sets of primers (tel 1b CGGTTTGGTTGGGTTT GGGTTTGGGTTT GGGTTTGGGTTT; tel 2b GGCTTGCCTTACCCTTACCCTTACCCTTACCCTTACCCT; HBG3: TGTGCTGGC CCATCACTTTG; HBG4: ACCAGCCACCACTTTCTGATAGG).

The primer concentrations were 100 nM tel 1b primer and 900 nM tel 2b primer for telomere amplification; 400 nM HBG3 primer and 400 nM HBG4 primer for single copy gene amplification. Genomic *Escherichia coli* DNA was added to the reaction mix, to act as a carrier DNA in order to reduce variation between replicates.

hTERT promoter methylation analysis

In 147 (130 BCP-ALL, 17 T-ALL) of 169 ALL patients, enough DNA was available for analysis of hTERT core promoter methylation. Sodium bisulfite-treated DNA (EZ DNA methylation kit; Zymed, Orange, CA, USA) was analyzed for methylation status using quantitative MethyLight, a fluorescence-based real-time PCR method [21]. MssI (New England Biolab, Ipswich, MA, USA)-treated leukocyte DNA (Promega, Madison, WI, USA) was used as reference sample and Alu sequences as controls in methylation-independent reactions as described [22]. The amplified hTERT promoter region correspond to positions 10996 to 11111 of hTERT sequence AF128893 (Gene Bank). PCR was performed in an ABI Prism 7900 HT sequence detection system (Applied Biosystems, Sundbyberg, Sweden). PCR program was 95°C for 10 minutes followed by 95°C for 15 seconds and 60°C for 1 minute repeated for 50 cycles. To determine if a sample was considered methylated or not, the following formula was used:

Percent methylated reference = $100 \times (\text{methylated reaction of the sample} / \text{control reaction of the sample}) / (\text{methylated reaction of the MssI-treated reference sample} / \text{control reaction of the MssI-treated reference sample})$.

Samples were considered positive for methylation if percent methylated reference > 10 [22].

Primer sequences used were:

ALU, Forward primer: GTTAGGTATAGTGGTTTATATTTGT AATTTTAGTA and reverse primer: ATTAACATAAATAATCTTA AACTCCTAACCTCA; probe: 6FAM-CCTACCTTAACCTCCC-MGBNFQ. TERT, forward primer: GGATTCGCGGGTATA-GA CGTT and reverse primer: CGAAATCCGCGGAAA; probe: 6FAM-CCCAATCC-CTCCGCCACGTAAAA-BHQ-1.

hTERT RT-PCR

Total RNA was extracted from a subset of samples from 18 BCP patients, from whom enough material was available, using Trazol (Invitrogen, Carlsbad, CA, USA), according to manufacturer's protocol. RNA was converted to complementary DNA with Super-script II (Invitrogen) and random hexameric primers (Applied Biosystems). Messenger RNA levels were quantified by real-time reverse transcription PCR using Light Cycler SYBR Green

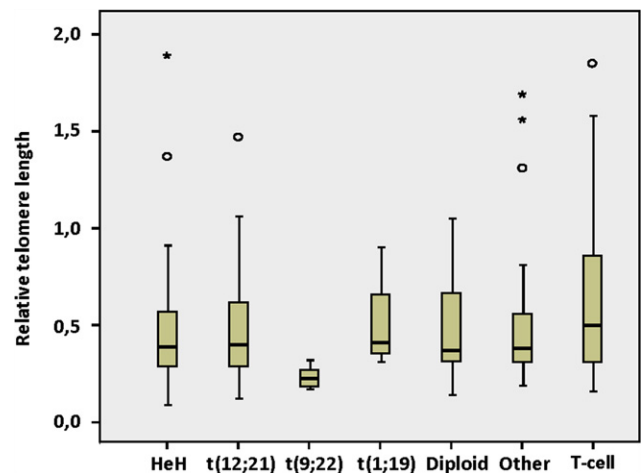


Figure 1. Relative telomere length distribution in different cytogenetic BCP subgroups and in T-ALL. The box plot presents the minimum, lower quartile, median, upper quartile, and maximum RTL, except for outliers.

I technology (Roche Diagnostics, Mannheim, Germany). Primer sequences in the study were as follows: hTERT, all splice variants, forward: CGAGCTGCTCAGGTCCTTCT and reverse: GCA CCC TCT TCA AGT GCT GT. Full length hTERT, forward: TAC GAC ACC ATC CCC CAG and reverse: ACGA ACT GTC GCA TGT ACG G. β -actin, forward: CCC AGC ACA ATG AAG ATC AAG ATC AT and reverse: ATC TGC TGG AAG GTG GAC AGC GA.

Statistical analyses

Differences in RTL, hTERT methylation status, and hTERT mRNA expression between groups were evaluated using the Mann-Whitney *U*-test. The Spearman's correlation test was used for correlation between telomere length and white blood cell count, S-phase, and flow cytometric DNA index and hTERT expression. Survival curves were plotted according to the Kaplan-Meier method and differences in outcomes between groups were tested by the log-rank test. The only event considered was relapse of disease. In survival analysis, children remaining in continuous complete remission were censored at the last known time of follow-up, and patients with death in remission were excluded ($n = 1$). The level of statistical significance was defined as $p < 0.05$ (two-sided). All data were analyzed using the SPSS software for Windows, version 15.0 (SPSS, Chicago, IL, USA).

Results

Telomere length in ALL subgroups

The median RTL for all ALL cases ($n = 169$, Fig. 1) at diagnosis was 0.37, which was significantly shorter than samples obtained at cessation of therapy (median RTL = 0.73, $n = 40$; $p < 0.001$, not shown in Fig. 1). There was a near-significant difference ($p = 0.063$) regarding RTL between BCP (median RTL = 0.36) and T-ALL (median RTL = 0.49) at diagnosis. Cases with t(9;22)(q11;q34)[BCR/ABL1] (although few) had significantly shorter telomeres compared to HeH (51–61 chromosomes) or t(12;21) cases as well as compared

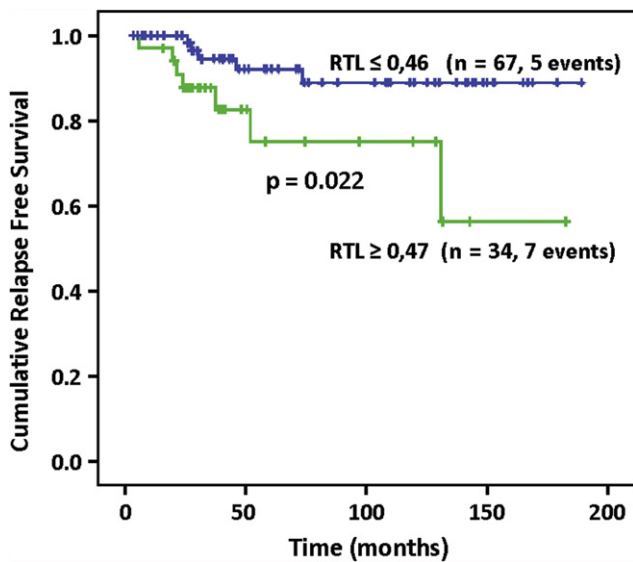


Figure 2. Cumulative relapse-free survival in low-risk BCP patients subdivided after telomere length. Patients having RTL above the mean value (RTL = 0.47), representing the upper third of patients with BCP-ALL had inferior outcomes compared to cases with shorter telomeres.

to all BCP-ALL combined ($p = 0.002$, $p = 0.001$, and $p = 0.001$ respectively). Telomere length distribution in the most prevalent cytogenetic subgroups is shown in Figure 1.

Telomere length was not associated with white blood cell count or S-phase ($p = 0.179$ and $p = 0.809$, respectively). Further, logistic regression analysis showed no correlation between telomere length and age at diagnosis.

Telomere length predicted outcomes among low-risk BCP patients

Survival analysis was performed in the largest risk group, i.e., low-risk BCP patients treated either with the NOPHO 92 or NOPHO 2000 protocols ($n = 101$) after subdivision according to telomere length values. The cutoff used for high and low RTL was decided by exploring RTL data using Cox's analyses dividing the material by the median, mean, and quartiles. The highest hazard ratio 3.5 (95% confidence interval, 1.1–11.2) was found for patients having RTL over the mean value (RTL = 0.47), representing the upper third of patients with BCP-ALL. In the low-risk BCP group, patients with long RTL (i.e. greater than the mean value) had inferior outcomes compared to patients with short RTL ($p = 0.022$, Fig. 2). For the high-risk BCP group ($n = 31$), no survival difference was found with regard to telomere length. In the total material, RTL was not a statistically significant parameter regarding outcomes. However, if t(9;22) cases (a group with very short telomeres, Table 1) were excluded from the survival analysis, a similar pattern emerged as for the low-risk BCP group, i.e., worse outcomes were found for patients with longer telomeres ($p = 0.046$, $n = 124$; not shown in figures).

Finally, because RTL is measured as a continuous variable, a Cox regression analysis was performed using

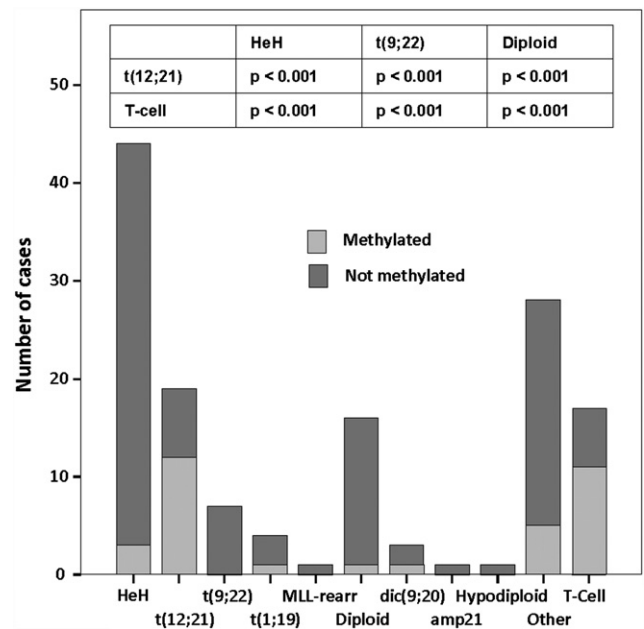


Figure 3. Frequency of *hTERT* promoter methylation in different cytogenetic BCP subgroups and in T-ALL. The insert table shows that t(12;21) and T-ALL cases differed significantly from HeH, t(9;22) and diploid ALL regarding *hTERT* methylation frequency.

noncategorized RTL values. In this setting, RTL had no predictive value on outcomes.

hTERT promoter methylation frequency differed between subgroups of ALL

Thirty-four of 147 patients investigated regarding methylation status of the *hTERT* core promoter were positive (23.1%). In methylated samples, the degree of methylation was regularly $>70\%$ (percent methylated reference). All samples taken in remission at cessation of therapy revealed an unmethylated *hTERT* promoter. There was a significant difference in the frequency of methylated samples comparing the BCP group ($n = 130$; 17.7% methylated) with the T-ALL cases ($n = 17$; 64.7% methylated) ($p < 0.001$).

Further subdivision based on cytogenetic data revealed that HeH cases ($n = 42$) were methylated in only 7% compared to 63% of the t(12;21) cases ($n = 19$) ($p < 0.001$). In line with these data, HeH cases classified according to flow cytometric DNA index were significantly less methylated at the *hTERT* promoter compared to diploid cases ($p = 0.006$). All t(9;22) cases were unmethylated and both the diploid and “other aberrations” groups demonstrated low *hTERT* methylation frequency. The data are summarized in Figure 3. *hTERT* methylation was not associated to clinical outcomes.

hTERT promoter methylation and telomere length

There was no association between *hTERT* promoter methylation and telomere length in the BCP or in the T-ALL groups. However, in the t(12;21) subgroup, the 12 methylated

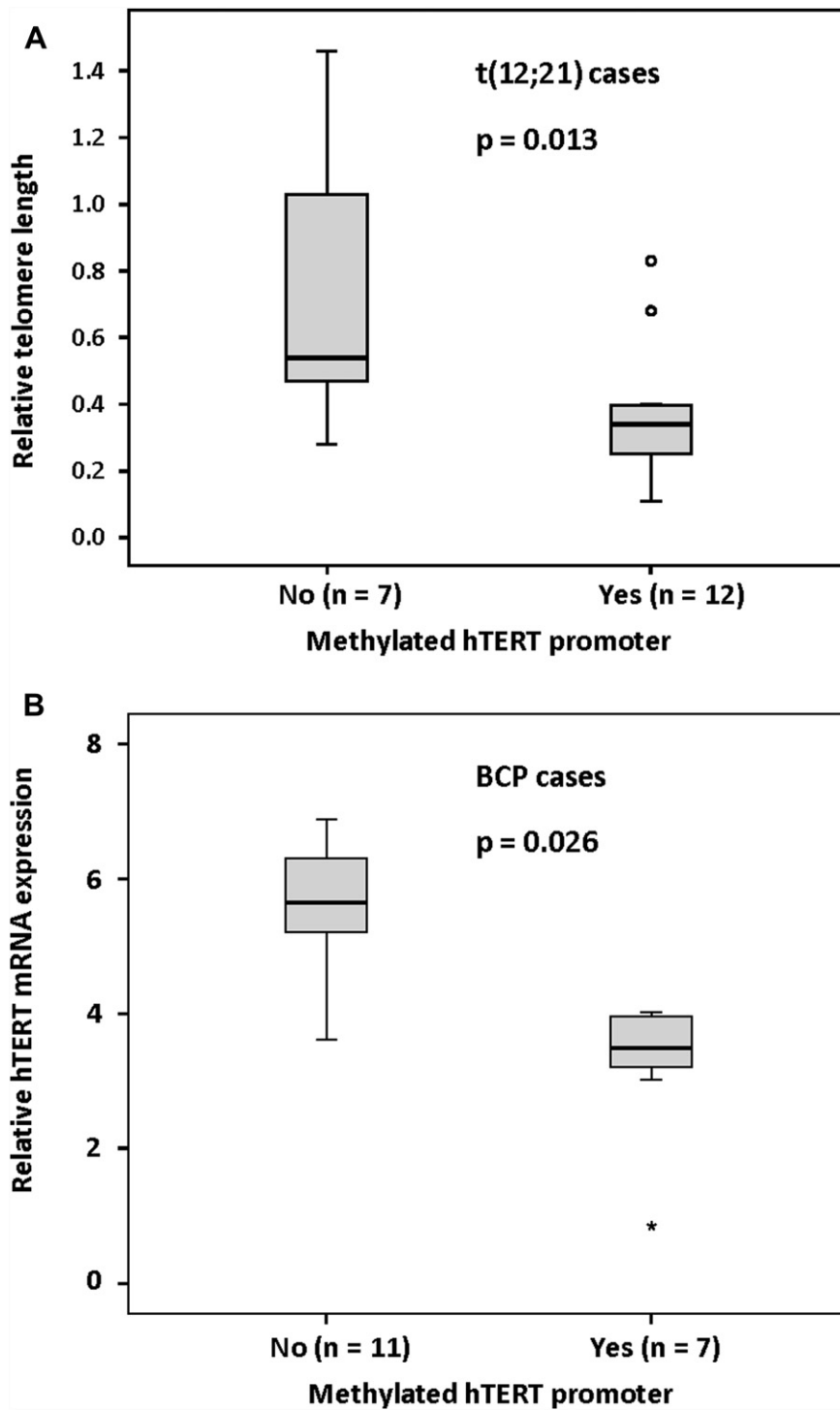


Figure 4. (A) Relative telomere length in relation to *hTERT* methylation in t(12;21) ALL cases. The box plot presents the minimum, lower quartile, median, upper quartile, and maximum RTL, except for outliers. (B) Relative *hTERT* gene expression normalized against β -actin in relation to *hTERT* methylation in BCP-ALL cases. The box plot show the minimum, lower quartile, median, upper quartile, and maximum expression of *hTERT*.

cases had significantly shorter telomeres compared to the seven nonmethylated cases ($p = 0.013$; Fig. 4A). The other subgroups were either too small or had too low methylation frequency for adequate statistical analysis.

hTERT RNA expression was lower in methylated BCP samples

We tested if *hTERT* promoter methylation did effect *hTERT* gene expression in a subset of 18 BCP-ALL cases with RNA

available for analysis. By alternative splicing, hTERT RNA can appear as different isoforms and the full-length variant is believed to generate the active enzyme. Primers for both the full-length variant and primers detecting variants of hTERT were used. All cases showed a high concordance between the expression of full-length hTERT RNA and of full-length + α/β splice variants ($p < 0.001$). The BCP cases with methylated *hTERT* promoters ($n = 11$) had significantly lower levels of *hTERT* expression compared to unmethylated cases ($n = 7$) ($p = 0.026$; Fig. 4B). HeH cases had significantly higher hTERT RNA expression in comparison with other BCP cases ($p = 0.007$). Accordingly, there was a significant correlation between flow cytometric DNA index and full-length *hTERT* gene expression ($p = 0.009$).

Discussion

In the present study, we have investigated two aspects of telomere biology, i.e., telomere length and *hTERT* gene methylation, in a series of childhood ALL. Telomeres were shorter at ALL diagnosis compared to at cessation of therapy, suggesting that the leukemia blasts had shorter telomeres than normal hematopoietic stem cells from which the bone marrow cells at remission originated. This is in line with previous data focusing on acute myeloid leukemia and chronic myeloid leukemia (CML), showing shorter telomeres in leukemic blasts compared to in normal blast and progenitor cells (reviewed in [23]). A possible explanation for this finding is a difference in the proliferative history (i.e., the number of cell cycle rounds) of leukemic vs normal blasts and progenitors. There was a trend toward longer telomeres in our T-ALL cases compared to BCP, however, this did not reach significance ($p = 0.063$). A recent study actually showed longer telomeres in T-ALL compared to B-ALL in adult patients [24].

Cases with t(9;22)(q11;q34)[BCR/ABL1] where completely unmethylated and had the shortest telomeres. Translocation t(9;22) is a hallmark of CML. Earlier studies reported that CML cells had shorter telomeres compared to Ph-translocation–negative T cells from the same patient, probably due to accelerated proliferation of the CML cells [25]. This explanation seemed unlikely in our material because t(9;22) cases had lower S-phase fractions than other cytogenetic subgroups at diagnosis (our unpublished data). However, we cannot rule out a higher proliferation rate at earlier stages of the disease. Apart from short telomeres, elevated expression of *hTERT* and telomerase activity have also been reported in CML cells, in line with our results with unmethylated promoter of t(9;22)-positive ALL samples. Of interest is that Drummond et al. observed lower hTR levels in CML cells compared to non-CML CD34⁺ cells, despite higher telomerase activity [26], which could in part explain short telomere length. Regarding these cases, it should be noted that they were few and a larger set of patients is needed to verify this finding.

Furthermore, telomere length seemed to be a marker for clinical outcomes in ALL. When studying prognostic indicators, the treatment has to be taken into account and thus the analysis should be performed in subgroups given similar treatment schedules. In the largest ALL subgroup, low-risk BCP, we found a better prognosis for cases with short telomeres. This is in contrast to previous reports on myelodysplastic syndrome, myeloma, and chronic lymphocytic leukemia, all showing shorter telomeres to be associated with progressive disease and a worse prognosis [20,27,28]. Regarding ALL, no previous data are available for comparison. A possible mechanism underlying our finding is that short telomeres are associated with enhanced chromosomal instability. In vitro experiments on solid tumor cell lines show that telomerase inhibition leading to telomere shortening sensitizes cells to anticancer agents [29]. Hence cancer cells with long telomeres might tolerate anticancer agents better. Telomere length did not predict outcomes in high-risk leukemia patients. The reason for this is probably the fact that BCP-ALL–positive for t(9;22) had shorter telomeres than all other groups. When the t(9;22) cases were excluded from the analysis, long telomeres were associated with worse outcomes for all ALL patients (data not shown). Further studies in other ALL cohorts are needed in order to verify this observation.

In the present study, we also investigated *hTERT* promoter methylation, showing that nearly one fourth of the tumor samples had a methylated promoter, whereas all normal bone marrow samples were unmethylated. However, there were large differences between the ALL subgroups regarding methylation incidence. *hTERT* methylation was more prevalent in T-ALL cases as compared to in BCP-ALL (72% vs 18%). However, within the BCP group, a significantly higher methylation rate was found in the t(12;21) (63%) compared to in the HeH cases (7%). These results strongly indicate that the *hTERT* promoter methylation status defines specific ALL subgroups and may have a functional relevance. The t(12;21) translocation gives rise to the ETV6/RUNX1 fusion protein and these leukemias have a similar immunophenotype as well as a similar gene expression pattern, indicating a defined molecular pathology [30]. There are indications that ETV6/RUNX1 can affect the chromatin state, even if direct evidence for involvement in DNA methylation metabolism is lacking [31]. Of interest is that other fusion genes involving *RUNX1(AML1)* (*AML1-ETO*) can interact with methyl transferases and alter the DNA methylation state [32]. A recent study even showed that the AML1-ETO fusion protein could regulate *hTERT* expression; however, probably not by binding to the *hTERT* promoter [33]. It would be interesting to further examine if the ETV6/RUNX1 fusion protein is directly involved in remodeling of the *hTERT* promoter.

It was recently shown that specific leukemic fusions genes, MLL/AF4 in infant ALL and AML1/MTG8, can control expression of hTERT [34]. Regarding MLL/AF4, data indicated that this effect was mediated by a direct

promoter effect of the HOXA7 transcription factor. There were no MLL rearranged ALL cases in our series, but these data illustrate that hTERT is regulated at multiple levels.

The general concept that promoter methylation suppresses gene transcription could not be demonstrated in the first studies of the *hTERT* gene [15,16,33], and even a positive correlation between methylation and gene expression was reported [35]. However, later studies have shown a methylation-dependent gene repression when the core promoter region close to the *hTERT* start site was analyzed [15]. This is in agreement with our data showing a significant, negative correlation between *hTERT* methylation and gene expression. T-ALL case had more frequent events of hTERT methylation compared to BCP cases. An earlier study found that T-ALL had lower expression *hTERT* compared to BCP cases [36], in line with our results. A recent study confirmed these results on adult T-ALL and B-ALL. This study also showed a significant correlation between *hTERT* levels and telomerase activity [24].

We do not know if methylation of the *hTERT* promoter arises early during leukemogenesis in t(12;21) and T-cell leukemia, or if it is a later event during disease progression. If promoter methylation and subsequent downregulation of hTERT occur early, it could facilitate telomere attrition and promote genome instability and thereby contribute to transformation.

Previous studies on methylation profiles have revealed methylation hotspots in ALL DNA compared to normal DNA. Differences between leukemias of B- and T-cell origin in DDX51 gene methylation have been described [8]. Other studies have revealed a general pattern with more methylated promoters in adult T-ALL vs B-ALL [37]. However, no gene showed such a strong difference in methylation frequency as hTERT in the present study.

Collected data indicate that epigenetic mechanisms are clinically important in ALL and associated with outcomes. A connection between survival and methylation status was described for CpG island methylator phenotype—positive t(12;21) ALL, showing worse outcomes compared to CpG island methylator phenotype—negative cases. The same was true for T-cell ALL [9,10]. In the present study, patients were too few in each subgroup to properly investigate a relationship between survival and *hTERT* methylation status. Future studies on the biological and clinical significance of *hTERT* methylation in ALL should be focused on t(12;21) and T-ALL cases, two subgroups characterized by a high degree of *hTERT* methylation. Further detailed characterization of the chromatin landscape, including DNA methylation, is important in order to identify patients who might benefit from treatment with chromatin-modifying drugs.

Conflict of interest disclosure

No financial interest/relationships with financial interest relating to the topic of this article have been declared.

Funding disclosure

Supported by grants from the Swedish Childhood Cancer Foundation, Swedish Cancer Society, the Medical Faculty, Umeå University, and Lion's Cancer Research Foundation, Umeå. The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2011 under grant agreement no. 200950.

References

1. Kurz DJ, Decary S, Hong Y, Trivier E, Akhmedov A, Erusalimsky JD. Chronic oxidative stress compromises telomere integrity and accelerates the onset of senescence in human endothelial cells. *J Cell Sci.* 2004;117:2417–2426.
2. von Zglinicki T. Oxidative stress shortens telomeres. *Trends Biochem Sci.* 2002;27:339–344.
3. Hug N, Lingner J. Telomere length homeostasis. *Chromosoma.* 2006; 115:413–425.
4. Anderson CJ, Hoare SF, Ashcroft M, Bilstrand AE, Keith WN. Hypoxic regulation of telomerase gene expression by transcriptional and post-transcriptional mechanisms. *Oncogene.* 2005;25:61–69.
5. Esteller M. Epigenetics in cancer. *N Engl J Med.* 2008;358:1148–1159.
6. Galm O, Herman JG, Baylin SB. The fundamental role of epigenetics in hematopoietic malignancies. *Blood Rev.* 2006;20:1–13.
7. Rice KL, Hormaeche I, Licht JD. Epigenetic regulation of normal and malignant hematopoiesis. *Oncogene.* 2007;26:6697–6714.
8. Taylor KH, Pena-Hernandez KE, Davis JW, et al. Large-scale CpG methylation analysis identifies novel candidate genes and reveals methylation hotspots in acute lymphoblastic leukemia. *Cancer Res.* 2007;67:2617–2625.
9. Roman-Gomez J, Jimenez-Velasco A, Agirre X, Prosper F, Heiniger A, Torres A. Lack of CpG island methylator phenotype defines a clinical subtype of T-cell acute lymphoblastic leukemia associated with good prognosis. *J Clin Oncol.* 2005;23:7043–7049.
10. Roman-Gomez J, Jimenez-Velasco A, Agirre X, Castillejo JA, Navarro G, Calasanz MJ, et al. CpG island methylator phenotype redefines the prognostic effect of t(12;21) in childhood acute lymphoblastic leukemia. *Clin Cancer Res.* 2006;12:4845–4850.
11. Atkinson SP, Hoare SF, Glasspool RM, Keith WN. Lack of telomerase gene expression in alternative lengthening of telomeres is associated with chromatin remodeling of the hTR and hTERT gene promoters. *Cancer Res.* 2005;65:7585–7590.
12. Guilleret I, Benhattar J. Unusual distribution of DNA methylation within the hTERT CpG island in tissues and cell lines. *Biochem Biophys Res Commun.* 2004;325:1037–1043.
13. Oikonomou P, Messinis I, Tsezou A. DNA methylation is not likely to be responsible for hTERT expression in premalignant cervical lesions. *Exp Biol Med.* 2007;232:881–886.
14. Shin K-H, Kang MK, Dicterow E, Park N-H. Hypermethylation of the *hTERT* promoter inhibits the expression of telomerase activity in normal oral fibroblasts and senescent normal oral keratinocytes. *Br J Cancer.* 2003;89:1473–1478.
15. Zinn RL, Pruitt K, Eguchi S, Baylin SB, Herman JG. hTERT is expressed in cancer cell lines despite promoter DNA methylation by preservation of unmethylated DNA and active chromatin around the transcription start site. *Cancer Res.* 2007;67:194–201.
16. Devereux TR, Horikawa I, Anna CH, Annab LA, Afshari CA, Barrett JC. DNA methylation analysis of the promoter region of the human telomerase reverse transcriptase (hTERT) gene. *Cancer Res.* 1999; 59:6087–6090.
17. Bechter OE, Eisterer W, Dlaska M, Kühn T, Thaler J. CpG island methylation of the hTERT promoter is associated with lower telomerase activity in B-cell lymphocytic leukemia. *Exp Hematol.* 2002;30:26–33.

18. Svenson U, Roos G. Telomere length as a biological marker in malignancy. *Biochim Biophys Acta*. 2009;1792:317–323.
19. Gustafsson G, Schmiegelow K, Forestier E, et al. Improving outcome through two decades in childhood ALL in the Nordic countries: the impact of high-dose methotrexate in the reduction of CNS irradiation. *Nordic Society of Pediatric Haematology and Oncology (NOPHO). Leukemia*. 2000;14:2267–2275.
20. Grabowski P, Hultdin M, Karlsson K, et al. Telomere length as a prognostic parameter in chronic lymphocytic leukemia with special reference to VH gene mutation status. *Blood*. 2005;105:4807–4812.
21. Weisenberger DJ, Campan M, Long TI, et al. Analysis of repetitive element DNA methylation by MethyLight. *Nucl Acids Res*. 2005;33:6823–6836.
22. Weisenberger DJ, Siegmund KD, Campan M, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet*. 2006;38:787–793.
23. Drummond MW, Balabanov S, Holyoake TL, Brummendorf TH. Concise review: telomere biology in normal and leukemic hematopoietic stem cells. *Stem Cells*. 2007;25:1853–1861.
24. Capraro V, Zane L, Poncet D, et al. Telomere deregulations possess cytogenetic, phenotype and prognostic specificities in acute leukemias. *Exp Hematol*. 2011;39:195–202.
25. Brummendorf TH, Holyoake TL, Rufer N, et al. Prognostic implications of differences in telomere length between normal and malignant cells from patients with chronic myeloid leukemia measured by flow cytometry. *Blood*. 2000;95:1883–1890.
26. Drummond MW, Hoare SF, Monaghan A, et al. Dysregulated expression of the major telomerase components in leukaemic stem cells. *Leukemia*. 2005;19:381–389.
27. Sieglóvá Z, Zilovcová S, Cermák J, et al. Dynamics of telomere erosion and its association with genome instability in myelodysplastic syndromes (MDS) and acute myelogenous leukemia arising from MDS: a marker of disease prognosis? *Leuk Res*. 2004;28:1013–1021.
28. Wu KD, Orme LM, Shaughnessy J Jr, Jacobson J, Barlogie B, Moore MAS. Telomerase and telomere length in multiple myeloma: correlations with disease heterogeneity, cytogenetic status, and overall survival. *Blood*. 2003;101:4982–4989.
29. Cerone MA, Londoño-Vallejo JA, Autexier C. Telomerase inhibition enhances the response to anticancer drug treatment in human breast cancer cells. *Mol Cancer Ther*. 2006;5:1669–1675.
30. Andersson A, Olofsson T, Lindgren D, et al. Molecular signatures in childhood acute leukemia and their correlations to expression patterns in normal hematopoietic subpopulations. *Proc Natl Acad Sci U S A*. 2005;102:19069–19074.
31. Zelent A, Greaves M, Enver T. Role of the TEL-AML1 fusion gene in the molecular pathogenesis of childhood acute lymphoblastic leukaemia. *Oncogene*. 2004;23:4275–4283.
32. Fazi F, Racanicchi S, Zardo G, et al. Epigenetic silencing of the myelopoiesis regulator microRNA-223 by the AML1/ETO Oncoprotein. *Cancer Cell*. 2007;12:457–466.
33. Dessain SK, Yu Hy, Reddel RR, Beijersbergen RL, Weinberg RA. Methylation of the human telomerase gene CpG island. *Cancer Res*. 2000;60:537–541.
34. Gessner A, Thomas M, Castro PG, et al. Leukemic fusion genes MLL/AF4 and AML/MTG8 support leukemic self-renewal by controlling expression of the telomerase subunit TERT. *Leukemia*. 2010;24:1751–1759.
35. Guilleret I, Yan P, Grange F, Braunschweig R, Bosman T, Benhattar J. Hypermethylation of the human telomerase catalytic subunit hTERT gene correlates with telomerase activity. *Int J Cancer*. 2002;101:335–341.
36. Cogulu O, Kosova B, Karaca E, et al. Evaluation of telomerase mRNA (hTERT) in childhood acute leukemia. *Leuk Lymphoma*. 2004;45:2477–2480.
37. Martin-Subero JI, Ammerpohl O, Bibikova M, et al. A comprehensive microarray-based DNA methylation study of 367 hematological neoplasms. *PLoS One*. 2009;4:e6986.

Promoter DNA Methylation Pattern Identifies Prognostic Subgroups in Childhood T-Cell Acute Lymphoblastic Leukemia

Magnus Borssén¹, Lars Palmqvist², Kristina Karrman³, Jonas Abrahamsson⁴, Mikael Behrendtz⁵, Jesper Heldrup⁶, Erik Forestier¹, Göran Roos¹, Sofie Degerman^{1*}

1 Department of Medical Biosciences, Pathology, Umeå University, Umeå, Sweden, **2** Department of Clinical Chemistry and Transfusion Medicine, Sahlgrenska University Hospital, University of Gothenburg, Gothenburg, Sweden, **3** Department of Clinical Genetics, University and Regional Laboratories, Skåne University Hospital, Lund University, Lund, Sweden, **4** Institute of Clinical Sciences, Department of Pediatrics, Sahlgrenska University Hospital, Gothenburg, Sweden, **5** Department of Pediatrics, Linköping University Hospital, Linköping, Sweden, **6** Department of Pediatrics, Skåne University Hospital, Lund, Sweden

Abstract

Background: Treatment of pediatric T-cell acute lymphoblastic leukemia (T-ALL) has improved, but there is a considerable fraction of patients experiencing a poor outcome. There is a need for better prognostic markers and aberrant DNA methylation is a candidate in other malignancies, but its potential prognostic significance in T-ALL is hitherto undecided.

Design and Methods: Genome wide promoter DNA methylation analysis was performed in pediatric T-ALL samples (n = 43) using arrays covering >27000 CpG sites. Clinical outcome was evaluated in relation to methylation status and compared with a contemporary T-ALL group not tested for methylation (n = 32).

Results: Based on CpG island methylator phenotype (CIMP), T-ALL samples were subgrouped as CIMP+ (high methylation) and CIMP- (low methylation). CIMP- T-ALL patients had significantly worse overall and event free survival (p = 0.02 and p = 0.001, respectively) compared to CIMP+ cases. CIMP status was an independent factor for survival in multivariate analysis including age, gender and white blood cell count. Analysis of differently methylated genes in the CIMP subgroups showed an overrepresentation of transcription factors, ligands and polycomb target genes.

Conclusions: We identified global promoter methylation profiling as being of relevance for subgrouping and prognostication of pediatric T-ALL.

Citation: Borssén M, Palmqvist L, Karrman K, Abrahamsson J, Behrendtz M, et al. (2013) Promoter DNA Methylation Pattern Identifies Prognostic Subgroups in Childhood T-Cell Acute Lymphoblastic Leukemia. PLoS ONE 8(6): e65373. doi:10.1371/journal.pone.0065373

Editor: Carmen J. Marsit, Dartmouth Medical School, United States of America

Received: October 16, 2012; **Accepted:** April 26, 2013; **Published:** June 6, 2013

Copyright: © 2013 Borssén et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by grants from the Swedish Cancer Society (GR), the Swedish Research Council (GR), the Swedish Childhood Cancer Foundation (GR, EF), the Medical Faculty, Umeå University (GR) and Lion's Cancer Research Foundation, Umeå (SD, MBo, GR). The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007–2011 under grant agreement no. 200950 (GR). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: sofie.degerman@medbio.umu.se

Introduction

Acute lymphoblastic leukemia (ALL) is the most common malignancy among children. A majority of the ALL cases originates from the B-cell compartment and is characterized by structural genetic aberrations. However, about 15% derive from the T-cell lineage (T-ALL). Over the past 30 years a remarkable improvement in ALL treatment has occurred, but the therapy success rate for T-ALL is lower than for precursor B-cell lymphoblastic leukemia. T-ALL harbors a multitude of genetic alterations, frequently associated with rearrangements of T cell antigen receptor genes with involved fusion genes, such as *TLX1*, *LMO1*, *LMO2*, *TAL1* and *LYL1* [1,2]. Mutations of *NOTCH1*, *FBXW7*, *PTEN* and *WT1* have also been reported in several studies [1,3]. These aberrations have provided new insights into the malignant transformation of T-cells, but the prognostic information given by the individual genetic abnormalities is

unclear or controversial [4]. Due to the lack of good prognostic markers all T-ALL patients have been enrolled in high intensive treatment programs with substantial therapy related side effects.

Apart from classical genetic aberrations, epigenetics has emerged as an important regulator of cell functions that has been implicated in cancer development [5]. The term “epigenetics” refers to control of gene expression without affecting the actual genetic sequence due to mechanisms like DNA methylation and post translational histone modifications. DNA methylation occurs on CpG (cytosine preceding a guanine) nucleotides distributed across the entire genome but are enriched in CpG islands (CGI). Transformation from a benign to a neoplastic cell is generally associated with a decrease in global CpG site methylation concomitant with an increase in promoter associated CGI methylation. Global demethylation is associated with chromosomal instability and gain of methylation at promoters might lead to silencing of tumor suppressor genes [6].

Further subclassification of hematological malignancies, besides morphologic, immunophenotypic, and genetic features, is of increasing importance in parallel with the development of more target directed therapies. In this context DNA methylation is of potential interest and has appeared as a candidate for subtype discrimination in breast cancer, glioma and colon cancer [7–9]. In many studies, methylation status of a limited number of promoter associated CpG sites or promoter regions have been used to define tumors as CGI methylator phenotype positive (CIMP+) or negative (CIMP–). In myelodysplastic syndrome (MDS) CIMP+ patients had increased risk of malignant transformation, shorter survival and could benefit from DNA methyltransferase inhibitor therapy [10,11]. In ALL, including T-ALL, the methylation pattern of certain CpG site positions in selected promoters were associated with treatment outcome, indicating that methylation on some of these CpG sites was associated with a worse outcome [12]. A Nordic study identified a set of CpG sites discriminating between different forms of ALL based on their methylation status and an association to prognosis was shown in subgroups of B-ALL, but not for T-ALL [13]. Noteworthy, Deneberg and coworkers recently demonstrated, that cytogenetically normal AML cases with a globally methylated promoter profile had a significantly better survival compared with less methylated cases [14]. Our knowledge about global promoter DNA methylation profiles in T-ALL is sparse and its possible role in prognostication and/or subgrouping has not been established.

In the present study we characterized the DNA methylation pattern in diagnostic pediatric T-ALL samples using genome-wide promoter focused methylation arrays identifying two subgroups with highly significant differences in clinical outcome.

Materials and Methods

Patient and Control Samples

Between January 1, 1992 and June 30, 2008, 75 infants, children, and adolescents <18 years were diagnosed with T-ALL at the Swedish regional pediatric oncologic centers in Lund (N = 18), Göteborg (N = 34), Linköping (N = 8) and Umeå (N = 15). 46 patients were treated according to the NOPHO (Nordic Society of Pediatric Hematology and Oncology) ALL 1992 protocol and 29 patients according to the NOPHO ALL 2000 protocol [15,16]. Diagnostic bone marrow material was available for DNA analysis from 43 of the 75 patients (57%) and was included in the promoter methylation investigation. The Regional Ethics Committee in Umeå, Lund and at the Karolinska Institute in Sweden approved the study, and the patients and/or their guardians provided informed consent and during the later years written informed consent. All samples were anonymized in figures and tables. The study was conducted in accordance with the Declaration of Helsinki. The T-ALL diagnosis was in each case based on morphology and flow cytometric immunophenotyping. Cytogenetic aberrations were established by G-band karyotyping and over time FISH analysis was increasingly used.

DNA from age-matched mononuclear normal bone marrow (n = 3) cells, and mitogen (WGA) stimulated primary lymphoblast T-cell cultures (L2 at 5 population doublings (PD) and P7/R2- at 14 PD) derived from mononuclear cells of healthy adult donors were included as controls [17,18].

Genome Wide Promoter CpG Methylation Profiling

Genomic DNA from diagnostic T-ALL samples, normal bone marrow and T-cell cultures were isolated by standard procedures and DNA purity and concentration were determined by spectrophotometry (NanoDrop, Thermo Scientific, Wilmington DE,

USA). The DNA (500 ng) was bisulfite treated and purified according to manufacturer's protocol (Zymo EZ DNA methylation kit, Zymo Research, CA, USA). M.SssI-modified DNA (all cytosines methylated) was included as a methylated reference and standard curve in the MethyLight assays. The bisulfite conversion was confirmed using the MethyLight technology for amplification and detection of methylated DNA regions in a iQTM5 Real-Time PCR Detection System (BioRad laboratories, Hercules, CA, USA). A C-LESS-C1 reaction was performed as described by Weisenberger *et al.* 2008 and with the specified primers and probes [19]. Since the C-LESS reaction amplifies both template strands of unconverted genomic DNA but only one strand of bisulfite-converted DNA, the bisulfite-converted DNA will amplify one PCR cycle later than unconverted DNA. In parallel, the *ALU* gene was analyzed with the methylation independent ALU-C4 primer/probe sets [20]. The ALU and C-LESS-C1 PCR reactions were run at 95°C for 10 minutes, followed by 50 cycles of 95°C 15 s, 60°C 60 s. All samples passed these preliminary control analyses before they were applied on the arrays.

Genome-wide promoter CpG site methylation profiling was performed using the Illumina Infinium HumanMeth27K BeadArray (Illumina Inc., San Diego, USA). These arrays generate data for 27578 CpG nucleotides, corresponding to 14473 individual gene promoter regions spanning from –1,000 to +500 from the transcription start site [21]. CpG nucleotides are preferentially located within CpG islands and the majority of CpG sites (~70%) on the array were classified as being within a CpG island as defined by Takai and Jones relaxed criteria [22]. The methylation assay was performed according to Illumina instructions. Briefly, bisulfite converted DNA was fragmented and amplified with supplied reagents and the pre-processed DNA was hybridized to the arrays. Each CpG site on the array was represented by two site-specific probes, one designed for the methylated locus (M) and another for the unmethylated locus (U). Single-base extension of the probes incorporates labeled ddNTP, which subsequently is stained with a fluorescence reagent. The arrays were scanned using a BeadArray Reader (Illumina). Image processing and intensity data extraction were performed according to Illumina instructions using the GenomeStudio software (V2010.3) with the methylation analysis module (version 1.8.5). The methylation level for the interrogated locus was determined by calculating the ratio (β value) of the fluorescent signals from the methylated (M) vs unmethylated (U) sites. $\beta = \text{Max}(M,0) / (\text{Max}(M,0) + \text{Max}(U,0) + 100)$ ranging in theory from 0, corresponding to completely unmethylated, to 1, representing fully methylated DNA.

As recommended by Illumina technical support, data were not normalized due to the nature of the β -value calculated as a ratio of the methylated and the unmethylated signal. Empirical testing of background normalization had minute effects on the results. CpG sites that failed (detected in ≤ 2 beads/array) in one or several arrays were omitted from the analysis. Furthermore, to avoid gender bias, all CpGs located on the X or Y chromosomes were excluded. Analysis was subsequently restricted to the remaining 26436 CpG sites. Unsupervised and supervised hierarchical clustering analysis with the Euclidean distance algorithm was used to create heat maps and dendograms. In order to identify the most variable methylated probes within the T-ALL group a standard deviation ≥ 0.30 of the β -value across the samples was used as cut-off. A summary over the CpG selection steps is shown in Figure 1. Methylation array data has been deposited to the NCBI Gene Expression Omnibus (GEO) database (GSE42079).

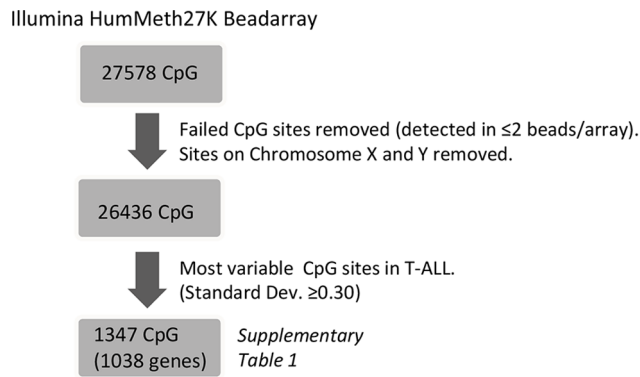


Figure 1. Schematic overview of the methylation array analysis.

doi:10.1371/journal.pone.0065373.g001

Verification of Array Data

A subset of 10 T-ALL samples was also separately analyzed at the SNP&SEQ Technology Platform laboratory in Uppsala, Sweden, using a high density array (485577 CpG sites) (HumanMeth450K bead array, Illumina Inc., San Diego, USA), including >90% of the CpG sites on the HumanMeth27K array. A methylated control sample (Human HCT116 DKO methylated DNA, Zymo Research) enzymatically methylated on all cytosines by M.SssI methyltransferase, and a non-methylated (<5%) control double knocked-out for DNA methyltransferases (DNMT1 $-/-$ and DNMT3B $-/-$) (Human HCT116 DKO non-methylated DNA, Zymo Research), were included. A replicate sample was also included to assess inter-assay reproducibility ($R^2 = 0.989$). For each sample, 500 ng DNA was bisulfite converted with the EZ-96 DNA Methylation-GoldTM Kit (Zymo Research) according to the instruction manual. 200 ng of bisulfite-converted DNA was applied to each array according to the Illumina HumanMeth450K protocol, arrays were scanned and data analyzed with the GenomeStudio Software.

In addition, nine T-ALL samples were sent to Genome Centre Queen Mary, University of London for targeted pyrosequencing on a selection of gene promoter regions overlapping with specific CpGs in the methylation array, including TAL1 (cg19797376), KLF4 (cg07309102) and HOXD8 (cg15520279). In brief, pyrosequencing was performed according to manufacturers protocols by bisulfite treatment of DNA (EZ DNA Methylation, Zymo Research), followed by PCR amplification and pyrosequencing using PyroMark Gold Q96 Reagents (Qiagen) in the PSQ 96MA instrument with PSQ 96MA software V2.1 (Qiagen). The following set of primers were used: TAL1_F:ATGGGGGT-TAGAGAGAGAATGA; TAL1_R:ACCTCCTCAACCAAT-CTC; TAL1_seq:GGGGGATTTTAAGGT; HOXD8_F:AGT-GATAGTAGTAGTAAGTGGGATTGAT; HOXD8_R:AACA-ACCCCCCACAACCCCC; HOXD8_seq:GTTTTGTATTT-GGAGTATAG; KLF4_F:AGGTTGTAGAGAAGTAAGTTA-TAAGTAAG; KLF4_R:CAACAACCTCCCCACCACTAT; KLF4_seq:ATACCCCCAAATAAACTAACTAC.

Gene Expression Array Analysis and q-RT-PCR Verification

Total RNA was isolated from 17 frozen bone marrow samples using TRIZOL Reagent (Invitrogen, Stockholm, Sweden), according to the manufacturer's protocol. Concentrations of total RNA were determined by spectrophotometry (NanoDrop, Thermo Scientific, Wilmington, DE, USA) and RNA quality was analyzed (Agilent 2100 Bioanalyzer, Agilent Technologies,

Santa Clara, CA, USA). 500 ng of total RNA from each sample was used for cRNA production by the Illumina TotalPrep RNA amplification kit (Ambion Inc, St.Austin, TX, USA) according to the protocol provided. The biotinylated cRNA was quantified (NanoDrop), and the quality of cRNA was evaluated using the RNA 6000 pico kit (Agilent 2100 Bioanalyzer, Agilent Technologies).

A total of 750 ng biotinylated cRNA was used for hybridization to a human HT12 Illumina Beadchip gene expression array (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. The arrays were scanned using the Illumina BeadArray Reader (Illumina). The Illumina GenomeStudio software was used for data analysis and normalization by the quantile algorithm. Genes with signals below background level were excluded from the analysis, and differentially expressed genes (2-fold) between CIMP subgroups were identified by fold change calculations. Gene expression data was merged with methylation data in the GenomeStudio software and exported to Excel 2010 (Microsoft Office) for visualization. Gene expression array data has been deposited to the NCBI GEO database (GSE41621).

TAL1 gene expression was verified by q-RT-PCR analysis. Briefly, cDNA was prepared (Superscript III, Invitrogen, Sweden) and PCR performed using TaqMan (Applied Biosystems, Inc., Foster city, CA, USA) gene expression assays for TAL1 (Hs01097987_m1) and TBP (Hs00427620_m1) genes with the TaqMan universal mastermix II, according manufacturers protocol in the ABI Prism 7900HT instrument (Applied Biosystems). A cell line, CCRF-CEM (CCL-119, ATCC; Manassas, VA, USA) was used in 5-fold dilutions as standard curve and the relative TAL1 mRNA level was normalized to the housekeeping gene TBP.

Statistical and Bioinformatics Analysis

The Statistical Package for the Social Sciences (SPSS Inc., Chicago, IL) software for Macintosh 11 was used for the statistical analyses. Estimates of event free survival (pEFS) and overall survival (pOS) are given at 5 years and were calculated using the Kaplan-Meier method and the different subgroups listed in Table 1 were compared using the log rank test. In Table 1, the chi-square/Fisher's Exact test was used to compare differences between variables among subgroups; the Mann-Whitney U test was used for continuous variables. For multivariate analysis the Cox proportional-hazard regression model was used, significance evaluated with Log rank. Factors included in the equation were CIMP status, gender, age at diagnosis, and WBC (Table 2). The significance limit for two-sided p-values was set to <0.05 in all tests. Time in first remission (CR1) was defined as time from diagnosis until first event, comprising induction failure, relapse, death of disease, death in remission, or second malignant neoplasm. In the OS analysis, death of any cause was the endpoint. The NOPHO leukemia registry is updated annually and follow-up data were extracted from the registry as of February, 2011.

The bioinformatics analysis was performed with the GeneGO MetaCoreTM software (version 6.7) (Thomson Reuters, New York, USA) and the Database for Annotation, Visualization and Integrated Discovery (DAVID v6.7, <http://david.abcc.ncifcrf.gov/>) [23]. The MetaCore software was used for "Enrichment by protein function" analysis, cell signaling pathway and network analysis, and the MapEditor tool was used for visualization.

The polycomb target gene lists from Bracken et al. [24] and Lee et al. [25] were compared with the differently methylated gene (DMG) lists in our T-ALL samples using Bioconductor (R statistical programming language). We hypothesized that the

Table 1. Clinical characteristics and outcome in 75 T-ALL of which 43 cases were CIMP classified.

	CIMP+, N = 21	CIMP-, N = 22	Not analysed, N = 32	p value	Total, N = 75
Gender M/F	17/4	19/3	28/4	ns	64/11
Median age (range, years)	8.7 (2.4–16.6)	8.4 (1.6–16.8)	8.8 (2.4–16.8)	ns	8.7 (1.6–16.8)
WBC × 10 ⁹ /l (range)	73.0 (0.8–540)	141 (2.0–572)	51.8 (1.1–768)	ns	73.0 (0.8–768)
Relapse (months to relapse)	2 (10–30)	14 (3–91)	6 (3–28)		22 (3–91)
DCR1	1	0	5		6
Ind death	0	1	2		3
RD	0	0	1		1
Death	3	12	12		27
pEFS 5y (SD)	0.86 (0.08)	0.36 (0.10)	0.56 (0.09)	0.001	0.59 (0.06)
pOS 5y (SD)	0.86 (0.08)	0.45 (0.11)	0.62 (0.09)	0.02	0.63 (0.06)

doi:10.1371/journal.pone.0065373.t001

proportion of polycomb target genes in our DMG list was as many as expected and tested our hypothesis with a binominal test. Pearson's Chi-squared test was used to compare the distribution of the number of CpG sites in polycomb target genes on the HumMeth27K array in relation to the entire HumMeth27K array, and to compare the distribution of CpG sites located in polycomb and non-polycomb target genes on the array in relation to the differently methylated CpG sites in T-ALL.

Results

Analysis of Promoter DNA Methylation in Pediatric T-ALL Samples and Controls

Genome-wide promoter methylation status was determined by the Illumina HumanMeth27K array in 43 diagnostic pediatric T-ALL samples and 5 controls (normal bone marrow and primary T-cell cultures). The array covers 27578 CpG sites located in 14473 gene promoters, distributed on all chromosomes. CpG sites lacking an average beta value due to ≤ 2 reported beads/array, as well as CpGs on the X and Y chromosomes were omitted from the analysis to avoid gender-related methylation biases, and analysis was subsequently restricted to the remaining 26436 CpG sites (Figure 1).

The robustness of the results was confirmed by reanalyzing four diagnostic samples with the HumanMeth27K array giving R^2 -values ranging from 0.940 to 0.978 (data not shown). In addition, to independently verify the HumanMeth27K results, 10 T-ALL samples were analyzed with the HumanMeth450K methylation array showing that CpG sites overlapping between the arrays (25978 CpGs) correlated at R^2 values 0.932+/-0.022. Unsupervised Euclidean clustering of the overlapping CpG sites showed

identical clustering for samples analyzed on the HumanMeth27K and 450 K arrays, further substantiating the data obtained (Figure S1).

Furthermore, selected gene promoter regions for TAL1(cg19797376), KLF4 (cg07309102), and HOXD8 (cg15520279), were separately analyzed by pyrosequencing in nine T-ALL samples. The CpG sites analyzed on the HumMeth27K and HumMeth450K arrays were compared with pyrosequencing data on the same position (Figure S5 D, S6 A-B), showing a strong correlation between the methods.

The mean methylation level for the 26436 CpG sites was calculated for each sample. T-ALL samples gave a mean β value of 0.348 (range: 0.284–0.437, $n = 43$) to be compared with 0.272 (range 0.259–0.294, $n = 5$) for controls (data not shown). These results indicated a higher global promoter methylation of the T-ALL samples, but also a wide range in methylation levels within the T-ALL group. We focused our analysis on these differences within T-ALL in order to investigate if promoter DNA methylation might be a relevant factor in subgrouping T-ALL.

Differently Methylated Genes (DMGs) within the T-ALL Group

CpG sites with the largest variation in methylation level within the T-ALL group (standard deviation ≥ 0.3 , 1347 CpG sites) were selected for further examination (Table S1). Subsequent hierarchical cluster analysis of these 1347 CpGs subdivided the samples into three methylation groups, designated low ($n = 22$), intermediate ($n = 6$) and high ($n = 15$) (Figure 2A). The normal bone marrow samples and stimulated T-cells showed a low methylation profile and were similar to the CIMP- samples (Figure 2A). The distribution of mean CpG methylation levels (26436 CpG sites) differed significantly between the groups ($p \leq 0.001$, respectively) (Figure 2B). The deviation was more obvious for CpG sites located within CpG islands (CGI) ($n = 19211$) compared to CpGs outside CGIs ($n = 7225$) (Figure 2C–D). However, the methylation level was generally higher in CpGs outside islands. Seventy percent of the CpGs covered in the HumanMeth27K array were located within CGIs, but among the DMGs (1347 sites) a majority (98%) was located within CGIs, indicating a preferential methylation (or demethylation) of these sites.

Table 2. Multivariate analysis.

Factors in equation	EFS p value (HR)	OS p value (HR)
Gender	ns	ns
Age	ns	ns
WBC	ns	ns
CIMP status	0.006 (1.9)	0.044 (1.64)

doi:10.1371/journal.pone.0065373.t002

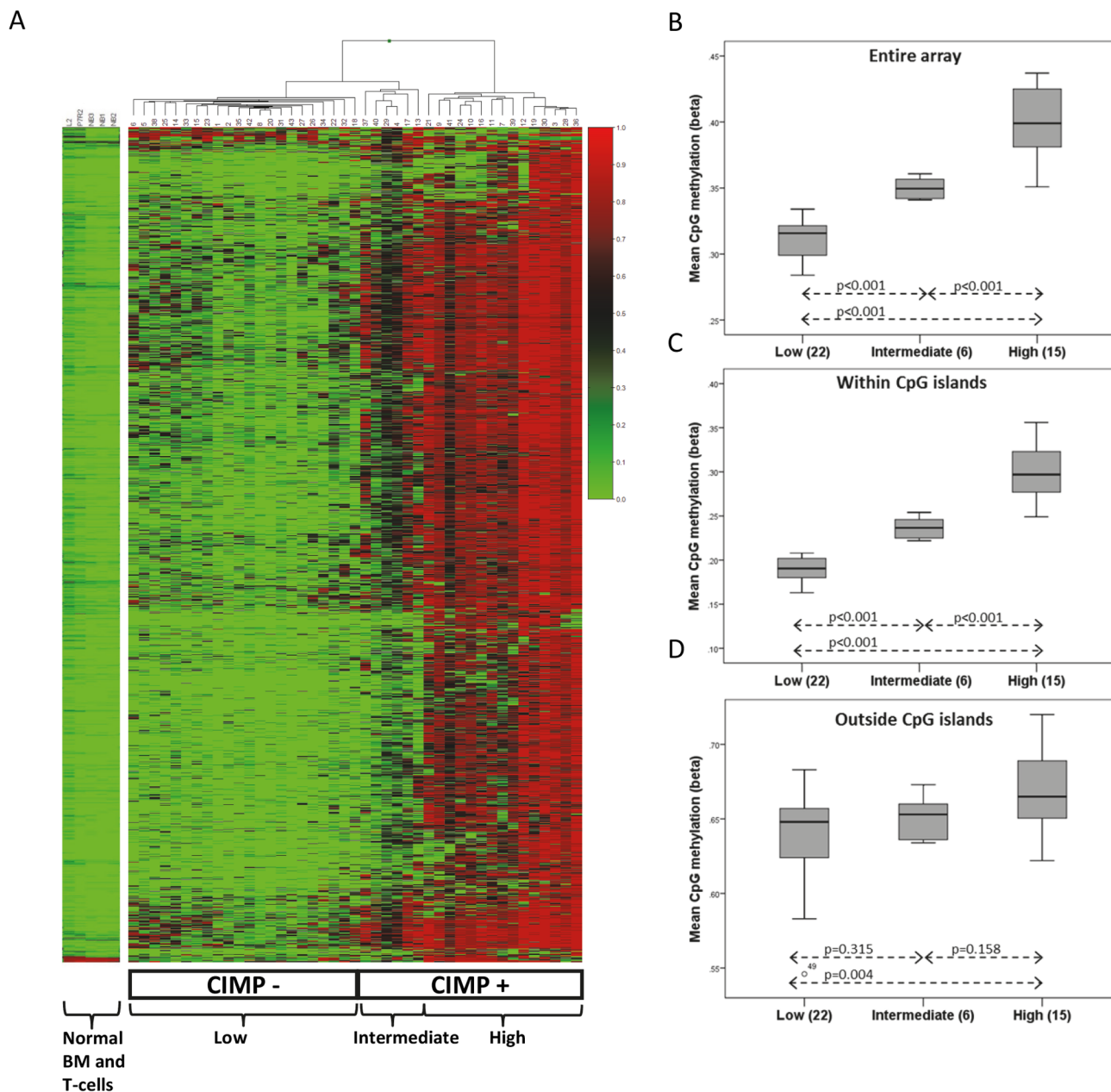


Figure 2. Hierarchical clustering of DNA methylation profiles. A) Unsupervised Euclidean hierarchical clustering was performed using beta values for the 1347 most variable CpG sites in T-ALL (standard deviation ≥ 0.30). In the dendrogram each individual is represented as a vertical row and the specific CpG sites are shown in the horizontal columns. High methylation levels are shown in red and low levels in green, according to the Beta value scale bar in the figure where 1.0 is fully methylated and 0 is unmethylated. The T-ALL methylation subtypes are marked in the figure as low, intermediate and high, and with the merged classification in CIMP- and CIMP+ cases. The methylation status of the 1347 CpG sites of normal bone marrow samples (N1–N3) and two stimulated primary T-cell lines (P7/R2 and L2) are shown next to the hierarchical cluster. B–D) The mean CpG methylation levels within and outside CpG islands were examined. The distribution of mean methylation levels in low ($n = 22$), intermediate ($n = 6$) and high ($n = 15$) methylation subtypes are shown for: B) the entire array (26436 sites), C) within CpG islands (19211 sites), D) outside CpG islands (7225 sites). The significance (independent samples t-test) for differences in methylation between groups is shown. doi:10.1371/journal.pone.0065373.g002

Treatment Outcome and Promoter Methylation at Diagnosis

The T-ALL samples included were diagnosed between years 1992 and 2008 and during this time period two different treatment protocols were used, NOPHO ALL 1992 and 2000. Event free and overall survival did not differ between these two protocols ($p = 0.99$, and $p = 0.3$, respectively) (Figure S2 A-B). The entire

cohort ($n = 75$) diagnosed with T-ALL during this time period had an EFS 5y of 0.59 (SE = 0.06) and OS 5y of 0.63 (SE = 0.06) with no difference for patients analyzed for methylation ($n = 43$) and cases not analyzed ($n = 32$) ($p = 0.9$ for both) (Figure 3A–B). Thereby we conclude that our study material was unbiased and had an expected clinical outcome for T-ALL patients during this time period [16].

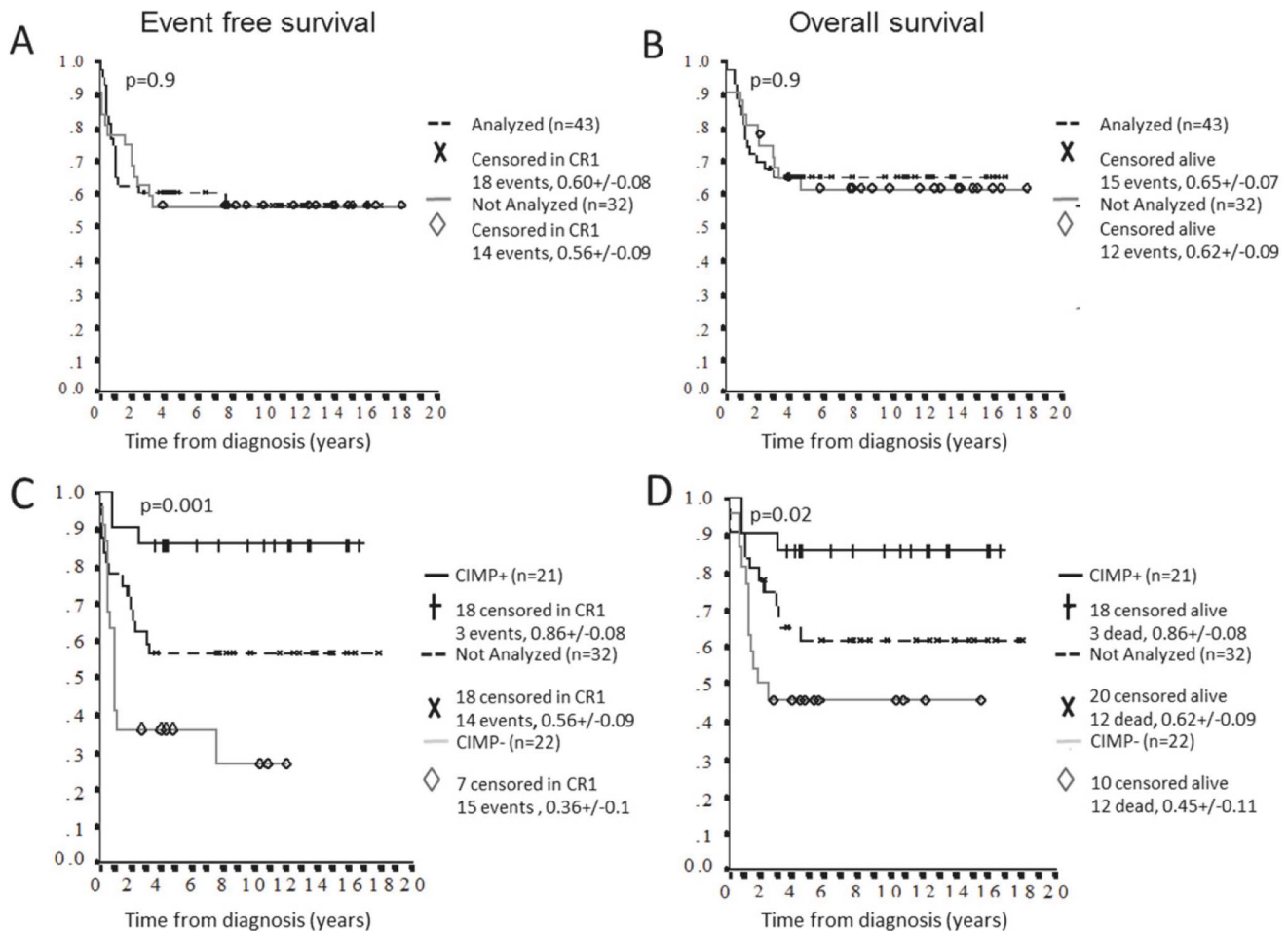


Figure 3. Event free survival (EFS) and overall survival (OS) in methylation subgroups in T-ALL. The 5 year EFS (A) and OS (B) for 43 T-ALL samples analyzed for CpG methylation and 32 T-ALL patients not analyzed but diagnosed during the same time-period showed no differences ($p=0.9$ and $p=0.9$ respectively). The total patient group ($n=75$) had an EFS 5y of 0.59 (SE=0.06) and OS 5y of 0.63 (SE=0.06). CIMP+ ($n=21$) and CIMP- ($n=22$) patients differed significant regarding both 5 year EFS (C) and OS (D) ($p=0.001$ and $p=0.02$, respectively), with cases lacking methylation data forming an intermediate group.
 doi:10.1371/journal.pone.0065373.g003

The T-ALL methylation groups, low ($n=22$), intermediate ($n=6$) and high ($n=15$) (Figure 2A) were examined for clinical outcome. The intermediate and high methylation groups showed no significant difference in outcome regarding EFS and OS ($p=0.25$ for both) and were merged into one group denominated CIMP+ (Figure S2 C–D). Accordingly, cases with low methylation were called CIMP- (Figure 2A).

Interestingly, using this subclassification CIMP+ patients had significantly better EFS (86% vs. 36%, $p=0.001$) and OS (86% vs. 45%, $p=0.02$) compared to CIMP- patients (Figure 3C–D). To confirm that our samples ($n=43$) were representative of the entire cohort of patients ($n=75$), CIMP+ ($n=21$) and CIMP- ($n=22$) cases were compared to the 32 T-ALL patients diagnosed during the same time period and for whom DNA and subsequently CIMP data were missing. Clinical data regarding gender, age, WBC, and outcome for the three groups are given in Table 1. No differences in distribution regarding sex, age or white cell count were observed between the CIMP+ and CIMP- groups (Table 1). The median observation time for patients in continuous CR1 was 127 months (range 32–214 months). The differences in EFS and OS between the CIMP +/- groups were retained in multivariate analysis including WBC, age, gender and CIMP status ($p=0.006$ and

$p=0.044$ respectively) (Table 2). Cases lacking methylation data formed an intermediate group between the CIMP+ and CIMP- groups (Figure 3C–D).

Polycomb Target Genes, Transcription Factors, and Genes in ATP Metabolism were Preferentially Methylated in CIMP+ Samples

To identify potential genes and cellular processes that were connected to the CIMP groups we used the GeneGO metacore software and the DAVID Functional Annotation Tool. The result showed that 1038 genes corresponding to the 1347 differently methylated CpG sites described above were most significantly associated with sequence specific binding, transcription regulator activity, protein binding, and promoter binding (data not shown).

Our initial GeneGO Metacore analysis of DMGs indicated an overrepresentation of genes associated with the polycomb repressive complexes (PRC) 1 and 2. This is in line with recent reports on AML as well as solid tumor types where hypermethylation of PRC target genes has been observed [14,26,27]. To further investigate this we compared previously published lists of PRC target genes identified in human embryonic stem cells (Lee et al.) [25] and

human embryonic fibroblasts (Bracken et al.) [24] with our list of most variable gene promoters in T-ALL (n = 1038) (Figure 4A). A high proportion of genes (62%) in our list were identified as polycomb target genes by Lee et al. and/or Bracken et al. The number of PRC target genes in our list was much higher than expected from a random selection ($p < 0.0001$), indicating a preferential methylation of these genes (Figure 4A). We analyzed if the HumMeth27K methylation array was biased regarding the number of CpG sites measured per gene. The Lee and Bracken polycomb target gene lists were combined and compared with the entire HumMeth27K array for distribution of 1, 2 and ≥ 3 CpG sites per gene. The distribution was comparable, but with a slightly higher number of CpGs in polycomb target genes on the array. However, there was a highly significant overrepresentation of polycomb target genes among the differently methylated CpG sites in T-ALL that cannot be explained by the small array bias (Figure S3 A and S3 B).

Furthermore, the differently methylated genes in T-ALL (1038 genes) and the 260 common PRC target genes in Bracken/Lee/T-ALL (Figure 4A) were evaluated for protein function by enrichment analysis using the Metacore software and compared with genes in the GeneGo database (n = 23868) and in the

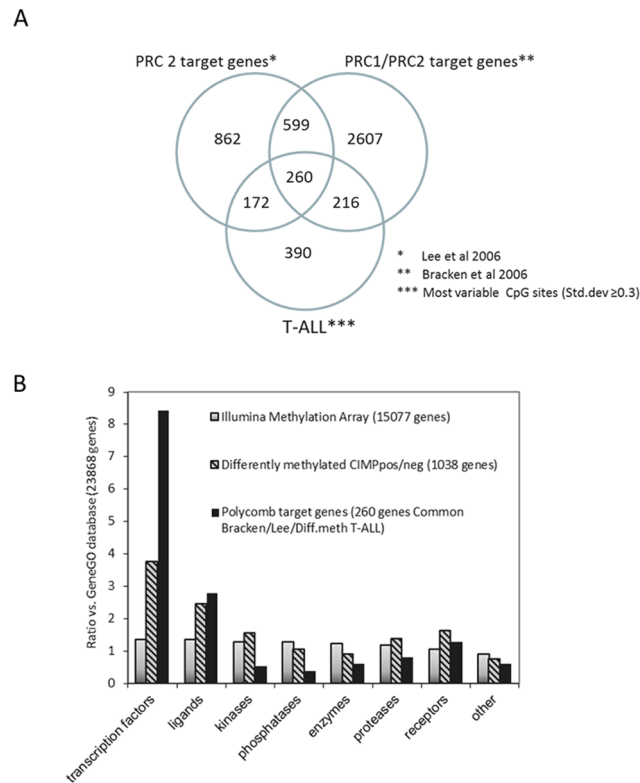


Figure 4. Differently methylated genes in T-ALL are overrepresented by transcription factors and ligands and enriched for polycomb target genes. A) Lists of polycomb target genes identified in human embryonic fibroblasts and embryonic stem cells were compared with the DMGs in T-ALL in a Venn diagram. 260 genes were commonly present in these lists. B) The most variable CpG sites in T-ALL (1347 CpGs/1038 genes) and the 260 common polycomb genes identified in Figure 4A were evaluated for protein function, and compared with the protein function distribution of genes within the GeneGo database (23868 genes) and the Illumina methylation array (15077 genes). Transcription factors and ligands were overrepresented in the DMG in T-ALL compared with the distribution within the database and the entire methylation array. doi:10.1371/journal.pone.0065373.g004

methylation array (n = 15077). The distribution of genes regarding protein function was similar in the array and in the GeneGO database. However, the DMGs in our T-ALL samples were clearly overrepresented by transcription factors and ligands, 3.8 and 2.5 fold more abundant than expected (Figure 4B). Small or no differences were seen in the following subgroups: kinases, phosphatases, enzymes, proteases and receptors (Figure 4B). The overrepresentation of transcription factors and ligands was further strengthened when focusing on the 260 polycomb target genes (8.4 and 2.8 fold increase, respectively) (Figure 4A–B).

Focused analysis of cell processes and metabolic pathways revealed a large number of genes involved in ATP metabolism and cAMP signaling pathways to be differently methylated within the T-ALL group (Figure 5). These genes were more methylated in the CIMP+ group compared to the CIMP- group (data not shown). These genes included adenylate cyclases (*ADCY1*, *ADCY2*, *ADCY4*, *ADCY5* and *ADCY8*) and phosphodiesterases (*PDE10A*, *PDE3A*, *PDE4B*, *PDE8A*, and *PDE1C*), genes that are distributed on different chromosomal locations.

Promoter Methylation in Relation to Gene Expression

In order to examine the correlation between promoter methylation and gene expression we performed gene expression array analysis (Illumina HT-12) on 17 T-ALL samples. A significant (Spearman correlation $Rho = -0.260$, $p < 0.001$) inverse correlation was observed between promoter methylation and gene expression (Figure S4 A).

By combining the list of differently methylated genes (1038 genes) with a list of differently expressed genes (2-fold up/down, 405 genes), 39 genes including *SOCS2*, *DCHS1*, *EPHX2*, *ATP9A*, *MARCKS*, *SLX6*, *TFAP2C* and *ANXA5* were identified as differently methylated and differently expressed (Figure S4 B). The majority of these 39 genes showed a negative correlation between methylation and gene expression, i.e. high promoter methylation was associated with low expression. Since only 47% of the DMGs where expressed above background signal level in the gene expression array, evaluation of a possible influence of promoter methylation on gene expression was not feasible for all genes, i.e. the majority of the cAMP/ATP signaling pathway genes.

We next focused our analysis on selected genes previously associated with T-ALL development, including *TAL1*, *LYL1*, *TLX1* and *TLX3*. None of these genes were identified as DMGs between CIMP subgroups. However, *TAL1* gene expression (HT-12 array) was negatively correlated ($R^2 = 0.611$) with *TAL1* methylation levels (TargetID cg19797376, HumMeth27K array) in 17 T-ALL samples (Figure S5 A). We verified *TAL1* gene expression levels in the array by q-RT-PCR analysis in six samples ($R^2 = 0.972$) (Figure S5 B). Detailed analysis of the Human-Meth450K methylation array (including 36 CpG sites distributed at different locations in the *TAL1* promoter and gene body) showed that methylation of CpG sites 1–1500 nucleotides upstream of transcription start site (TSS) was associated with gene expression, in contrast to CpG sites located within the *TAL1* gene body (Figure S5 C). Methylation level of the *TAL1* CpG site (cg19797376) located close to transcription start site (TSS 200 region) was verified by repeated array analysis and pyrosequencing (Figure S5 D).

An early T cell precursor (ETP) gene expression signature associated with a worse prognosis in T-ALL was identified by Gutierrez et al. in 2010 [28]. We used this signature for analysis of the 17 T-ALL samples studied in the gene expression array. However, the ETP signature did not discriminate CIMP+ from CIMP- individuals (Figure S7).

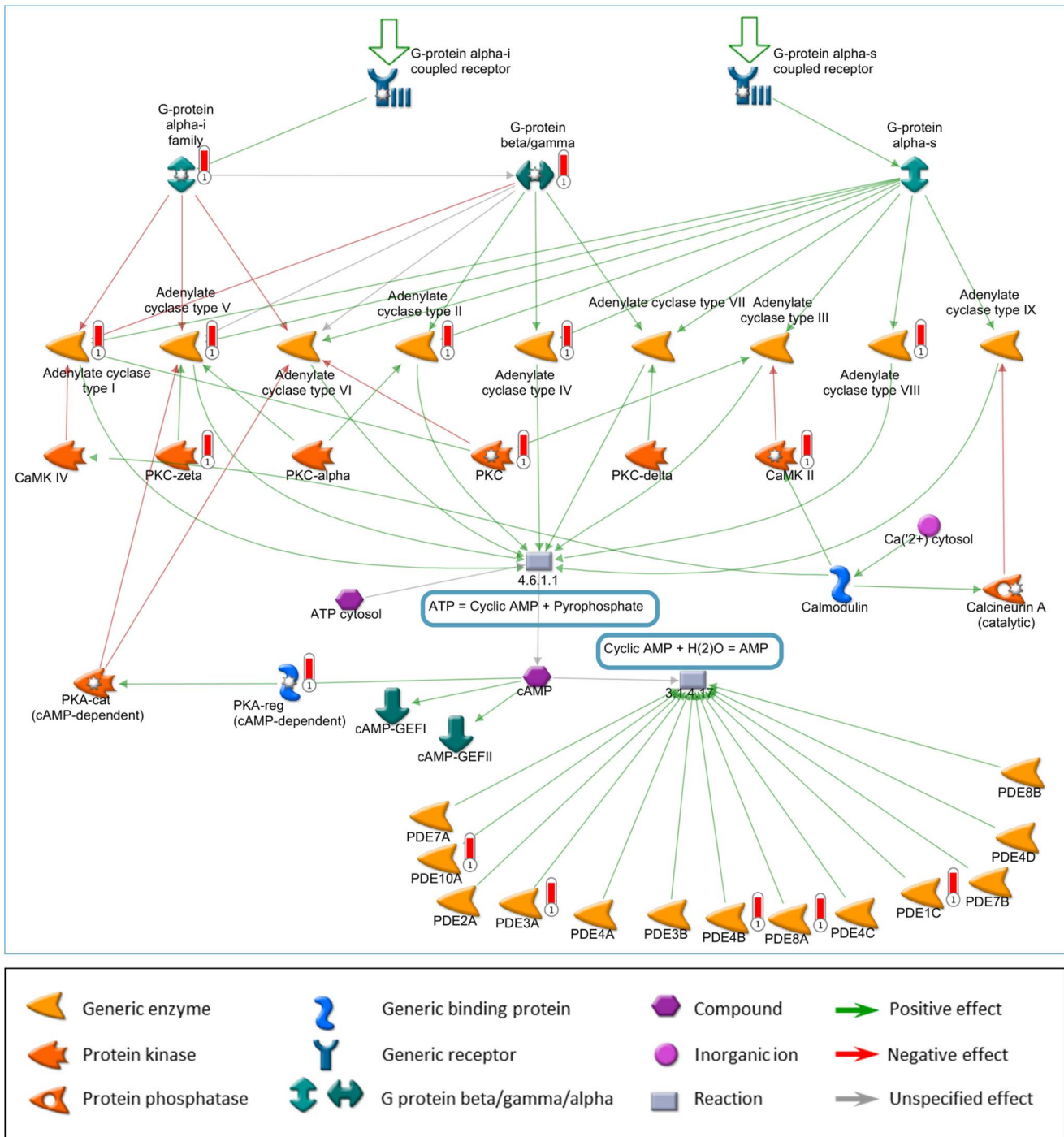


Figure 5. Genes in cAMP signaling and ATP metabolism among the DMGs in T-ALL. A cAMP signaling and ATP metabolism network was created in the MapEditor tool of the Metacore software, and the list of differently methylated genes (DMGs) (1347 CpG/1038 genes, std. dev ≥ 0.3 within T-ALL) was applied on the network. Each red thermometer highlights a DMG within the network. doi:10.1371/journal.pone.0065373.g005

Discussion

In the present study we demonstrated a prognostic significance of global promoter DNA methylation status at diagnosis in T-ALL. We observed a high variability in CpG site promoter methylation between individual T-ALL samples, which could be subdivided as CIMP+ or CIMP- experiencing large differences in clinical outcome. Previous studies indicate that a CIMP phenotype exists

in T-ALL, but are inconclusive regarding a potential connection to prognosis [12,13,29–31]. One explanation for the diverging reports is the lack of a common CIMP definition and that diverse methodologies have been used. We utilized an array approach to study genome-wide promoter methylation [32,33] whereas other laboratories have studied a selection of few promoter CpGs [34]. Our classification was based on CpG sites (n = 1347) with the most variable methylation levels within the T-ALL group and revealed

a significantly better 5 year OS ($p = 0.02$) for the CIMP+ T-ALL group (86%) compared with the CIMP- group (45%). Since relapse is associated with poor outcome EFS was very similar to OS, 86% in CIMP+ compared with 36% in CIMP- group. Interestingly, an association between global promoter methylation and good prognosis has recently been observed in both childhood pre-B ALL and in cytogenetically normal adult AML using the same array platform as in the present study [14,35]. The similar findings in two clinically separate diseases strongly suggest that promoter methylation status can give clinically useful information in separate hematological malignancies. It is notable that the CIMP- T-ALL samples experiencing a worse outcome had a similar methylation profile as normal bone marrow cells and stimulated T-cells. However, in order to compare T-ALL samples with normal samples of the same cellular origin, purified pre-T cells are needed.

DMGs distinguishing CIMP- from CIMP+ samples were in our samples enriched for targets of the polycomb complexes (PRC 1 and 2). These complexes have chromatin-modifying capacity, influence CpG methylation and are involved in regulation of developmental and differentiation processes [24,25]. PRC members are commonly deregulated in cancer, including leukemia and lymphoma [36,37]. Recently whole-genome sequencing of early T-cell precursor ALL identified alterations in PRC2 genes, including SUZ12, EZH2 and EED [38]. In another study, mice with bi-allelic deletions of EZH2 showed high frequency of T-ALL and low levels of histone H3K27 tri-methylation [39]. We analyzed DMGs in T-ALL according to their protein function and observed an overrepresentation of transcription factors and ligands compared with an expected random event. The overrepresentation was even stronger when the analysis was restricted to DMGs overlapping with polycomb target genes identified in embryonic stem cells and embryonic fibroblasts [24,25]. Lee et al showed a preferential binding of SUZ12, the DNA binding component of PRC2, to transcription factors associated with developmental processes [25]. Accordingly, the molecular function of the DMGs in T-ALL was strongly associated with sequence specific binding, transcription regulator activity, protein binding and promoter binding.

One interesting question is if the increased methylation in CIMP+ cases is random or not [40]. Our analysis of DMGs indicated a nonrandom process with preferential methylation of transcription factors, ligands and polycomb targets genes for at least specific sets of genes. Furthermore, methylation seemed to be favored at CpG sites located within CpG islands. In accordance, previous studies have shown that CpG sites within CpG islands are more prone to undergo tumor specific methylation compared to CpG sites outside islands [41].

One important issue is the biological background to the association between DNA methylation and clinical outcome, which we only can speculate upon. It is obvious that the CIMP- subgroup with a poor outcome is more resistant to the treatment received compared to CIMP+ ALL. On the other hand, the better outcome for CIMP+ cases might suggest that specific methylation events are important for a therapy sensitive phenotype. Combinations of cytotoxic drugs are used in pediatric T-ALL therapy targeting a wide range of cellular processes in order to induce DNA damage and cell death [16]. The bioinformatics analysis revealed that many DMGs in CIMP+ T-ALL were genes involved in ATP metabolism and cAMP signaling processes, which have been considered as potential targets for cancer therapy [42–44]. However, the relevance of these methylation changes for therapy response in T-ALL needs further studies.

In concordance with other studies we observed a significant negative and weak correlation between methylation and gene expression at a global level [14]. The rather weak global correlation has also been previously observed [14,45,46]. Jung *et al.* combined methylation data from the Illumina Golden Gate methylation array (1505 CpG sites/807 genes) with Affymetrix gene expression data in 193 multiple myeloma samples and found only 2.1% of the CpG sites to be correlated with gene expression [45]. The correlation between methylation and gene expression might also depend on where the CpG sites are located, e.g. within/outside CpG islands and the CpG site gene location (e.g. promoter, gene body) [47]. The number of CpGs per promoter in the 27K array did not permit a detailed analysis, but future analysis by high resolution methylation arrays or next generation sequencing might reveal a deeper understanding of the relation between methylation alterations and gene expression. However, we identified several genes with a negative correlation between gene expression and promoter CpG methylation, including EPHX2, DSCHS1, MARSCS, ANXA5, and the T-ALL associated TAL1 gene. Whether or not these genes contributed to the poor prognosis in the CIMP- group remains to be evaluated. It has previously been shown that T-ALL cases with an ETP phenotype, representing up to 15% of all T-ALL cases, have a worse prognosis [28,38,48]. We analyzed the gene expression profile associated with the ETP signature and found no association to CIMP status.

In conclusion, our study identified two CIMP groups of T-ALL with significantly different prognosis regarding risk of relapse and overall survival. Few prognostic T-ALL markers are available today and the potential for CIMP status as a predictor of clinical outcome and as a guide for therapy seems promising but needs to be further investigated in larger patient cohorts.

Supporting Information

Figure S1 Hierarchical clustering of 10 T-ALL samples analyzed with the HumanMeth27K and 450 K arrays.

Ten T-ALL samples were analyzed on both the HumanMeth27K and HumanMeth450K Illumina methylation arrays. Unsupervised Euclidean hierarchical clustering of the 25978 overlapping CpG sites covered by both arrays in the A) HumanMeth27K and B) HumanMeth450K arrays. Methylation groups identified in Figure 2A is shown in the figure. (TIF)

Figure S2 Event free survival (EFS) and overall survival (OS) in T-ALL.

EFS (A) and OS (B) for the 2 different treatment protocols used during years 1992–2008 at which the diagnosis T-ALL samples were collected showing no significant difference between these protocols ($p = 0.99$ and $p = 0.3$, respectively). EFS (C) and OS (D) for the methylation subgroups (low $n = 22$, intermediate $n = 6$, and high $n = 15$) identified by hierarchical clustering of the most variable CpG sites in T-ALL (1347 CpGs, Figure 2A). The intermediate and high methylation subgroups showed no difference regarding these parameters (EFS $p = 0.25$, OS $p = 0.25$) while the low methylation subgroup differed significantly from these groups (EFS low/high $p = 0.005$, low/intermediate $p = 0.01$). The intermediate and high methylation groups were hereafter collectively grouped as CIMP+ and the low group CIMP-. (TIF)

Figure S3 Distribution of CpG sites on the HumanMeth27K array. A) The HumanMeth27K array was examined for possible bias regarding the number of CpG sites in polycomb

target genes compared to the entire array. The Lee et al. (2006) and Bracken et al. (2006) polycomb target gene lists were combined and compared with the entire HumMeth27K array for distribution of 1, 2 and ≥ 3 CpG sites per gene. B) The distribution of CpG sites located in polycomb and non-polycomb target gene promoters on the HumMeth27K array was compared with the distribution of the differently methylated CpG sites within T-ALL. Pearson's Chi squared test was used to test the distribution in both analyses.

(TIF)

Figure S4 Correlation of gene expression with CpG promoter methylation. A) Mean CpG methylation levels (beta-value, x-axis) is plotted against mean gene expression levels (log average signal, y-axis) from 17 T-ALL samples. Gene expression data below background level was excluded from the analysis and the remaining 18500 CpG sites is shown. Spearman correlation $Rho = -0.260$, $P < 0.001$. B) Examples of 8 selected genes identified as differently methylated (T-ALL Std. Dev. ≥ 0.3 , 1347 sites) and differently expressed (2-fold up/down CIMP+/-, 405 genes). Gene expression (log Average signal) is shown on y-axis, and CpG methylation (beta-value) on x-axis for 17 T-ALL patients, and R^2 correlation coefficient shown in figure.

(TIF)

Figure S5 TAL1 promoter methylation and gene expression. A) TAL1 gene expression (Average signal, HT12 array) and CpG methylation (Average beta, HumanMeth27K array) correlated at $R^2 = 0.611$. B) TAL1 gene expression was verified by TaqMan gene expression q-RT-PCR analysis in 6 samples. Array data and q-RT-PCR data correlated at $R^2 = 0.972$. C) Detailed analysis of mean CpG site methylation by the HumanMeth450K array in promoter regions; Gene body, TSS200 (0–200 nt upstream of transcription start site) and TSS1500 (200–1500 nt upstream of TSS) and in correlation to the *TAL1/TBP* gene expression ratio. D) Verification of the TAL1 CpG site (cg19797376) by pyrosequencing. Methylation status (% methylated) revealed by pyrosequencing was compared with Human-

Meth27K and HumanMeth450K methylation status on the same CpG position.

(TIF)

Figure S6 Verification of array data by pyrosequencing. The methylation status of selected gene promoters, HOXD8 (cg15520279), and KLF4 (cg07309102), was determined by pyrosequencing and compared with the HumMeth27K array data.

(TIF)

Figure S7 ETP gene signature analysis. The early T cell precursor (ETP) gene expression signature identified by Gutierrez et al. in 2010 [28] was used for analysis of 17 T-ALL samples analyzed by the HT-12 gene expression array. The heat map shows log. Average signal of the ETP associated genes for each sample. Euclidean hierarchical cluster analysis did not discriminate CIMP+ from CIMP- samples.

(TIF)

Table S1 Most variable (Std.dev. ≥ 0.3) CpGs (n = 1347) in T-ALL (n = 43).

(XLSX)

Acknowledgments

Methylation profiling with the HumanMeth450K array was performed by the SNP&SEQ Technology Platform in Uppsala, Sweden. Pyrosequencing was performed by the Genome Centre Queen Mary, University of London, England. The authors would like to thank Inger Cullman and Emma Andersson Evelönn; Medical Biosciences, Umeå University, for technical assistance with the HumanMeth27K arrays, and Patrik Ryden and Mattias Landfors, Mathematics and mathematical statistics, Umeå University, for assistance with statistical analysis.

Author Contributions

Conceived and designed the experiments: SD MBo GR EF. Performed the experiments: SD. Analyzed the data: SD MBo GR EF. Contributed reagents/materials/analysis tools: EF KK JH MBe LP JP. Wrote the paper: SD MBo GR EF.

References

- Aifantis I, Raetz E, Buonamici S (2008) Molecular pathogenesis of T-cell leukaemia and lymphoma. *Nat Rev Immunol* 8: 380–390.
- Karman K, Forestier E, Heyman M, Andersen MK, Autio K, et al. (2009) Clinical and cytogenetic features of a population-based consecutive series of 285 pediatric T-cell acute lymphoblastic leukemias: rare T-cell receptor gene rearrangements are associated with poor outcome. *Genes Chromosomes Cancer* 48: 795–805.
- Weng AP, Ferrando AA, Lee W, Morris JPt, Silverman LB, et al. (2004) Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science* 306: 269–271.
- Ma J, Wu M (2012) The indicative effect of Notch1 expression for the prognosis of T-cell acute lymphocytic leukemia: a systematic review. *Mol Biol Rep* 39: 6095–6100.
- Jones PA, Baylin SB (2007) The epigenomics of cancer. *Cell* 128: 683–692.
- Esteller M (2008) Epigenetics in cancer. *N Engl J Med* 358: 1148–1159.
- Ang PW, Loh M, Liem N, Lim PL, Grieu F, et al. (2010) Comprehensive profiling of DNA methylation in colorectal cancer reveals subgroups with distinct clinicopathological and molecular features. *BMC Cancer* 10: 227.
- Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, et al. (2010) Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell* 17: 510–522.
- Van der Auwera I, Yu W, Suo L, Van Neste L, van Dam P, et al. (2011) Array-based DNA methylation profiling for breast cancer subtype discrimination. *PLoS One* 5: e12616.
- Shen L, Kantarjian H, Guo Y, Lin E, Shan J, et al. (2010) DNA methylation predicts survival and response to therapy in patients with myelodysplastic syndromes. *J Clin Oncol* 28: 605–613.
- Quintas-Cardama A, Santos FP, Garcia-Manero G (2010) Therapy with azanucleosides for myelodysplastic syndromes. *Nat Rev Clin Oncol* 7: 433–444.
- Roman-Gomez J, Jimenez-Velasco A, Agirre X, Castillejo JA, Navarro G, et al. (2006) Promoter hypermethylation and global hypomethylation are independent epigenetic events in lymphoid leukemogenesis with opposing effects on clinical outcome. *Leukemia* 20: 1445–1448.
- Milani L, Lundmark A, Kiialainen A, Nordlund J, Flaegstad T, et al. (2010) DNA methylation for subtype classification and prediction of treatment outcome in patients with childhood acute lymphoblastic leukemia. *Blood* 115: 1214–1225.
- Deneberg S, Guardiola P, Lennartsson A, Qu Y, Gaidzik V, et al. (2011) Prognostic DNA methylation patterns in cytogenetically normal acute myeloid leukemia are predefined by stem cell chromatin marks. *Blood* 118: 5573–5582.
- Gustafsson G, Schmiegelow K, Forestier E, Clausen N, Glomstein A, et al. (2000) Improving outcome through two decades in childhood ALL in the Nordic countries: the impact of high-dose methotrexate in the reduction of CNS irradiation. *Nordic Society of Pediatric Haematology and Oncology (NOPHO). Leukemia* 14: 2267–2275.
- Schmiegelow K, Forestier E, Hellebostad M, Heyman M, Kristinsson J, et al. (2010) Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. *Leukemia* 24: 345–354.
- Degerman S, Siwicki JK, Osterman P, Lafferty-Whyte K, Nicol Keith W, et al. (2010) Telomerase upregulation is a postcrisis event during senescence bypass and immortalization of two Nijmegen breakage syndrome T cell cultures. *Aging Cell*.
- Siwicki JK, Hedberg Y, Nowak R, Loden M, Zhao J, et al. (2000) Long-term cultured IL-2-dependent T cell lines demonstrate p16(INK4a) overexpression, normal pRb/p53, and upregulation of cyclins E or D2. *Experimental Gerontology* 35: 375–388.
- Weisenberger DJ, Trinh BN, Campan M, Sharma S, Long TI, et al. (2008) DNA methylation analysis by digital bisulfite genomic sequencing and digital MethyLight. *Nucleic Acids Res* 36: 4689–4698.
- Weisenberger DJ, Campan M, Long TI, Kim M, Woods C, et al. (2005) Analysis of repetitive element DNA methylation by MethyLight. *Nucleic Acids Res* 33: 6823–6836.
- Bibikova M, Le J, Barnes B et al (2009) Genome-wide DNA methylation profiling using Infinium assay. *Epigenomics* 1: 177–200.

22. Takai D, Jones PA (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci U S A* 99: 3740–3745.
23. Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44–57.
24. Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K (2006) Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev* 20: 1123–1136.
25. Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, et al. (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 125: 301–313.
26. Bracken AP, Helin K (2009) Polycomb group proteins: navigators of lineage pathways led astray in cancer. *Nat Rev Cancer* 9: 773–784.
27. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, et al. (2007) Epigenetic stem cell signature in cancer. *Nat Genet* 39: 157–158.
28. Gutierrez A, Dahlberg SE, Neuberg DS, Zhang J, Grebliunaite R, et al. (2010) Absence of biallelic TCRgamma deletion predicts early treatment failure in pediatric T-cell acute lymphoblastic leukemia. *J Clin Oncol* 28: 3816–3823.
29. Gutierrez MI, Siraj AK, Bhargava M, Ozbek U, Banavali S, et al. (2003) Concurrent methylation of multiple genes in childhood ALL: Correlation with phenotype and molecular subgroup. *Leukemia* 17: 1845–1850.
30. Kraszczyńska MD, Dawidowska M, Larmonic NS, Kosmalka M, Sedek L, et al. (2011) DNA methylation pattern is altered in childhood T-cell acute lymphoblastic leukemia patients as compared with normal thymic subsets: insights into CpG island methylator phenotype in T-ALL. *Leukemia*.
31. Milani L, Lundmark A, Nordlund J, Kiiäläinen A, Flaegstad T, et al. (2009) Allele-specific gene expression patterns in primary leukemic cells reveal regulation of gene expression by CpG site methylation. *Genome Res* 19: 1–11.
32. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, et al. (2011) High density DNA methylation array with single CpG site resolution. *Genomics*.
33. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, et al. (2006) High-throughput DNA methylation profiling using universal bead arrays. *Genome Res* 16: 383–393.
34. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB (1996) Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci U S A* 93: 9821–9826.
35. Sandoval J, Heyn H, Mendez-Gonzalez J, Gomez A, Moran S, et al. (2013) Genome-wide DNA methylation profiling predicts relapse in childhood B-cell acute lymphoblastic leukaemia. *Br J Haematol* 160: 406–409.
36. Ntzachristos P, Tsigros A, Van Vlierberghe P, Nedjic J, Trimarchi T, et al. (2012) Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. *Nat Med* 18: 298–301.
37. Martin-Perez D, Piris MA, Sanchez-Beato M (2010) Polycomb proteins in hematologic malignancies. *Blood* 116: 5465–5475.
38. Zhang J, Ding L, Holmfeldt L, Wu G, Heatley SL, et al. (2012) The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* 481: 157–163.
39. Simon C, Chagraoui J, Kros J, Gendron P, Wilhelm B, et al. (2012) A key role for EZH2 and associated genes in mouse and human adult T-cell acute leukemia. *Genes Dev*.
40. Costello JF, Fruhwald MC, Smiraglia DJ, Rush LJ, Robertson GP, et al. (2000) Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat Genet* 24: 132–138.
41. Martin-Subero JI, Ammerpohl O, Bibikova M, Wickham-Garcia E, Agirre X, et al. (2009) A comprehensive microarray-based DNA methylation study of 367 hematological neoplasms. *PLoS One* 4: e6986.
42. Abecassis I, Maes J, Carrier JL, Hillion J, Goodhardt M, et al. (2008) Re-expression of DNA methylation-silenced CD44 gene in a resistant NB4 cell line: rescue of CD44-dependent cell death by cAMP. *Leukemia* 22: 511–520.
43. Lucchi S, Calebiro D, de Filippis T, Grassi ES, Borghi MO, et al. (2011) 8-Chloro-cyclic AMP and protein kinase A I-selective cyclic AMP analogs inhibit cancer cell growth through different mechanisms. *PLoS One* 6: e20785.
44. Rajendra Prasad VV, Peters GJ, Lemos C, Kathmann I, Mayor YC (2011) Cytotoxicity studies of some novel fluoro acridone derivatives against sensitive and resistant cancer cell lines and their mechanistic studies. *Eur J Pharm Sci* 43: 217–224.
45. Jung S, Kim S, Gale M, Cherni I, Fonseca R, et al. (2012) DNA methylation in multiple myeloma is weakly associated with gene transcription. *PLoS One* 7: e52626.
46. Sanders YY, Ambalavanan N, Halloran B, Zhang X, Liu H, et al. (2012) Altered DNA methylation profile in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 186: 525–535.
47. van Eijk KR, de Jong S, Boks MP, Langeveld T, Colas F, et al. (2012) Genetic analysis of DNA methylation and gene expression levels in whole blood of healthy human subjects. *BMC Genomics* 13: 636.
48. Coustan-Smith E, Mullighan CG, Onciu M, Behm FG, Raimondi SC, et al. (2009) Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. *Lancet Oncol* 10: 147–156.

DNA Methylation Adds Prognostic Value to Minimal Residual Disease Status in Pediatric T-Cell Acute Lymphoblastic Leukemia

Magnus Borssén, MD,^{1#} Zahra Haider, MSc,^{1#} Mattias Landfors, PhD,¹ Ulrika Norén-Nyström, MD, PhD,² Kjeld Schmiegelow, MD, PhD,³ Ann E. Åsberg, MD, PhD,⁴ Jukka Kanerva, MD, PhD,⁵ Hans O. Madsen, PhD,⁶ Hanne Marquart, MD, PhD,⁶ Mats Heyman, MD, PhD,⁷ Magnus Hultdin, MD, PhD,¹ Göran Roos, MD, PhD,¹ Erik Forestier, MD, PhD,¹ and Sofie Degerman, PhD^{1,4*}

Background. Despite increased knowledge about genetic aberrations in pediatric T-cell acute lymphoblastic leukemia (T-ALL), no clinically feasible treatment-stratifying marker exists at diagnosis. Instead patients are enrolled in intensive induction therapies with substantial side effects. In modern protocols, therapy response is monitored by minimal residual disease (MRD) analysis and used for postinduction risk group stratification. DNA methylation profiling is a candidate for subtype discrimination at diagnosis and we investigated its role as a prognostic marker in pediatric T-ALL. **Procedure.** Sixty-five diagnostic T-ALL samples from Nordic pediatric patients treated according to the Nordic Society of Pediatric Hematology and Oncology ALL 2008 (NOPHO ALL 2008) protocol were analyzed by HumMeth450K genome wide DNA methylation arrays. Methylation status was analyzed in relation to clinical data and early T-cell precursor (ETP) phenotype. **Results.** Two distinct CpG island

methylator phenotype (CIMP) groups were identified. Patients with a CIMP-negative profile had an inferior response to treatment compared to CIMP-positive patients (3-year cumulative incidence of relapse (CIR_{3y}) rate: 29% vs. 6%, $P = 0.01$). Most importantly, CIMP classification at diagnosis allowed subgrouping of high-risk T-ALL patients (MRD $\geq 0.1\%$ at day 29) into two groups with significant differences in outcome (CIR_{3y} rates: CIMP negative 50% vs. CIMP positive 12%; $P = 0.02$). These groups did not differ regarding ETP phenotype, but the CIMP-negative group was younger ($P = 0.02$) and had higher white blood cell count at diagnosis ($P = 0.004$) compared with the CIMP-positive group. **Conclusions.** CIMP classification at diagnosis in combination with MRD during induction therapy is a strong candidate for further risk classification and could confer important information in treatment decision making. Pediatr Blood Cancer 0000;00:000–000. © 2016 Wiley Periodicals, Inc.

Key words: childhood leukemia; DNA methylation; MRD; prognosis; T-ALL; WBC

Additional supporting information can be found in the supporting information tab for this article.

Abbreviations: CIMP, CpG island methylator phenotype; CIR, cumulative incidence of relapse; CIR_{3y}, 3-year cumulative incidence of relapse; CpG, cytosine-phosphate-guanine dinucleotide; EGIL, European group for the immunological classification of leukemias; ETP, early T-cell precursor; HRM, high-resolution melting curve; MRD, minimal residual disease; OS, overall survival; pOS_{3yr}, 3-year overall survival rate; WBC, white blood cell

¹Department of Medical Biosciences, Umeå University, Umeå, Sweden; ²Department of Clinical Sciences and Pediatrics, Umeå University, Umeå, Sweden; ³Department of Paediatrics and Adolescent Medicine, University Hospital Rigshospitalet, Copenhagen, Denmark; ⁴Department of Paediatrics, University Hospital of Trondheim, Norway; ⁵Hospital for Children and Adolescents, University of Helsinki, Helsinki, Finland; ⁶Department of Clinical Immunology, University Hospital Rigshospitalet, Copenhagen, Denmark; ⁷Department of Woman and Child health, Karolinska Institute, Stockholm, Sweden

[#]Magnus Borssén and Zahra Haider contributed equally to this study.

Grant sponsor: Swedish Cancer Society; Grant sponsor: Swedish Research Council; Grant sponsor: Swedish Childhood Cancer Foundation; Grant sponsor: Medical Faculty of Umeå University; Grant sponsor: Danish Childhood Cancer Foundation; Grant sponsor: Lion's Cancer Research Foundation, Umeå; Grant sponsor: Umeå Paediatric Clinic Research Foundation; Grant sponsor: Magnus Bergvalls stiftelse; Grant sponsor: Uppsala-Umeå Comprehensive Cancer Consortium.

Conflict of interest: Nothing to declare.

*Correspondence to: Sofie Degerman, Department of Medical Biosciences, Umeå University, Blg 6M, 2nd floor, SE-90185 Umeå, Sweden.

© 2016 Wiley Periodicals, Inc.

DOI 10.1002/pbc.25958

Published online in Wiley Online Library (wileyonlinelibrary.com).

INTRODUCTION

T-cell acute lymphoblastic leukemia (T-ALL) accounts for 10–15% of childhood ALL, but is more common among adolescents and adults.[1] The malignant transformation involves several genomic changes altering the normal control of T-cell development and proliferation.[2] Although the pathogenesis of T-ALL has been extensively studied, few clinically useful prognostic markers exist beyond minimal residual disease (MRD) status during the first months of therapy.[3]

During the last decade, “epigenetics” (that is, functionally relevant changes in the genome that might influence gene expression without affecting the nucleotide sequence), has emerged as an important player in tumor development. One general finding during malignant transformation is a decrease in global DNA methylation, contributing to genomic instability, and an increase in promoter associated CpG island methylation associated with downregulation of tumor suppressor genes.[4] We and others have demonstrated both subtype classification and prognostic relevance of aberrant DNA methylation patterns in various hematological disorders.[5–9] However, the definition of methylation classification into CpG island methylator phenotype (CIMP) subgroups is poorly defined and ranges from classification of single predefined genes to unsupervised genome-wide methylation profiling in different studies.[10] A few other papers have been published on pediatric T-ALL, methylation, and prognosis, but these studies have classified CIMP status based on single selected genes and are not comparable with our genome-wide approach.[11,12]

E-mail: sofie.degerman@umu.se

Received 3 December 2015; Accepted 3 February 2016

Based on our previous finding of a strong prognostic significance of promoter-associated DNA methylation in Swedish T-ALL patients treated according to the Nordic Society of Pediatric Hematology and Oncology (NOPHO) ALL 1992/2000 protocols,[6] we here explore its prognostic relevance in a new independent Nordic cohort of patients treated with the currently used NOPHO ALL 2008 protocol. In contrast to the ALL 1992/2000 protocols, the ALL 2008 protocol uses postinduction MRD levels for risk group stratification.[1]

METHODS

T-Cell Acute Leukemia and Control Samples

Between July 2008 and March 2013, 113 children (age <18 years) were diagnosed in the Nordic countries with T-ALL and treated according to the common NOPHO ALL 2008 protocol.[1] Sixty-five diagnostic bone marrow/peripheral blood samples were available in the NOPHO leukemia biobank in Uppsala, Sweden, and were analyzed for methylation status. The T-ALL diagnosis was in each case based on morphology and flow cytometric immunophenotyping. Cytogenetic aberrations were explored by G-band karyotyping and targeted FISH analysis.[1] Clinical data including white blood cell (WBC) count, immunophenotype, cytogenetic aberrations at diagnosis, and MRD status at treatment day 29 (end of induction) were evaluated in relation to methylation status at diagnosis. MRD was monitored by PCR and/or flow cytometry. PCR analysis of clonal gene rearrangements was recommended for MRD quantification in T-cell ALL and such MRD data [13] were used when available (n = 41). However, if no PCR-based MRD was available, flow cytometric quantification of MRD [14] was used (n = 20). Four cases lacked both PCR and flow MRD data and were excluded from the survival analyses that included MRD.

T-cell maturation stage was evaluated in the diagnostic samples. The European Group for the Immunological Classification of Leukemias (EGIL) criteria were used as defined by Bene et al.[3] The immature subgroup of T-ALL described by Coustan-Smith et al. [15] as “early T-cell precursor” (ETP) is characterized by hematopoietic stem cell (HSC) and myeloid progenitor markers.[3,15–17]

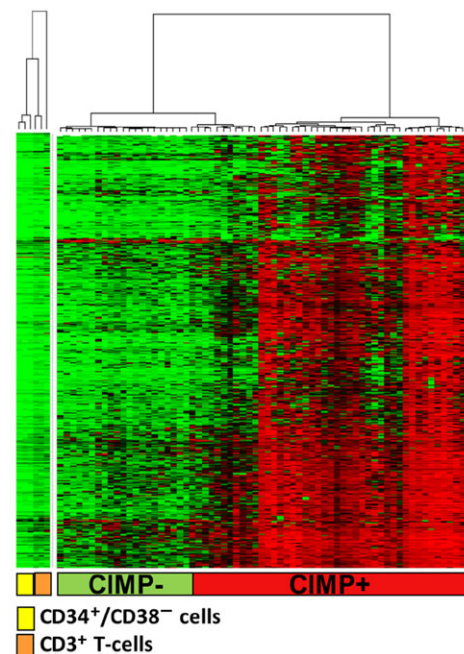
Publicly available methylation data (NCBI GEO database, GSE49618) of sorted T-cells (CD3⁺) and HSCs (CD34⁺, CD38⁻) from healthy donors were used as reference samples in the methylation heat map (Fig. 1A) to illustrate methylation profiles of normal immature and mature hematopoietic cells.

The study was approved by the regional and/or national ethics committees, and the patients and/or their guardians provided informed consent in accordance with the Declaration of Helsinki.

Genome Wide CpG Site Methylation Profiling

DNA was extracted by the AllPrep DNA/RNA kit (Qiagen, Hilden, Germany) and diagnostic T-ALL samples were analyzed by the HumMeth450K methylation array (Illumina, San Diego, CA) covering 485,577 CpG sites. Bisulfite conversion was performed according to manufacturer’s manual (Zymo Research, Irvine, CA) and array analysis including preprocessing and normalization was performed as previously described.[18] CpG probes that align to multiple loci in the genome or were

A



B

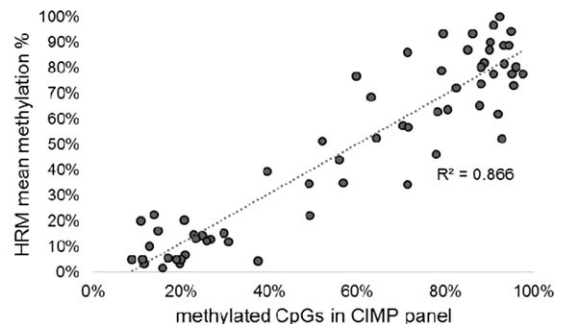


Fig. 1. CIMP classification of T-ALL samples. (A) Heat map showing 1,293 CpG sites in the CIMP panel for the 65 diagnostic T-ALL samples and sorted normal CD34⁺/CD38⁻ and CD3⁺ cells. The samples were CIMP classified based on the percentage of methylated CpGs in the CIMP panel ($\leq 40\%$ CIMP negative and $>40\%$ CIMP positive; Supplementary Table SII). (B) Verification of CIMP classification by a six gene HRM panel. The mean methylation percentage of the HRM-panel versus percentage of methylated CpGs in 1,293 array-CIMP panel showed a high correlation ($R^2 = 0.87$).

located less than 3 bp from a known single nucleotide polymorphism were excluded.[19] The fluorescence intensities were extracted using the Methylation Module (1.9.0) in the Genome Studio software (V2011.1). The methylation level (β value) of each CpG site ranged from 0 (no methylation) to 1 (complete methylation). The quality of each individual array was evaluated with built-in controls. Two replicate samples were included to assess interassay reproducibility ($R^2 = 0.97$ – 0.99). The methylation array data were deposited in the NCBI GEO database, GSE69954. The normalized β values were used as measures of

TABLE I. Clinical Characteristics and Outcome in 113 T-ALL Cases

	CIMP Analyzed N = 65	Not analyzed N = 48	P-value	Total N = 113
Gender male/female	43/22	35/13	ns	78/35
Median age (range, years)	7 (1–17)	7 (2–17)	ns	7 (1–17)
Median WBC x 10 ⁹ /l (range)	150 (1.6–983)	67.1 (0.7–938)	0.03	98.3 (0.7–983)
MRD day 29 <0.1%/≥0.1%/NA	28/33/4	29/18/1	ns	57/51/5
pCIR _{3y} (SE)	0.15 ± 0.05	0.13 ± 0.06	ns	0.15 ± 0.04
pOS _{3y} (SE)	0.79 ± 0.05	0.79 ± 0.07	ns	0.79 ± 0.04

NA, not analyzed; ns, not significant; pCIR_{3y}, 3-year cumulative incidence of relapse; pOS_{3y}, 3-year overall survival rate; SE, standard error.

methylation levels and downstream analysis of the data was performed using R (v2.15.0).

All samples were CpG island methylator phenotype (CIMP) classified according to the previously described CIMP panel [6] using 1,293 CpG sites present in the HumMeth450K array. In short, in order to identify T-ALL methylation subgroups, the CIMP panel was defined as the most variable CpG sites in the 27K array (Illumina) within diagnostic T-ALL samples. The CpG sites within the CIMP panel were characterized by being located within CpG islands and associated with polycomb-target genes.[6]

In order to standardize CIMP classification independent of clustering, a cut-off level for “percentage of methylated CpGs within CIMP panel” was defined for the CIMP subgroups. Diagnostic T-ALL samples with >40% methylated CpG sites (each CpG site was considered methylated if the beta value was >0.4) in the CIMP panel (1293 CpGs) were classified as CIMP positive, whereas samples with ≤40% methylated sites were denoted CIMP negative. The cut-off was set to reflect the previously identified clusters to be the most discriminating with respect to prognosis.[6]

Verification of Methylation Array Data by High-Resolution Melting Curve Analysis

High-resolution melting (HRM) assays were designed for a selection of genes in 1,293 CpG site CIMP panel, representing CpG sites with distinct differences in methylation levels between CIMP subgroups in the array. The genes included *KLF4*, *TFAP2C*, *IGFBP3*, *WNT3A*, *GATA4*, and *EYA4*. Each 25 μl HRM reaction mix contained 1X Epiect HRM PCR Mastermix (Qiagen), 0.75 μM of forward and reverse primers (Supplementary Table SI), and 10 ng DNA template. The analysis was run in a RotorGene instrument (Qiagen) as follows: 95°C for 5 min, followed by 35 cycles of 95°C for 10 sec, 55°C for 30 sec, and 72°C for 20 sec. HRM was conducted by melting from 60 to 90°C, rising by 0.1° each step. A standard curve was prepared by mixing 100% methylated DNA (*M.SssI* treated) in different ratios with DNA from mitogen (wheat germ agglutinin) stimulated primary lymphoblast T-cell cultures P7/R2 (theoretically 0% methylated mononuclear cells). The methylation level of each gene region covered in the HRM assay was estimated in relation to the standard curve, and a mean methylation level (%) of the six-gene HRM panel was calculated. DNA was available from 63 of 65 array-analyzed T-ALL samples. Data were analyzed using Rotor-Gene® software v1.7 (Qiagen).

Pediatr Blood Cancer DOI 10.1002/pbc

Statistical Analysis

The Statistical Package for the Social Sciences (SPSS Inc., Chicago, IL) software for Macintosh 22 was used for the statistical analyses. The chi-square/Fisher's exact test was used to compare differences between subgroups among categorical variables and the Mann–Whitney U test was used for continuous variables. Estimates of 3-year cumulative incidence of relapse (pCIR_{3yr}) and overall survival (pOS_{3yr}) rates were calculated using the Kaplan–Meier method and the subgroups listed in tables were compared using the log rank test. The significance level used in all tests was 0.05. Time in first remission (CR1) was defined as time (month) from diagnosis until first event. In CIR analysis, relapse was the endpoint. In the overall survival (OS) analysis, death from any cause was the endpoint. Moreover, the cumulative incidence of death in remission (CIDCR1) was compared between CIMP subgroups. The NOPHO leukemia registry is updated annually, and follow-up data were extracted from the registry as of May 2015.

RESULTS

Demographic Data

One hundred thirteen children (age ≤18 years) were diagnosed with T-ALL in the Nordic countries between July 2008 and March 2013, and treated according to the Nordic study protocol NOPHO ALL 2008. Clinical response to induction therapy with dexamethasone, vincristine, doxorubicin, and intrathecal methotrexate [1] was evaluated by MRD at day 29. The response to induction therapy determined whether the patients were assigned to antimetabolite-based intermediate risk therapy (MRD < 0.1%) or intensive myelosuppressive high risk (MRD ≥0.1%) block therapy, respectively.[1,20]

A total of 65 diagnostic T-ALL samples in the NOPHO biobank were available for methylation analysis. Apart from higher WBC counts at diagnosis for those included ($P = 0.03$), there were no significant differences between the analyzed ($n = 65$) and not analyzed ($n = 48$) samples; that is, no statistically significant differences were found regarding age, gender, or MRD day 29 status (fraction <0.1%/≥0.1%; Table I). CIR and OS analysis confirmed that the 65 patients with available samples in the NOPHO biobank were representative for all pediatric T-ALL cases diagnosed in the Nordic countries during the study period with no difference in pCIR_{3yr} between patients analyzed for methylation ($n = 65$) or not analyzed ($n = 48$, 15% vs. 13%, $P = 0.82$) or pOS_{3yr} (79% vs. 79%, $P = 0.85$; Table I).

TABLE II. Clinical Characteristics and Outcome in 65 CIMP Classified T-ALL

		CIMP+ N = 40	CIMP- N = 25	P-value	Total N = 65
Gender	male/female	27/13	16/9	ns	43/22
Median age (range, years)		7 (2–17)	8 (1–15)	ns	7 (1–17)
Median WBC x 10 ⁹ /l (range)		77.1 (1.6–825)	228 (6.9–983)	0.03	150 (1.6–983)
MRD day 29	<0.1%/≥0.1%/NA	18/18/4	10/15/0	ns	28/33/4
ETP phenotype ¹	Yes/No/NA	6/28/6	2/20/3	ns	8/48/9
EGIL class ²	Immature/cortical /mature/NA	13/19/2/6	7/11/4/3	ns	20/30/6/9
Follow-up status	CR1 (median follow up; range, months)	35 (50; 2–76)	17 (58; 24–73)		52 (52; 2–76)
	Relapse (median time to relapse, months)	2 (18)	7 (13)		9 (13)
	DCR1	2	1		3
	Induction failure	1	0		1
	Resistant disease	0	0		0
	Dead/alive	5/35	8/17		13/52
pCIR _{3y} (SE)		0.06 ± 0.04	0.29 ± 0.09	0.01	0.15 ± 0.05
pCIDCR1 _{3y} (SE)		0.05 ± 0.04	0.05 ± 0.05	ns	0.05 ± 0.03
pOS _{3y} (SE)		0.87 ± 0.06	0.68 ± 0.1	0.08	0.79 ± 0.05

¹Ref. [15]. ²Ref. [3]. NA, not analyzed; ns, not significant; CR1, first complete remission; DCR1, dead in CR1; SE, standard error; pCIR_{3y}, 3-year cumulative incidence of relapse; pCIDCR1_{3y}, 3-year cumulative incidence of death in CR1; pOS_{3y}, 3-year overall survival rate.

CIMP Classification and Verification

The diagnostic T-ALL samples were analyzed by HumMeth450K arrays, and CpG island methylation phenotype (CIMP) classified by a previously defined panel of 1,293 CpG sites.[6] Twenty-five T-ALL patients were classified as CIMP negative and 40 were classified as CIMP positive (Fig. 1A and Supplementary Table SII). The CIMP-negative samples had a methylation profile very similar to normal CD3⁺ T-cells and CD34⁺ HSCs (Fig. 1A and Supplementary Table SII).

The validity of the methylation arrays was determined by HRM analysis of six gene regions covering a selection of CpG sites with distinct differences in methylation levels between CIMP subgroups in the array CIMP panel. The mean HRM methylation level (%) of the six-gene HRM-CIMP panel correlated well with the percentage of methylated CpG sites in the array-CIMP panel ($R^2 = 0.87$, Fig. 1B).

CIMP Classes and Clinical Characteristics Including T-Cell Maturation Stage

CIMP-negative (n = 25) and CIMP-positive (n = 40) T-ALL patients showed no significant differences regarding gender, age, or MRD status (<0.1%/≥0.1%) at day 29, but higher WBC values were found in CIMP-negative compared to CIMP-positive cases (median WBC: 228 vs. 77 × 10⁹/l; $P = 0.03$) (Table II).

Among the 65 methylation analyzed patients, 56 had sufficient data to allow for T-cell maturation stage classification (EGIL class and ETP phenotype). We found no significant association between T-cell maturation stage (EGIL class or ETP phenotype) and CIMP class (Table II).

Survival Analysis

The pCIR_{3y} analysis of patients classified according to CIMP status showed significant differences ($P = 0.01$) with the best prognosis for CIMP-positive cases (pCIR_{3y} 6%) compared to CIMP-negative cases (pCIR_{3y} 29%) (Fig. 2A and Table II).

The pOS_{3y} also differed between the CIMP groups (pOS_{3y}: 87% for CIMP positive and 68% for CIMP negative), but was not significant ($P = 0.08$) (Fig. 2A and Table II). The cumulative incidence of death in remission (pCIDCR1_{3y}) did not differ between CIMP subgroups ($P = 0.86$) (Table II) and was therefore not corrected for.

MRD at day 29 is used for risk group stratification in the current NOPHO ALL 2008 protocol. The pCIR_{3y} and pOS_{3y} analysis of the 61 samples with available MRD data showed that MRD ≥0.1% day 29 identified individuals with poor outcome (pCIR_{3y} MRD ≥0.1%, 30% and MRD <0.1%, 0%, $P = 0.003$, and pOS_{3y} MRD ≥0.1%, 65% and MRD <0.1%, 100%, $P = 0.001$; Fig. 2B).

By combining CIMP status at diagnosis and MRD level at day 29, this prognostic information was enhanced ($P < 0.001$; Fig. 2C). The MRD ≥0.1%/CIMP negative group could be identified as a group with very poor prognosis (pCIR_{3y} 50% and pOS_{3y} 45%), whereas the MRD ≥0.1%/CIMP-positive patients had a much better outcome (pCIR_{3y} 12% and pOS_{3y} 83%) ($P = 0.02$ [CIR], $P = 0.03$ [OS]; Table III). Interestingly, CIMP status did not play a role for patients with low MRD, since there were no events in the MRD <0.1% group, irrespective of the CIMP status (pCIR_{3y} 100% and pOS_{3y} 100%) (Fig. 2C and Table III).

When comparing the MRD ≥0.1%/CIMP negative group with the MRD ≥0.1%/CIMP positive group for clinical data including age, gender, WBC at diagnosis, ETP, and EGIL status, only age and WBC differed significantly. Thus, the MRD ≥0.1%/CIMP negative group had higher WBC count (Median: 241 × 10⁹/l vs. 72 × 10⁹/l, $P = 0.004$) and were younger (Median: 7.0 vs. 11.5 years, $P = 0.02$) at diagnosis compared with the MRD ≥0.1%/CIMP positive group (Table III).

To further study their relation, WBC were plotted together with CIMP and MRD status (Fig. 3) showing that our identified poor prognostic, MRD ≥0.1%/CIMP negative group, was characterized by high WBC counts at diagnosis.

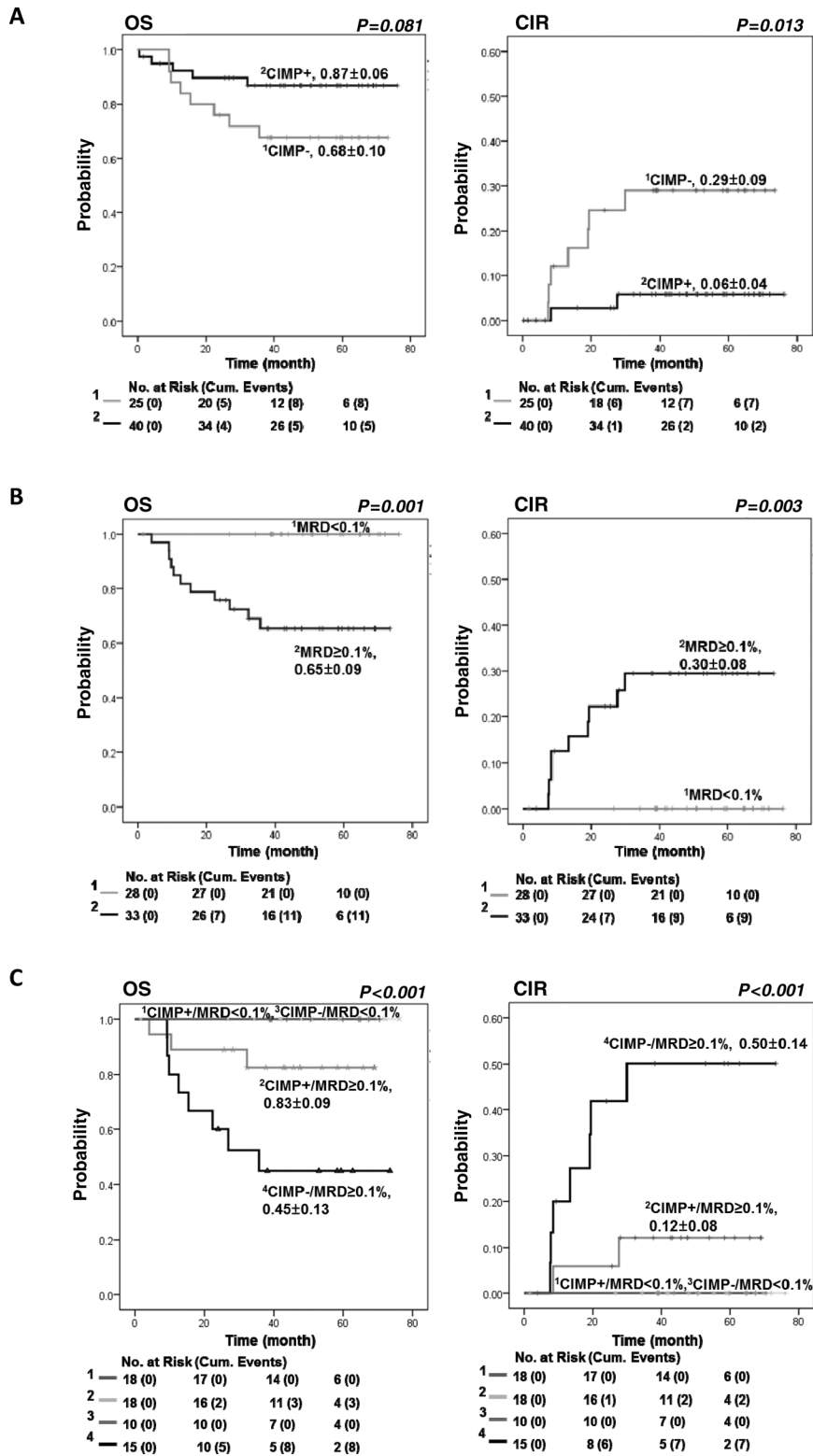


Fig. 2. Survival analysis based on CIMP classification and MRD level at day 29. Overall survival (OS) and cumulative incidence of relapse (CIR) in (A) CIMP subgroups, CIMP negative (n = 25), and CIMP positive (n = 40). (B) MRD ≥0.1% (n = 33) and MRD <0.1% (n = 28) subgroups at day 29. (C) CIMP+/MRD ≥0.1% (n = 18), CIMP+/MRD <0.1% (n = 18), CIMP-/MRD ≥0.1% (n = 15), CIMP-/MRD <0.1% (n = 10) subgroups.

TABLE III. Clinical Characteristics, Outcome, and Therapy Stratification in 61 T-ALL in CIMP/MRD Subgroups

		MRD \geq 0.1%			MRD < 0.1%		
		CIMP+ (N = 18)	CIMP- (N = 15)	P-value	CIMP+ (N = 18)	CIMP- (N = 10)	P-value
Gender	male/female	13/5	10/5	ns	12/6	6/4	ns
Median age (range, years)		11.5 (2–17)	7 (1–15)	0.02	5 (2–14)	12 (1–15)	ns
Median WBC $\times 10^9/l$ (range)		72 (3–560)	241 (52–983)	0.004	74 (4–825)	126 (7–492)	ns
ETP ¹	Yes/no/NA	3/14/1	1/13/1	ns	2/11/5	1/7/2	ns
EGIL ²	Immature/cortical /mature/NA	6/9/2/1	4/8/2/1	ns	5/8/0/5	3/3/2/2	ns
Therapy stratification after day 29	Standard risk	0	0		0	0	
	Intermediate risk	0	0		17	9	
	High-risk chemo	16	13		1	1	
	High-risk SCTCR1	2	2		0	0	
	CR1	15	7		18	10	
Follow-up status	Induction failure	0	0		0	0	
	Resistant disease	0	0		0	0	
	Relapse	2	7		0	0	
	DCR1	1	1		0	0	
	Dead/alive	3/15	8/7		0/18	0/10	
pCIR _{3y} (SE)		0.12 \pm 0.08	0.50 \pm 0.14	0.02	0	0	ns
pOS _{3y} (SE)		0.83 \pm 0.09	0.45 \pm 0.13	0.03	1	1	ns

¹Ref. [15]. ²Ref. [3]. NA, not analyzed; SCTCR1, stem cell transplantation in CR1; pCIR_{3y}, 3-year cumulative incidence of relapse; pOS_{3y}, 3-year overall survival rate; SE, standard error.

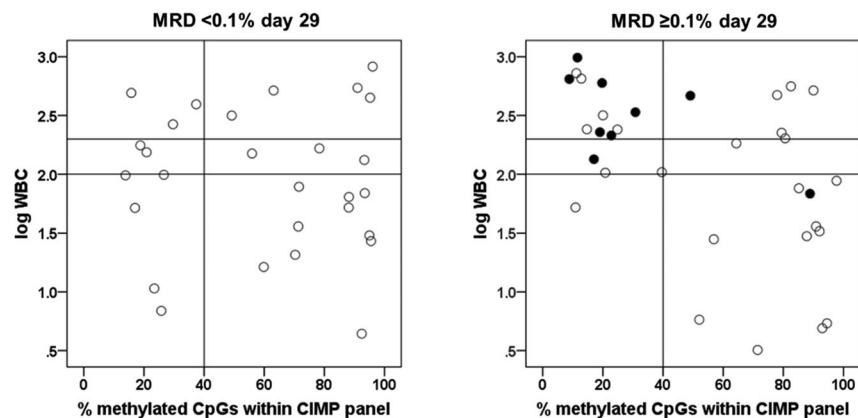


Fig. 3. CIMP status, white blood cell count, and MRD level. Percentage of methylated CpGs in the CIMP panel versus log WBC count at diagnosis in MRD < 0.1% and MRD \geq 0.1% samples at day 29. CIMP status (\leq 40% or $>$ 40% methylated CpGs in CIMP panel) and WBC at $100 \times 10^9/l$ and $200 \times 10^9/l$ are separated with lines in the figure; relapses are marked with filled circles.

DISCUSSION

Even though the outcome of pediatric T-ALL has improved significantly over the last decades, the high intensive chemotherapy treatment currently used may cause short- or long-term adverse effects. Furthermore, the cure rate of relapsed T-ALL has remained dismal, calling for improved risk stratification that allow those with the highest risk of relapse to be allocated to novel treatment strategies and/or HSC transplantation in first remission.

We have previously shown a strong prognostic impact of DNA methylation classification into CIMP groups in pediatric T-ALL patients treated according to the NOPHO ALL 1992/2000 protocols.[6] However, in those protocols the cure rate for T-cell ALL was generally poor. The present study in an independent Nordic cohort confirms that CIMP classification is a relevant clinical prognostic factor for childhood T-ALL, also in the setting of contemporary treatment pro-

grams that integrate MRD measurements in the risk group allocation. Although the overall pCIR rate has been improved with the current ALL2008 protocol, CIMP status at diagnosis was a significant prognostic factor. Of special importance is that a combination of CIMP status at diagnosis and MRD status at treatment day 29 could further identify patients with significantly different clinical outcomes. In the current NOPHO ALL 2008 protocol (NCT00816049) for T-ALL, MRD levels $>$ 0.1% at day 29 stratifies patients to an intensified treatment schedule.[1] Whether a future combination of CIMP classification and postinduction MRD levels will allow down grading of CIMP-positive T-ALL patients to less intensive chemotherapy must be evaluated in prospective trials. Importantly, the prognosis for CIMP negative/MRD \geq 0.1% patients is so dismal that novel treatment approaches are needed. Methylation status does not seem to influence the initial response to therapy since there were no significant difference in

remaining leukemic cells after induction therapy (MRD status day 29) in CIMP subgroups. The vast majority of leukemic cells were eliminated by therapy regardless of methylation status. However, the relapse frequency was higher in the CIMP-negative leukemias. The reason for this we can only speculate upon but it might reflect different efficiency of eliminating leukemic initiating cells in CIMP-positive and CIMP-negative leukemias.

Our CIMP profile reflecting 1,293 CpGs is based on methylation array analysis, and is a robust and informative, but yet relatively expensive (\approx 350 USD/sample) and time consuming technique (3 days). As an alternative technique for CIMP classification and to confirm the array CIMP classification, we used HRM analysis based on six gene regions selected from the CIMP panel to represent CpG sites with distinct differences in methylation levels between the CIMP subgroups. The HRM classification correlated well with the array-based CIMP classification, and it can be run on single samples and be completed within 1 day. Therefore, HRM CIMP classification is an attractive alternative to array CIMP classification. However, the array analysis has its advantage in gaining information about genome-wide methylation patterns and can be used for additional analysis, including copy number alterations.[21]

Prognostic impact of DNA methylation signatures has also been observed in other hematological malignancies, including myelodysplasia, acute myeloid leukemia (AML), and BCP-ALL, and methylation modifiers play a therapeutic role.[8,22,23] The prognostic relevance of CIMP classification in pediatric T-ALL shown in this paper confirms our previous finding in a separate T-ALL cohort. However, there are conflicting results of CIMP status and prognosis in pediatric T-ALL,[11,12] but these studies used 14–20 predefined genes for methylation classification in contrast to our genome-wide array approach. In order to gain the full potential of DNA methylation classification as a prognostic marker in T-ALL, further validation of larger cohorts is needed.

A deeper functional understanding of the complex role of DNA methylation aberrations in the development and progression of hematological malignancies is still missing.

Genetic subtypes of childhood ALL and AML have been associated with DNA methylation signatures.[24–26] A recent study by Amabile et al. showed that in a murine chronic myeloid leukemia (CML) model, induction of BCR-ABL could trigger DNA methylation changes. Furthermore, it was shown that aberrant DNA methylation had the potential to contribute to leukemia progression in primary CML cells.[27]

Genome-wide methylation studies of the mouse hematopoietic system identified specific signatures associated with lineage commitment and T-cell maturation stage and showed that cells committed to a myeloid lineage had lower global methylation levels than cells committed to a lymphoid lineage.[28] Immature T-ALL and ETP-ALL, with myeloid molecular characteristics, have been associated with poor outcome.[15,29–32] However, we found no significant association of CIMP subgroup with T-cell maturation stage based on ETP phenotype or EGIL classification.

The MRD \geq 0.1%/CIMP negative group was associated with high WBC counts at diagnosis. WBC count is not used as a stratifier for therapy in T-ALL in the current protocol but analysis are ongoing whether or not it should be included in fu-

ture risk assessments.[33] The CIMP, MRD, and WBC factors partly covaried and it seems that combining two of these three factors may be useful to further define different risk groups. However, larger sample size is needed to study their relation/independence in detail by multivariate Cox regression analysis.

DNA methylation alterations have been shown to accumulate in cells over time,[18,34,35] and our recent publication on a T-cell culture model for immortalization showed overlapping methylation alterations in long-term cultured immortalized cell cultures and CIMP-positive T-ALL.[18] Age-related alterations in the DNA methylation patterns have been studied in the ontogeny of HSCs showing a significant global DNA hypermethylation in older HSCs. The same study found that hypermethylation occurred at polycomb repressive complex 2 (PRC2) target loci upon forced proliferation of HSCs.[36] Our CIMP profile is enriched for CpG sites located in polycomb target genes.[6] The fact that the CIMP-negative subgroup showed a similar methylation profile to normal CD3⁺ and CD34⁺/CD38⁻ cells suggests that these cases might have undergone fewer rounds of replications.

Recurring loss-of-function mutations and deletions have been identified in the PRC2 components (EZH2, SUZ12, and EED) in pediatric T-ALL.[32] In adult T-ALL, mutations in RUNX1 and the DNA methyltransferase DNMT3A gene were associated with poor prognosis.[37,38] Recently, DNMT3A mutated preleukemic HSCs have been identified in AML that were resistant to chemotherapy and persisted in remission, indicating that they might represent a reservoir from which relapse arises.[39] It remains to be determined, if mutations or genetic aberrations in DNA methylation associated genes or oncogenes can be associated with CIMP subgroups.

To further gain insight into characteristics of CIMP-positive and CIMP-negative cells, the next step will be to perform transcriptome and genome sequence analysis of diagnostic and relapse samples and relate to DNA methylation and histone modification alterations. *In vitro* cytotoxicity analysis of CIMP-positive and CIMP-negative cells will gain deepened knowledge of drug resistance and cell signaling response in cells of different methylation status.

To conclude, in this collaboration study between the Nordic countries we show that DNA methylation patterns in diagnostic T-ALL samples hold important prognostic information. Of special interest was that CIMP status at diagnosis could separate postinduction MRD-positive ($>$ 0.1%, day 29) patients into two risk groups. Further dissection of the biology behind CIMP status will hopefully result in the identification of novel therapeutic targets.

ACKNOWLEDGMENTS

This study was supported by grants from the Swedish Cancer Society, the Swedish Research Council (Dnr 340-2013-5185), the Swedish Childhood Cancer Foundation, the Medical Faculty of Umeå University, the Danish Childhood Cancer Foundation, Lion's Cancer Research Foundation (Umeå), Umeå Paediatric Clinic Research Foundation, Magnus Bergvalls stiftelse, and Uppsala-Umeå Comprehensive Cancer Consortium. Financial support was provided through regional agreement between Umeå University and Västerbotten County Council on

cooperation in the field of medicine, odontology, and health. We thank Helene Sandström for laboratory assistance.

Authors' contributions

M. B., Z. H., M. Hu., G. R., E. F., and S. D. conceived and designed the experiments; U. N. N., K. S., A. Å., J. K., H. O. M., H. M., M. He. were involved in collection of data; S. D. and Z. H. performed the experiments; M. B., Z. H., M. L., E. F., and S. D. analyzed the data; and M. B., Z. H., K. S., G. R., and S. D. wrote the paper with contribution from all coauthors.

REFERENCES

- Toft N, Birgens H, Abrahamsson J, Bernell P, Griskevicius L, Hallbook H, Heyman M, Holm MS, Hulegårdh E, Klausen TW, Marquart HV, Jonsson OG, Nielsen OJ, Quist-Paulsen P, Taskinen M, Vaitkeviciene G, Vettenranta K, Asberg A, Schmiegelow K. Risk group assignment differs for children and adults 1–45 yr with acute lymphoblastic leukemia treated by the NOPHO ALL-2008 protocol. *Eur J Haematol* 2013;90:404–412.
- Van Vlierbergh P, Ferrando A. The molecular basis of T cell acute lymphoblastic leukemia. *J Clin Invest* 2012;122:3398–3406.
- Bene MC, Castoldi G, Knapp W, Ludwig WD, Matutes E, Orfao A, van't Veer MB. Proposals for the immunological classification of acute leukemias. European Group for the Immunological Characterization of Leukemias (EGIL). *Leukemia* 1995;9:1783–1786.
- Esteller M. Epigenetics in cancer. *N Engl J Med* 2008;358:1148–1159.
- Borssén M, Cullman I, Noren-Nyström U, Sundström C, Porwit A, Forestier E, Roos G. hTERT promoter methylation and telomere length in childhood acute lymphoblastic leukemia associations with immunophenotype and cytogenetic subgroup. *Exp Hematol* 2011;39:1144–1151.
- Borssén M, Palmqvist L, Karrman K, Abrahamsson J, Behrendtz M, Heldrup J, Forestier E, Roos G, Degerman S. Promoter DNA methylation pattern identifies prognostic subgroups in childhood T-cell acute lymphoblastic leukemia. *PLoS One* 2013;8:e65373.
- Deneberg S, Grovdal M, Karim M, Jansson M, Nahi H, Corbacioglu A, Gaidzik V, Dohner K, Paul C, Ekstrom TJ, Hellstrom-Lindberg E, Lehmann S. Gene-specific and global methylation patterns predict outcome in patients with acute myeloid leukemia. *Leukemia* 2010;24:932–941.
- Milani L, Lundmark A, Kialainen A, Nordlund J, Flaegstad T, Forestier E, Heyman M, Jonmundsson G, Kanerva J, Schmiegelow K, Soderhall S, Gustafsson MG, Lonnerholm G, Syvanen AC. DNA methylation for subtype classification and prediction of treatment outcome in patients with childhood acute lymphoblastic leukemia. *Blood* 2010;115:1214–1225.
- Shen L, Kantarjian H, Guo Y, Lin E, Shan J, Huang X, Berry D, Ahmed S, Zhu W, Pierce S, Kondo Y, Oki Y, Jelinek J, Saba H, Estey E, Issa JP. DNA methylation predicts survival and response to therapy in patients with myelodysplastic syndromes. *J Clin Oncol* 2010;28:605–613.
- Hughes LA, Melotte V, de Schrijver J, de Maat M, Smit VT, Bovee JV, French PJ, van den Brandt PA, Schouten LJ, de Meyer T, van Criekinge W, Ahuja N, Herman JG, Weijnen MP, van Engeland M. The CpG island methylator phenotype: What's in a name? *Cancer Res* 2013;73:5858–5868.
- Kraszewska MD, Dawidowska M, Larnonnie NS, Kosmalska M, Sedek L, Szczepaniak M, Grzeszczak W, Langerak AW, Szczepanski T, Witt M, Polish Pediatric Leukemia Lymphoma Study G. DNA methylation pattern is altered in childhood T-cell acute lymphoblastic leukemia patients as compared with normal thymic subsets: Insights into CpG island methylator phenotype in T-ALL. *Leukemia* 2012;26:367–371.
- Takeuchi S, Matsushita M, Zimmermann M, Ikezoe T, Komatsu N, Seriu T, Schrappe M, Bartram CR, Koefler HP. Clinical significance of aberrant DNA methylation in childhood acute lymphoblastic leukemia. *Leuk Res* 2011;35:1345–1349.
- van der Velden VH, Cazzaniga G, Schrauder A, Hancock J, Bader P, Panzer-Grumayer ER, Flohr T, Sutton R, Cave H, Madsen HO, Cayuela JM, Trka J, Eckert C, Foroni L, Zur Stadt U, Beldjord K, Raff T, van der Schoot CE, van Dongen JJ. European Study Group on MRDdALL. Analysis of minimal residual disease by Ig/TCR gene rearrangements: Guidelines for interpretation of real-time quantitative PCR data. *Leukemia* 2007;21:604–611.
- Bjorklund E, Matinlauri I, Tierens A, Axelsson S, Forestier E, Jacobsson S, Ahlberg AJ, Kaurec G, Mantymaa P, Osnes L, Penttila TL, Marquart H, Savolainen ER, Siitonen S, Torikka K, Mazur J, Porwit A. Quality control of flow cytometry data analysis for evaluation of minimal residual disease in bone marrow from acute leukemia patients during treatment. *J Pediatr Hematol Oncol* 2009;31:406–415.
- Coustan-Smith E, Mullighan CG, Onciu M, Behm FG, Raimondi SC, Pei D, Cheng C, Su X, Rubnitz JE, Basso G, Biondi A, Pui CH, Downing JR, Campana D. Early T-cell precursor leukaemia: A subtype of very high-risk acute lymphoblastic leukaemia. *Lancet Oncol* 2009;10:147–156.
- Chopra A, Bakshi S, Pramanik SK, Pandey RM, Singh S, Gajendra S, Gogia A, Chandramohan J, Sharma A, Kumar L, Seth R, Rai S, Kumar R. Immunophenotypic analysis of T-acute lymphoblastic leukemia. A CD5-based ETP-ALL perspective of non-ETP T-ALL. *Eur J Haematol* 2014;92:211–218.
- Hayhoe FG. Classification of acute leukaemias. *Blood Rev* 1988;2:186–193.
- Degerman S, Landfors M, Siwicki JK, Revie J, Borssen M, Evelonn E, Forestier E, Chrzanoska KH, Ryden P, Keith WN, Roos G. Immortalization of T-cells is accompanied by gradual changes in CpG methylation resulting in a profile resembling a subset of T-cell leukemias. *Neoplasia* 2014;16:606–615.
- Nordlund J, Backlin CL, Wahlberg P, Busche S, Berglund EC, Eloranta ML, Flaegstad T, Forestier E, Frost BM, Harila-Saari A, Heyman M, Jonsson OG, Larsson R, Palle J, Ronnblom L, Schmiegelow K, Sinnott D, Soderhall S, Pastinen T, Gustafsson MG et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. *Genome Biol* 2013;14:r105.
- Frandsen TL, Heyman M, Abrahamsson J, Vettenranta K, Asberg A, Vaitkeviciene G, Pruunsild K, Toft N, Birgens H, Hallbook H, Quist-Paulsen P, Griskevicius L, Helt L, Hansen BV, Schmiegelow K. Complying with the European Clinical Trials directive while surviving the administrative pressure: An alternative approach to toxicity registration in a cancer trial. *Eur J Cancer* 2014;50:251–259.
- Feber A, Guilhamon P, Lechner M, Fenton T, Wilson GA, Thirlwell C, Morris TJ, Flanagan AM, Teschendorff AE, Kelly JD, Beck S. Using high-density DNA methylation arrays to profile copy number alterations. *Genome Biol* 2014;15:R30.
- Figuerola ME, Lugthart S, Li Y, Erpelinck-Verschueren C, Deng X, Christos PJ, Schifano E, Booth J, van Putten W, Skrabanek L, Campagne F, Mazumdar M, Greally JM, Valk PJ, Lowenberg B, Delwel R, Melnick A. DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer Cell* 2010;17:13–27.
- Griffiths EA, Gore SD. Epigenetic therapies in MDS and AML. *Adv Exp Med Biol* 2013;754:253–283.
- Figuerola ME, Chen SC, Andersson AK, Phillips LA, Li Y, Sotzen J, Kundu M, Downing JR, Melnick A, Mullighan CG. Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia. *J Clin Invest* 2013;123:3099–3111.
- Nordlund J, Backlin CL, Zachariadis V, Caveller L, Dahlberg J, Ofverholm I, Barbany G, Nordgren A, Overnas E, Abrahamsson J, Flaegstad T, Heyman MM, Jonsson OG, Kanerva J, Larsson R, Palle J, Schmiegelow K, Gustafsson MG, Lonnerholm G, Forestier E et al. DNA methylation-based subtype prediction for pediatric acute lymphoblastic leukemia. *Clin Epigenetics* 2015;7:11.
- Qu X, Davison J, Du L, Storer B, Stirewalt DL, Heimfeld S, Estey E, Appelbaum FR, Fang M. Identification of differentially methylated markers among cytogenetic risk groups of acute myeloid leukemia. *Epigenetics* 2015;10:526–535.
- Amabile G, Di Ruscio A, Muller F, Welner RS, Yang H, Ebralidze AK, Zhang H, Levantini E, Qi L, Martinelli G, Brummelkamp T, Le Beau MM, Figuerola ME, Bock C, Tenen DG. Dissecting the role of aberrant DNA methylation in human leukaemia. *Nat Commun* 2015;6:7091.
- Ji H, Ehrlich LI, Seita J, Murakami P, Doi A, Lindau P, Lee H, Aryee MJ, Irizarry RA, Kim K, Rossi DJ, Inlay MA, Serwold T, Karsunky H, Ho L, Daley GQ, Weissman IL, Feinberg AP. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* 2010;467:338–342.
- Allen A, Sireci A, Colovai A, Pinkney K, Sulis M, Bhagat G, Alobaid B. Early T-cell precursor leukemia/lymphoma in adults and children. *Leuk Res* 2013;37:1027–1034.
- Haydu JE, Ferrando AA. Early T-cell precursor acute lymphoblastic leukaemia. *Curr Opin Hematol* 2013;20:369–373.
- Patrick K, Wade R, Goulden N, Mitchell C, Moorman AV, Rowntree C, Jenkinson S, Hough R, Vora A. Outcome for children and young people with Early T-cell precursor acute lymphoblastic leukemia treated on a contemporary protocol, UKALL 2003. *Br J Haematol* 2014;166:421–424.
- Zhang J, Ding L, Holmfeldt L, Wu G, Heatley SL, Payne-Turner D, Easton J, Chen X, Wang J, Rusch M, Lu C, Chen SC, Wei L, Collins-Underwood JR, Ma J, Roberts KG, Pounds SB, Ulyanov A, Becksfort J, Gupta P et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* 2012;481:157–163.
- Hastings C, Gaynon PS, Nachman JB, Sather HN, Lu X, Devidas M, Seibel NL. Increased post-induction intensification improves outcome in children and adolescents with a markedly elevated white blood cell count ($>1 = 200 \times 10^9/l$) with T cell acute lymphoblastic leukaemia but not B cell disease: a report from the Children's Oncology Group. *Br J Haematol* 2015;168:533–546.
- Tommasi S, Zheng A, Weninger A, Bates SE, Li XA, Wu X, Hollstein M, Besaratinia A. Mammalian cells acquire epigenetic hallmarks of human cancer during immortalization. *Nucleic Acids Res* 2013;41:182–195.
- Koch CM, Jousset S, Schellenberg A, Lin Q, Zenke M, Wagner W. Monitoring of cellular senescence by DNA-methylation at specific CpG sites. *Aging Cell* 2012;11:366–369.
- Beerman I, Bock C, Garrison BS, Smith ZD, Gu H, Meissner A, Rossi DJ. Proliferation-dependent alterations of the DNA methylation landscape underlie hematopoietic stem cell aging. *Cell Stem Cell* 2013;12:413–425.
- Grossmann V, Haferlach C, Weissmann S, Roller A, Schindela S, Poetzinger F, Stadler K, Bellos F, Kern W, Haferlach T, Schnittger S, Kohlmann A. The molecular profile of adult T-cell acute lymphoblastic leukemia: Mutations in RUNX1 and DNMT3A are associated with poor prognosis in T-ALL. *Genes Chromosomes Cancer* 2013;52:410–422.
- Mok MM, Du L, Wang CQ, Tergaonkar V, Liu TC, Yin Kham SK, Sanda T, Yeoh AE, Osato M. RUNX1 point mutations potentially identify a subset of early immature T-cell acute lymphoblastic leukaemia that may originate from differentiated T-cells. *Gene* 2014;545:111–116.
- Shlush LI, Zandi S, Mitchell A, Chen W, Brandwein JM, Gupta V, Kennedy JA, Schimmer AD, Schuh AC, Yee KW, McLeod JL, Doedens M, Medeiros JI, Marke R, Kim HJ, Lee K, McPherson JD, Hudson TJ, Consortium HP-LGP, Brown AM et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature*, 2014;506:328–333.

DNA methylation holds prognostic information in relapsed precursor B-cell acute lymphoblastic leukemia

Magnus Borssén¹, Jessica Nordlund², Zahra Haider¹, Mattias Landfors¹, Pär Larsson¹, NOPHO collaborators³, Erik Forestier¹, Mats Heyman⁴, Magnus Hultdin¹, Gudmar Lönnnerholm², Ann-Christine Syvänen², and Sofie Degerman^{1*}.

¹ Department of Medical Biosciences, Umeå University, Umeå, Sweden,

² Department of Medical Sciences, Uppsala University, Sweden

³ Nordic society of pediatric hematology and oncology

⁴ Department of Paediatrics, University Hospitals, Astrid Lindgrens Barnsjukhus, Stockholm, Sweden

*Corresponding author Sofie Degerman PhD Medical Biosciences, Pathology, Umeå University, Blg 6M, 2nd floor SE-90185 Umeå SWEDEN, sofie.degerman@umu.se, phone +4690-7852873, fax +46907854484.

Key words: DNA methylation, BCP-ALL, CIMP, relapse, time to relapse, relapse location, childhood leukemia, prognosis

Abstract

Background: Few biological markers exist that enable risk stratification of patients with relapsed B-cell precursor acute lymphoblastic leukemia (BCP-ALL). Here, the prognostic relevance of CIMP (CpG island methylator phenotype) classification was investigated at diagnosis and at relapse.

Procedure: Six-hundred-one diagnostic BCP-ALL samples from Nordic pediatric patients treated by the NOPHO ALL1992/2000 protocols were CIMP classified and analyzed in relation to clinical data.

Results: At diagnosis, CIMP subgroups were non-randomly distributed among the cytogenetic aberration groups and the CIMP low profile was associated with lower overall survival ($p=0.019$). Most importantly, relapsed patients with a CIMP low profile at diagnosis had an unfavorable response to treatment (pOS_{5yr} 33%) compared to CIMP high patients (pOS_{5yr} 65%) ($p=0.001$). Moreover, CIMP status further separated the prognosis of individuals with very early and early relapses.

Conclusions: CIMP classification at diagnosis holds prognostic relevance in pediatric BCP-ALL and is a strong candidate for further risk classification at relapse.

Introduction

With contemporary combination chemotherapy protocols for childhood acute lymphoblastic leukemia (ALL) more than 80% of patients survive. Stratification based on cytogenetic features, immunophenotype and molecular detection of therapy response has allowed intensified treatment among high-risk patients and poor responders (Schmiegelow et al., 2010).

Despite improved treatment response, relapse still occur. Fewer well-defined risk factors exist enabling risk stratification of patients at relapse, as compared to at

diagnosis. The prognosis is also significantly worse after relapse with a crude survival rate of $\approx 55\%$ in the Nordic countries (Oskarsson et al., 2016).

An accumulating body of evidence has highlighted the importance of epigenetics in cancer biology. DNA methylation is a central epigenetic mechanism and the most widely studied epigenetic process (Guillamot et al., 2016). We have previously shown prognostic relevance of promoter associated DNA methylation in T-cell acute lymphoblastic leukemia (T-ALL) (Borssén et al., 2013). The CIMP (CpG Island Methylator Phenotype) low group

with a methylation profile resembling that of to normal CD3+ cells displayed a worse prognosis as compared with patients exhibiting the CIMP high phenotype. Importantly, in a recent Nordic collaboration project the prognostic potential of CIMP classification was confirmed in patients treated by the current NOPHO (Nordic society of pediatric hematology and oncology) ALL 2008 protocol where CIMP classification allowed subgrouping of high-risk T-ALL patients (MRD>0.1% at day 29) (Borssen et al., 2016). In this study we aimed to investigate whether the prognostic DNA methylation profile identified in T-ALL patients also have prognostic value in B-cell precursor acute lymphoblastic leukemia (BCP-ALL).

Materials and methods

BCP-ALL samples

This study included 601 pediatric patients (1-18 years) diagnosed with B-cell precursor ALL (BCP-ALL) between years 1996 and 2008 in the Nordic countries and treated according to the common NOPHO (Nordic society of pediatric hematology and oncology) ALL 1992 and 2000 protocols (Nordlund et al., 2013; Schmiegelow et al., 2010). Clinical follow up data was extracted from the NOPHO leukemia registry in June 2016 and the mean follow up time for patients was 115 months (range 0-221). The regional and/or national ethics committees approved the study, and the patients and/or their guardians provided informed consent in accordance with the Declaration of Helsinki.

Methylation array analysis and CIMP classification

Infinium HumanMethylation 450k BeadChip (Illumina, San Diego, CA, USA) array analysis was previously performed on lymphocytes isolated from bone marrow or peripheral blood cells from ALL patients that included >80% leukemic blasts at diagnosis. The data is publicly available at

Gene Expression Omnibus (GEO) with accession number GSE49031(Nordlund et al., 2013). Array data was preprocessed (normalized, filtered for CpG probes that align to multiple genomic locations or within 3 bp from a SNP, and CpG located on X and Y chromosomes were excluded) and samples were classified according to our previously defined CIMP panel (Borssen et al., 2016).

The CIMP subgroups were defined based on percentage of methylated CpG sites (methylated CpG site defined as a β -value >0,4) within the 1293 CpG site CIMP panel, with 40% or more being classified as CIMP high in T-ALL. Since approximately 1/3 of the CpG sites in the CIMP panel seemed relevant only for T-ALL, the CIMP percentage separating CIMP low and CIMP high was adjusted for the BCP-ALL dataset. Samples with $\leq 25\%$ methylated CpG sites were classified as CIMP low, and samples > 25% methylated CpG sites were classified as CIMP high (Supplementary Figure S1).

Statistical analysis

Methylation data analysis was carried out in the R environment (v2.15.0). The Statistical Package for the Social Sciences (SPSS Inc., Chicago, IL) software was used for the statistical analyses. The chi-square and the Mann-Whitney U tests were used to determine whether differences among subgroups existed for discrete and continuous variables, respectively. Estimates of event free survival (pEFS), overall survival (pOS), and cumulative incidence of relapse (pCIR) are given at 5 years and were calculated using the Kaplan-Meier method and subgroups compared using the log rank test. The significance limit for two-sided p-values was set to <0.05 in all tests. Time in first remission (CR1) was defined as time from diagnosis until first event, comprising induction failure, resistant disease, death in remission, relapse, or second malignant neoplasm. All-cause mortality was the endpoint in the analysis of overall survival (OS).

Results and discussion

In this study 601 diagnostic BCP-ALL samples from children (1-18 years) diagnosed in the Nordic countries between years 1996 and 2008 and treated according to the NOPHO ALL 1992 and 2000 protocols were included (Nordlund et al., 2013).

Methylation data from Illumina Human methylation 450K arrays were used for classification of samples as CIMP high and CIMP low by analysis of our previously described CIMP panel, consisting of a subset of 1293 CpG sites with shown prognostic relevance in T-ALL (Borssen et al., 2016). The diagnostic BCP-ALL samples were classified according to the adjusted BCP-ALL CIMP definition (Supplementary Figure S1) and resulted in 175 CIMP low and 426 CIMP high patients. The cytogenetic groups were non-randomly distributed among CIMP subgroups ($p < 0.001$) (Supplementary Table S1). CIMP high samples were enriched for patients positive for ETV6-RUNX1 (t(12;21)) translocation and CIMP low were enriched for patients positive for BCR-ABL1 (t(9;22)) and TCF3-PBX1 (t(1;19)) (Supplementary Table S1). Of special notice is that the majority of infant leukemias (often with MLL/11q23 rearrangements) were classified as CIMP low, but excluded in this analysis due to the use of different protocols for treatment of these patients (data not shown).

The overall survival was significantly lower in the CIMP low group compared with the CIMP high group (pOS_{5yr} 85% vs 92%, $p = 0.019$, Figure 1A), but CIMP status at diagnosis did not predict relapse ($pCIR_{5yr}$) or event free survival ($pEFS_{5yr}$). Considering the skewed distribution of cytogenetic aberrations between CIMP groups, detection of differences in clinical outcome between the CIMP groups was expected. However, the fact that CIMP status did not predict $pCIR_{5yr}$ but showed differences in pOS_{5yr} indicated differences

between CIMP groups regarding response to relapse treatment.

We further analyzed all relapsed cases in our cohort ($n = 137$) separately, of which 42 were classified as CIMP low and 95 CIMP high at diagnosis. There was no significant difference between the groups with regard to age at diagnosis, white blood cell counts at diagnosis, initial therapy stratification, time to relapse from diagnosis, and whether or not patients had undergone bone marrow transplantation after relapse (Table 1). Although non-significant, a trend could be observed for relapse site location between CIMP groups ($p = 0.077$), where 73% of CIMP low relapses were isolated bone marrow relapses in contrast to 53% of the CIMP high relapses (Table 1).

Overall survival analysis showed that pOS_{5yr} among all relapsed patients were 55%, but for CIMP low patients only 33% ($\pm 0.8\%$), as compared with 65% ($\pm 5\%$) for CIMP high patients ($p = 0.001$) (Figure 1A).

Even though current relapse treatment protocols do not include cytogenetic aberrations in risk stratification, several studies have shown that high risk genetic aberrations at diagnosis has prognostic significance even at relapse (Oskarsson et al., 2016). This was further evaluated in separate overall survival analysis after relapse based on the primary risk stratification at diagnosis (Supplementary Figure S2). Patients initially (after primary diagnosis) treated with SR/IR (standard risk/intermediate risk) protocols had an pOS_{5y} of 63% ($\pm 6\%$) and HR (high risk) patients 42% ($\pm 7\%$) after relapse ($p = 0.008$) (data not shown). Of importance, CIMP status at diagnosis could add further prognostic information at relapse. The SR/IR patients classified as CIMP high at diagnosis had an OS_{5y} of 69% ($\pm 6\%$) after relapse compared to 45% ($\pm 12\%$) for SR/IR CIMP low patients ($p = 0.06$) (Supplementary Figure S2 A). This disparity was even more pronounced in the HR group. HR patients classified as CIMP high at diagnosis had an OS_{5y} of 56% (\pm

9%) after relapse compared with only 21% (+/-9%) for HR/CIMP low patients ($p=0.016$) (Supplementary Figure S2 B). These results indicate that CIMP status at diagnosis holds prognostic information at relapse regardless of initial risk stratification at diagnosis, although it did not reach significance in the SR/IR group. In line with our results, a previous small study by Sandoval et al., showed that a CIMP low like phenotype was associated to inferior clinical outcome (Sandoval et al., 2013).

Time in first complete remission is the most important prognostic factor in relapsed ALL and is a main stratifying factor for relapse therapy (Nguyen et al., 2008; Oskarsson et al., 2016). In our cohort, very early relapse (<18 months from primary diagnosis) had the worst outcome followed by early relapse (≥ 18 months from diagnosis to ≥ 6 months post treatment) with a pOS_{5y} at 26% (+/-9%) and 46% (+/-7%), respectively. Patients relapsing later than six months after completion of treatment (late relapse) had a survival rate of 77% (+/-6%). By combining CIMP status and time from diagnosis to relapse into 6 groups, CIMP status at diagnosis clearly had an impact on outcome ($p \leq 0,0001$). Our CIMP profile could separate very early and early relapse cases into two groups with different outcome (Figure 1C) ranging from only pOS_{5y} of 9% in CIMP low patients with very early relapse to pOS_{5y} of 59% in children with early relapse and CIMP high methylation profile at diagnosis. However, CIMP status did not have prognostic relevance for the late relapse patients. The reason for this could only be speculated upon but clonal evolution/selection might have contributed, where early relapses might represent leukemias with characteristics more similar to the clone dominating at diagnosis (Hogan et al., 2011). Still we do not know why CIMP status lacks prognostic information for risk of relapse at diagnosis in BCP-ALL but is of importance in predicting OS after

relapse. The prognostic relevance in predicting survival after relapse could not be explained by whether a patient was stem cell transplanted (SCT) or not, since the proportion of SCT-treated individuals after relapse were similar in both CIMP groups. The observation that our CIMP profile adds prognostic information among early relapse BCP-ALL patients and high risk (MRD>0.1% at treatment day 29) T-ALL patients, making our panel a potentially useful stratifier for high risk cases in both T-ALL and BCP-ALL. The panel adds information in areas where there is a lack of stratifying markers.

BCP-ALL in general and this cohort in particular have been well characterized from a DNA methylation perspective, and methylation pattern has been shown to reflect both cytogenetic aberrations and immunophenotype (Lee et al., 2015; Nordlund et al., 2015). Upfront methylation array analysis at diagnosis could be implemented as a complementary method for diagnosis of ALL and used as a support for prediction of the cytogenetic subtypes along with adding epigenetic phenotype information (CIMP status) with prognostic relevance at diagnosis in T-ALL and at relapse in BCP-ALL. Along with genetic analysis and cytogenetic classification, epigenetic classification could add information about biological characteristics of the leukemia and serve as a potential target in the search for new treatment approaches.

The biology behind CIMP subgroups in ALL remains to be determined, i.e. if mutations or genetic aberrations in DNA methylation associated genes or oncogenes can be associated with CIMP subgroups. Several mechanisms have been shown to influence the epigenetic landscape of cells and recurrent mutations in genes involved in epigenetic regulation have been reported (Mullighan et al., 2011; Zhang et al., 2012). Mutations in the CREBBP gene, which encodes the transcriptional coactivator and histone acetyltransferase CREB-binding

protein, have been associated with relapse (Mullighan et al., 2011). Another more general mechanism is accumulation of hypermethylation in malignant cells by proliferation, a phenomenon not only seen in leukemias but also in solid tumors (Lin and Wagner, 2015). To further elucidate this question integrated genetic and epigenetic analysis in combination with functional studies are needed.

To conclude, this study together with our recently published study on T-ALL shows that CIMP classification has the potential to separate high risk pediatric ALL patients and could confer important information in treatment decision-making.

Conflict of interest statement

The authors declare no competing interests.

Authors' contributions

Conceived and designed the study: MB, ZH, MHu, SD. Collection of data: JN, GL, ACS, MHe, NOPHO collaborators. Performed the experiments: JN, ZH. Analyzed the data: MB, ZH, ML, PL, MHu, EF, SD, JN, GL. Wrote the paper: MB, SD wrote the first draft and all authors contributed to the final manuscript.

Acknowledgements

Supported by grants from the Swedish Cancer Society, the Swedish Childhood Cancer Foundation, the Medical Faculty of Umeå University, Lion's Cancer Research Foundation, Umeå, Umeå Paediatric Clinic Research Foundation, the Kempe foundations, Magnus Bergvalls foundation and Uppsala-Umeå Comprehensive Cancer Consortium. Financial support was provided through regional agreement between Umeå University and Västerbotten County Council on cooperation in the field of Medicine, Odontology and Health.

References

- Borssen, M., Haider, Z., Landfors, M., Noren-Nystrom, U., Schmiegelow, K., Asberg, A.E., Kanerva, J., Madsen, H.O., Marquart, H., Heyman, M., *et al.* (2016). DNA Methylation Adds Prognostic Value to Minimal Residual Disease Status in Pediatric T-Cell Acute Lymphoblastic Leukemia. *Pediatric blood & cancer* *63*, 1185-1192.
- Borssén, M., Palmqvist, L., Karrman, K., Abrahamsson, J., Behrendtz, M., Heldrup, J., Forestier, E., Roos, G., and Degerman, S. (2013). Promoter DNA methylation pattern identifies prognostic subgroups in childhood T-cell acute lymphoblastic leukemia. *PLoS One Jun 6;8*, e65373.
- Guillamot, M., Cimmino, L., and Aifantis, I. (2016). The Impact of DNA Methylation in Hematopoietic Malignancies. *Trends Cancer* *2*, 70-83.
- Hogan, L.E., Meyer, J.A., Yang, J., Wang, J., Wong, N., Yang, W., Condos, G., Hunger, S.P., Raetz, E., Saffery, R., *et al.* (2011). Integrated genomic analysis of relapsed childhood acute lymphoblastic leukemia reveals therapeutic strategies. *Blood* *118*, 5218-5226.
- Lee, S.T., Muench, M.O., Fomin, M.E., Xiao, J., Zhou, M., de Smith, A., Martin-Subero, J.I., Heath, S., Houseman, E.A., Roy, R., *et al.* (2015). Epigenetic remodeling in B-cell acute lymphoblastic leukemia occurs in two tracks and employs embryonic stem cell-like signatures. *Nucleic Acids Res* *43*, 2590-2602.
- Lin, Q., and Wagner, W. (2015). Epigenetic Aging Signatures Are Coherently Modified in Cancer. *PLoS Genet* *11*, e1005334.
- Mullighan, C.G., Zhang, J., Kasper, L.H., Lerach, S., Payne-Turner, D., Phillips, L.A., Heatley, S.L., Holmfeldt, L., Collins-Underwood, J.R., Ma, J., *et al.* (2011). CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature* *471*, 235-239.

- Nguyen, K., Devidas, M., Cheng, S.C., La, M., Raetz, E.A., Carroll, W.L., Winick, N.J., Hunger, S.P., Gaynon, P.S., Loh, M.L., *et al.* (2008). Factors influencing survival after relapse from acute lymphoblastic leukemia: a Children's Oncology Group study. *Leukemia* 22, 2142-2150.
- Nordlund, J., Backlin, C.L., Wahlberg, P., Busche, S., Berglund, E.C., Eloranta, M.L., Flaegstad, T., Forestier, E., Frost, B.M., Harila-Saari, A., *et al.* (2013). Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. *Genome biology* 14, r105.
- Nordlund, J., Backlin, C.L., Zachariadis, V., Cavelier, L., Dahlberg, J., Ofverholm, I., Barbany, G., Nordgren, A., Overnas, E., Abrahamsson, J., *et al.* (2015). DNA methylation-based subtype prediction for pediatric acute lymphoblastic leukemia. *Clin Epigenetics* 7, 11.
- Oskarsson, T., Soderhall, S., Arvidson, J., Forestier, E., Montgomery, S., Bottai, M., Lausen, B., Carlsen, N., Hellebostad, M., Lahteenmaki, P., *et al.* (2016). Relapsed childhood acute lymphoblastic leukemia in the Nordic countries: prognostic factors, treatment and outcome. *Haematologica* 101, 68-76.
- Sandoval, J., Heyn, H., Mendez-Gonzalez, J., Gomez, A., Moran, S., Baiget, M., Melo, M., Badell, I., Nomdedeu, J.F., and Esteller, M. (2013). Genome-wide DNA methylation profiling predicts relapse in childhood B-cell acute lymphoblastic leukaemia. *Br. J. Haematol.* 160, 406-409.
- Schmiegelow, K., Forestier, E., Hellebostad, M., Heyman, M., Kristinsson, J., Soderhall, S., and Taskinen, M. (2010). Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. *Leukemia* 24, 345-354.
- Zhang, J., Ding, L., Holmfeldt, L., Wu, G., Heatley, S.L., Payne-Turner, D., Easton, J., Chen, X., Wang, J., Rusch, M., *et al.* (2012). The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* 481, 157-163.

Table 1: Clinical characteristics of 137 relapsed pediatric BCP-ALL which were CIMP classified at diagnosis.

<i>Relapsed BCP-ALL samples</i>	CIMP low n=42	CIMP high n=95	<i>p-value</i>
Sex male/female	28/14	51/44	<i>n.s</i>
Age at diagnosis (months)	48	70	<i>n.s</i>
(median, range)	(16-185)	(12-211)	
WBC x 10⁹/l at diagnosis	20,7	18,6	<i>n.s</i>
(median, range)	(2,2-269)	(1,3-274)	
Cytogenetic group at diagnosis			<i>0.003</i>
dic(9;20)	0	5	
t(1;19)	3	1	
t(12;21)	4	32	
t(9;22)	6	3	
Chr <45	1	1	
11q23	1	3	
HeH	14	18	
iAMP21	1	3	
Non-recurrent	4	20	
No result	8	9	
Initial therapy			<i>n.s</i>
SR/IR	23	64	
HR	19	31	
BMT after relapse			<i>n.s</i>
Yes	15	43	
No	27	52	
Relapse site*			<i>n.s</i>
Bone marrow isolated	31	50	
Bone marrow and extramedullary	6	24	
Extramedullary	5	20	
Time to relapse (months)	30	35	<i>n.s</i>
(range)	(5-124)	(1-172)	
Very early (VE)	11	12	
Early (E)	16	38	
Late (L)	15	45	
pOS _{5y}	0.33+/-0.08	0.65+/-0.05	<i>0.001</i>

ns=not significant, WBC=white blood cell count, CR1=first complete remission, DCR1=dead in CR1, SMN=second malignancy, SR/IR= standard risk/intermediate risk, HR=high risk, BMT= bone marrow transplanted

**One patient data missing*

Figure 1

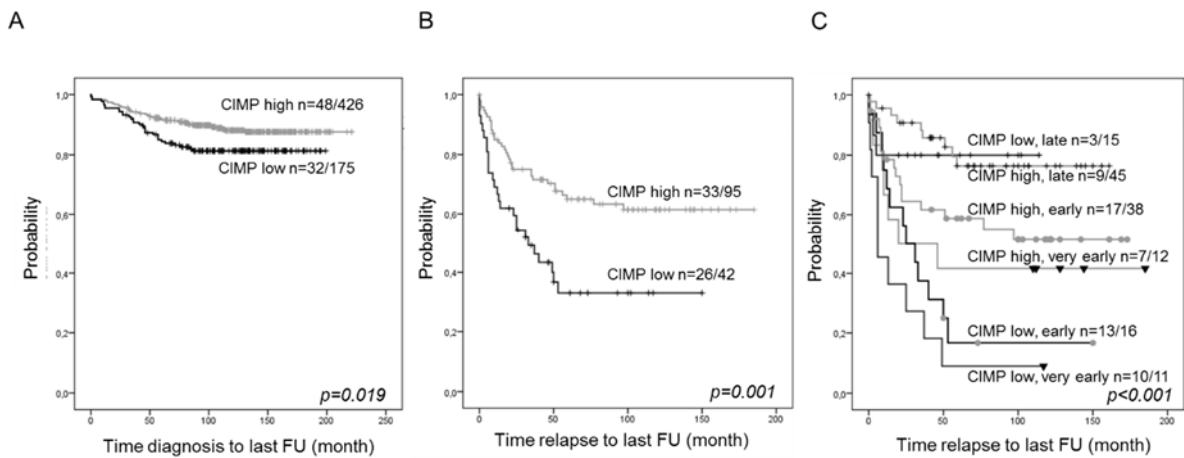


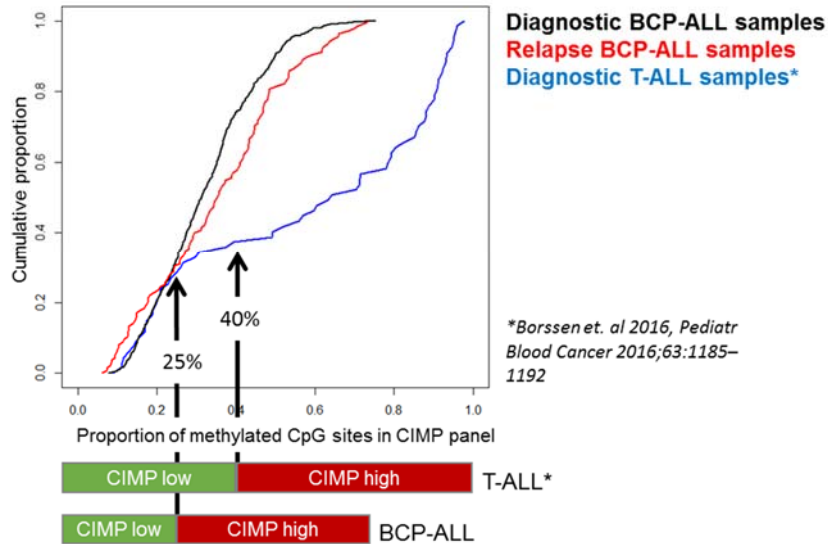
Figure 1: Kaplan-Meier overall survival analysis based on CIMP subgroups in BCP-ALL. Overall survival analysis in A) 601 diagnostic BCP-ALL samples CIMP classified at diagnosis. Follow up time (month) from diagnosis to last follow up. B) 137 relapsed BCP-ALL samples CIMP classified at diagnosis. Follow up time (month) from relapse to last follow up. C) 137 relapsed BCP-ALL samples with combined CIMP status at diagnosis and time to relapse. Time to relapse was classified as: very early (<18 months from primary diagnosis), early (≥ 18 months from diagnosis to ≥ 6 months post treatment), and late (>6 months post treatment). Follow up time (month) from relapse to last follow up.

Supplementary Table S1: Clinical characteristics of 601 diagnostic pediatric BCP-ALL.

<i>Diagnostic BCP-ALL samples</i>	CIMP low n=175	CIMP high n=426	<i>p-value</i>
Sex male/female	98/77	227/199	<i>n.s</i>
Age at diagnosis (months)	56	60	<i>n.s</i>
(median, range)	(12-211)	(12-226)	
WBC x 10⁹/l at diagnosis	14,1	12,6	<i>n.s</i>
(median, range)	(0,8-986,1)	(0,5-336)	
Cytogenetic group			
dic(9;20)	2	18	<i><0.001</i>
t(1;19)	19	4	
t(12;21)	17	130	
t(9;22)	12	6	
<45	1	3	
>67	0	2	
11q23	4	10	
HeH	64	111	
iAMP	2	8	
Non-recurrent	23	70	
No result	31	64	
Follow up status			<i>n.s</i>
CR1	124	313	
induction failure	2	4	
resistant disease	3	4	
relapse	42	95	
DCR1	1	7	
SMN	3	3	
pCIR _{5y}	0.20+/-0.02	0.21+/-0.03	<i>n.s</i>
pEFS _{5y}	0.76+/-0.02	0.75+/-0.03	<i>n.s</i>
pOS _{5y}	0.85+/-0.03	0.92+/-0.13	<i>0.019</i>

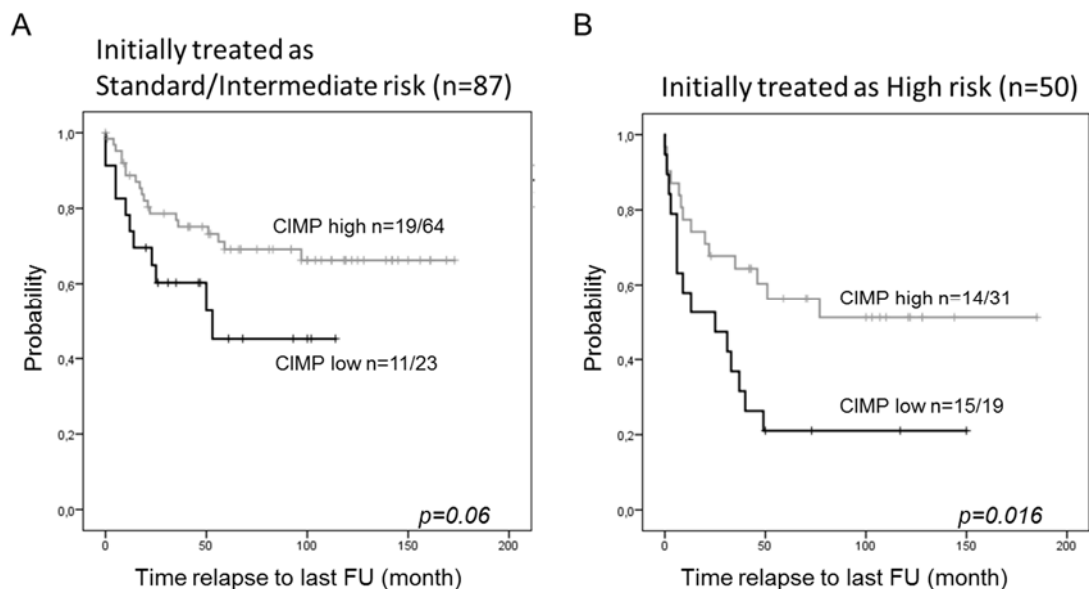
ns=not significant, WBC=white blood cell count, CR1=first complete remission, DCR1=dead in CR1, SMN=second malignancy

Supplementary Figure S1:



Supplementary Figure S1: CIMP classification definition. The CIMP subgroups were defined based on percentage of methylated CpG sites (methylated CpG site defined as a β -value $>0,4$) within the 1293 CpG site CIMP panel. Since approximately 1/3 of the CpG sites in the CIMP panel seemed to be relevant only for T-ALL the CIMP percentage separating CIMP low and CIMP high was adjusted for BCP-ALL. Samples with $\leq 25\%$ methylated CpG sites were classified as CIMP low, and samples $> 25\%$ methylated CpG sites were classified as CIMP high.

Supplementary Figure S2:



Supplementary Figure S2: Kaplan-Meier overall survival analysis of relapse BCP-ALL patients. Overall survival analysis in 137 relapsed BCP-ALL samples CIMP classified at diagnosis. A) 78 patients initially treated according to Standard/Intermediate Risk protocols. B) 50 patients initially treated according to High Risk protocols