



UPPSALA
UNIVERSITET

*Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Medicine 1452*

Resolving the Genomic Complexity of Pediatric Acute Lymphoblastic Leukemia

YANARA MARINCEVIC-ZUNIGA



ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2018

ISSN 1651-6206
ISBN 978-91-513-0302-4
urn:nbn:se:uu:diva-345085

Dissertation presented at Uppsala University to be publicly examined in E10:1307-1309, Navet, SciLifeLab (BMC), Husargatan 3, Uppsala, Friday, 25 May 2018 at 09:00 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English. Faculty examiner: Dr Caroline Heckman (Institute for Molecular Medicine Finland (FIMM), University of Helsinki).

Abstract

Marincevic-Zuniga, Y. 2018. Resolving the Genomic Complexity of Pediatric Acute Lymphoblastic Leukemia. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 1452. 47 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-0302-4.

Acute lymphoblastic leukemia (ALL) is the most common pediatric cancer in the Nordic countries. Structural chromosomal rearrangements are a hallmark of ALL and represent key markers for diagnosis, risk stratification and prognosis. Nevertheless, a substantial proportion of ALL cases (~25%) lack known risk-stratifying markers and are commonly referred to as the B-other subgroup. Improved delineation of structural alterations within this subgroup could provide additional information for diagnosis, prognosis and treatment decisions. Therefore, the aim of this thesis was to decipher the genetic alterations in pediatric ALL, focusing on patients in the B-other subgroup that lack known risk-stratifying markers, and to gain further understanding of the prognostic relevance of aberrant chromosomal changes in ALL.

This thesis comprises four studies. **In study I** we identified a novel and recurrent fusion gene (*PAX5-ESRRB*) in four B-other patients using a combination of RNA-sequencing and copy number analysis. These patients displayed a distinct gene expression and DNA-methylation pattern that differed from other subtypes of ALL. **In study II** we further explored the fusion gene landscape in ALL by applying RNA-sequencing to 134 patient samples assigned to different subtypes, including the B-other subgroup. We detected several novel and recurrent fusion gene families in approximately 80% of the B-other patients of which several were associated with distinct DNA methylation and gene expression profiles. Following on from study II, **in study III** we utilized subtype-specific DNA methylation patterns to design DNA methylation-based classifiers to screen for subtype membership in ~1100 ALL samples including a large group of B-other samples (25%). Re-classification of B-other samples into a new subtype using DNA methylation as the sole marker for subtype classification was validated by RNA-sequencing, which identified previously unknown fusion genes. **In study IV**, “linked-read” whole genome sequencing was applied to 13 ALL samples for in-depth analysis of chromosomal rearrangements. We detected all known pathogenic variants with this technique and also identified previously unknown structural aberrations at a resolution beyond that obtained by traditional karyotyping.

Together, these studies provide novel insights into the structural variation present in ALL and their potential clinical relevance, which may contribute to improved treatment stratification and risk-evaluation of children diagnosed with ALL in the future.

Keywords: Acute lymphoblastic leukemia (ALL), 450k array, DNA methylation, RNA-sequencing, linked-read WGS, B-other, fusion gene, chromosomal translocation

Yanara Marincevic-Zuniga, Department of Medical Sciences, Molecular Medicine, Akademiska sjukhuset, Uppsala University, SE-75185 Uppsala, Sweden.

© Yanara Marincevic-Zuniga 2018

ISSN 1651-6206

ISBN 978-91-513-0302-4

urn:nbn:se:uu:diva-345085 (<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-345085>)

*Dedicado a mi familia
y a la memoria de mi mamá*

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

- I **Marincevic-Zuniga Y.***, Zachariadis V.*, Cavelier L., Castor A., Barbany G., Forestier E., Fogelstrand L., Heyman M., Abrahamsson J., Lönnerholm G., Nordgren A., Syvänen A-C., Nordlund J. (2016) *PAX5-ESRRB* is a recurrent fusion gene in B-cell precursor pediatric acute lymphoblastic leukemia. *Hematologica*, 101(1):e20–3.
- II **Marincevic-Zuniga Y.**, Dahlberg J., Nilsson S., Raine A., Nystedt S., Lindqvist C. M., Berglund C. E., Abrahamsson J., Cavelier L., Forestier E., Heyman M., Lönnerholm G., Nordlund J., Syvänen A-C. (2017) Transcriptome sequencing in pediatric acute lymphoblastic leukemia identifies fusion genes associated with distinct DNA methylation profiles. *Journal of Hematology & Oncology*, 10(1):148.
- III **Marincevic-Zuniga Y.**, Nystedt S., Nilsson S., Almlöf J., Lilljebjörn H., Fioretos T., Flaegstad T., Forestier E., Heyman M., Kanerva J., Schmiegelow K., Lönnerholm G., Nordlund J. DNA methylation classification in combination with RNA-sequencing for subtype discovery in pediatric B-cell precursor acute lymphoblastic leukemia. *Manuscript*.
- IV Nordlund J., **Marincevic-Zuniga Y.**, Cavelier L., Raine A., Martin T., Lundmark A., Abrahamsson J., Norén-Nyström U., Lönnerholm G., Syvänen A-C. Refined detection and phasing of structural aberrations in pediatric acute lymphoblastic leukemia with linked-read whole genome sequencing. *Manuscript*.

*Equally contributing authors.

Reprints were made with permission from the respective publishers.

Contents

Introduction.....	9
Acute lymphoblastic leukemia.....	9
Cytogenetic subtypes of ALL.....	10
BCP-ALL.....	11
T-ALL.....	13
Risk-associated markers and clinical management in ALL.....	13
ALL in the “omics”-era and discovery of new subtypes.....	15
Genomic characterization in ALL.....	15
Characterization of the DNA methylome in ALL.....	19
Present investigations.....	22
Thesis aims.....	22
Materials and methods.....	22
Patient samples and clinical data.....	22
DNA methylation analysis.....	23
RNA-sequencing and fusion gene detection.....	23
Gene expression profiling.....	24
Copy number analysis.....	24
Linked-read whole genome sequencing.....	25
Results.....	26
Identification of the <i>PAX5-ESRRB</i> fusion gene (Study I).....	26
Fusion genes with distinct DNA methylation profiles (Study II).....	27
DNA methylation classifiers for novel ALL subtypes (Study III).....	28
Digital karyotyping by linked-read WGS (Study IV).....	29
Summarized discussion.....	31
Concluding remarks.....	36
Acknowledgements.....	37
References.....	40

Abbreviations

ALL	Acute lymphoblastic leukemia
BCP-ALL	B-cell precursor acute lymphoblastic leukemia
chr	Chromosome
CIMP	CpG island methylator phenotype
CNA	Copy number alteration
CpG	CG dinucleotide
dic(9;20)	Dicentric rearrangement between chromosomes 9 and 20
DMC	Differentially methylated CpG
DNMT	DNA methyltransferase
FISH	Fluorescence in situ hybridization
FPKM	Fragments per kilobase per million mapped reads
HeH	High hyperdiploid
HMW DNA	High-molecular weight DNA
iAMP21	Intrachromosomal amplification of chromosome 21
Ig	Immunoglobulin
MRD	Minimal residual disease
NGS	Next-generation sequencing
NOPHO	Nordic Society of Pediatric Hematology and Oncology
PCA	Principal component analysis
RACE	Rapid amplification of cDNA ends
RIN	RNA integrity number
RT-PCR	Reverse transcription polymerase chain reaction
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
T-ALL	T-cell lineage acute lymphoblastic leukemia
TAPS	Tumor Aberration Prediction Suite
TCR	T-cell receptor
WGS	Whole genome sequencing

Introduction

The hematopoietic stem cell is the founder unit to all the cells in the blood system. These cells undergo a series of hierarchical differentiation steps during hematopoiesis, giving rise to a wide plethora of mature blood cells. Leukemia is a heterogeneous malignant disorder that arises upon the acquisition of genetic alterations in immature blood cells during different stages of the hematopoietic hierarchy. These changes perturb multiple cellular pathways, ultimately leading to abnormal cellular expansion of leukemic cells. Genomic profiling to uncover the pathogenic alterations underlying leukemic cell development is essential in order to understand which of the molecular features that are of clinical relevance. This thesis focuses on using advanced technologies for identifying factors that may have a functional role in the pathogenesis of a specific sub-class of lymphoid leukemia.

Acute lymphoblastic leukemia

Acute lymphoblastic leukemia (ALL) is a disease of the white blood cells that arises from the early B-cell precursor (BCP-ALL) or T-cell (T-ALL) lineages in the bone marrow. The malignant transformation triggers a rapid clonal expansion of the leukemic cells, which out-competes the normal functioning cells, and ultimately disseminate into the peripheral blood system. The term “acute” in ALL reflects the quick progression of this transformation, and if not treated immediately, the outcome is dismal. While relatively uncommon in adults, ALL is the most common cancer amongst children with a peak age of incidence between two and five years of age [1]. Each year, there are approximately 200 children diagnosed with ALL in the Nordic countries.

Although ALL is a malignant disease, the modern advancements in combination therapy and improved diagnostics have been a true success story. Today over 90% of ALL patients reach complete clinical remission after induction therapy [2, 3], which is in contrast to the approximately 20-40% overall survival rates reported during the 1960-1970s [4-6]. However, despite the improved outcome for ALL patients, 15-20% of patients still suffer from relapse or do not respond to treatment, and are the main reasons for ALL-related mortality in children [7]. Relapse in ALL is typically associated with a much poorer outcome, however, recent studies have shown that the 5-

year overall survival rate for relapsed ALL cases has improved over the last two decades reaching over 50% [7]. Together, these marked improvements can be attributed to several landmarks within treatment advancement, such as the discovery of targeted therapies using tyrosine kinase inhibitors, but also several crucial biological findings [4].

In this introduction, I will highlight some of the key genetic findings throughout the modern history of ALL that have increased our knowledge of the underlying biology and further aided the advancement of successful treatment in ALL.

Cytogenetic subtypes of ALL

ALL is genetically a heterogeneous disease that is broadly classified into two immunophenotypic groups based on the cell of origin. ALL originating from the B-cell lineage is the most frequent immunophenotype accounting for almost 85% of the cases, while only 15% of the ALL cases arise from the T-cell lineage. At the molecular level, ALL is characterized by a variety of distinct genetic abnormalities. BCP-ALL is further subdivided into several cytogenetic subtypes that are observed in varying frequencies (**Figure 1**). In general, BCP-ALL can be divided into four different subgroupings: 1) cases with abnormal chromosomal numbers (aneuploidies); 2) cases with translocations giving rise to recurrent and expressed fusion genes; 3) additional recognized chromosomal abnormalities; 4) genetically unclassified cases with or those without any detectable cytogenetic aberration.

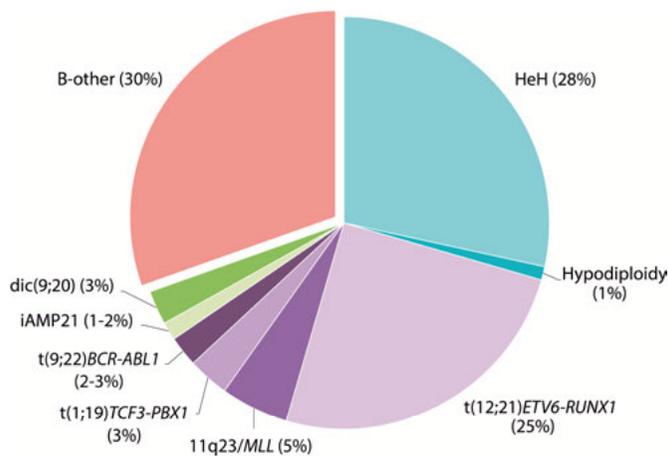


Figure 1. Distribution of risk-stratifying subtypes and B-other in BCP-ALL patients diagnosed in the Nordic countries between 1996 and 2012.

BCP-ALL

Aneuploidies

Numerical chromosomal copy number alterations (aneuploidies) are frequently detected in ALL. In fact, the most common subtype is characterized by the presence of 51-67 chromosomes per leukemic blast cell and is termed high hyperdiploid (HeH). This subtype accounts for approximately 25-30% of the diagnosed pediatric BCP-ALLs. The acquirement of extra chromosomes is a non-random event and may involve any chromosome, but typical patterns of trisomies are seen for chromosomes X, 4, 6, 10, 14, 17, 18, and 21 [8]. Gains of chromosome 21, manifested by one to five extra copies, are frequent in HeH and detected in over 90% of the cases [9]. In contrast to HeH, a much rarer subtype present in approximately 1% of ALL, is defined by <45 chromosomes and termed hypodiploid.

Translocations

Chromosomal rearrangements giving rise to expressed fusion genes are a hallmark of ALL. The most recurrent translocations used for cytogenetic classification are t(12;21), t(1;19), t(9;22) and translocations involving *KMT2A* (referred to as the 11q23/*MLL* subtype throughout this thesis).

The second most common subtype in ALL is defined by the t(12;21) translocation giving rise to the *ETV6-RUNX1* fusion gene, which is detected in approximately 25% of BCP-ALL. *RUNX1* is an essential transcription factor for definitive hematopoiesis, supported by the lack of primary hematopoiesis during embryogenesis in *RUNX1* knock-out mice models [10]. *ETV6* also plays an important role in later stages of hematopoiesis and acts as a transcriptional repressor [11, 12]. The current model for the disease triggering mechanism of t(12;21)*ETV6-RUNX1* implies a two-hit process, where a prenatally acquired fusion gene alone does not have the ability to develop overt ALL, and hence additional mutation(s) are required [13, 14]. Such a two-hit model is supported by several findings such as detection of t(12;21)*ETV6-RUNX1* in neonatal blood spots from children that developed *ETV6-RUNX1*-positive ALL several years after birth, even after a latency period up to 14 years, and a low concordance rate of *ETV6-RUNX1*-positive ALL in monozygotic twins (5%) [15, 16]. Interestingly, studies have shown that t(12;21)*ETV6-RUNX1* occurs frequently during normal fetal development and is observed in 1-5% of healthy newborns, although very few of these children develop ALL, which further supports the low oncogenic potential of this fusion gene [15, 17].

Two subtypes that historically have been associated with dismal prognosis are the t(1;19)*TCF3-PBX1* and the t(9;22)*BCR-ABL1* subtypes. Together, these subtypes represent approximately 6-8% of the BCP-ALL cases. The well described t(9;22) translocation, also known as the Philadelphia chromosome, gives rise to the chimeric protein BCR-ABL1 that functions as a con-

stitutively active tyrosine kinase. Presence of BCR-ABL1 leads to aberrant phosphorylation of various downstream target proteins and thus to the activation of multiple pathways that promote cell survival and growth [18]. The t(1;19) translocation causes the disruption and joining of two important transcription factors *TCF3* and *PBX1*, that are regulators of normal hematopoiesis [19, 20]. The previously poor prognosis associated with t(1;19)*TCF3-PBX1* has markedly improved by contemporary treatment [21]. Similarly, the discovery of tyrosine kinase inhibitors, such as Imatinib, has led to a substantial improvement of the outcome for patients with t(9;22)*BCR-ABL1* and is considered a paradigm for targeted cancer therapies [22].

Rearrangement of the *KMT2A* gene (*MLL*) is the genetic alteration that defines the 11q23/*MLL* subtype, which is present in approximately 5% of the ALL cases. The *KMT2A* gene is a promiscuous gene known to form fusion genes with over 80 gene partners, including the most common partner *AFF1* (*AF4*), *MLLT3* (*AF9*), *MLLT1* (*ENL*) and *MLLT10* (*AF10*) [23]. *KMT2A* encodes a histone methyltransferase that methylates lysine residues of the histone H3, and plays an essential role in regulating gene expression by chromatin remodeling [24]. It has been proposed that the *KMT2A*-related fusion genes are prenatally acquired and are by themselves sufficient for development of overt leukemia, which is supported by the young age of disease onset (typically <1 year) and a close to 100% concordance rate for BCP-ALL between monozygotic twins [25].

Structural abnormalities of specific chromosomes

In contrast to the aforementioned subtypes, the intrachromosomal amplification of chromosome 21 (iAMP21) and dicentric rearrangements involving chromosome 9 and 20 (dic(9;20)), represent more recently identified subtypes that are now recognized as recurrent and non-random events in ALL. iAMP21 is characterized by the amplification of a 5.1 Mb region on chromosome 21q involving the *RUNX1* gene [26]. This subtype, constituting roughly 2% of the cases, was initially discovered by routine FISH analysis of the t(12;21)*ETV6-RUNX1* subtype, revealing an absence of t(12;21) and instead several additional copies of *RUNX1* [27]. iAMP21 is thus defined as having a total of 5 or more copies of *RUNX1* per leukemic cell by FISH analysis [28]. The dic(9;20) subtype encompasses approximately 3% of the cases and is characterized by a dicentric chromosome resulting from the joining of the chr9p and chr20q with the retention of both centromeres from the respective chromosomes [29, 30]. In some cases, the breakpoint in chromosome 9 has been located within the *PAX5* gene, which results in expressed fusion genes together with a handful of gene partners on chromosome 20, such as *NOL4L*, *ASXLI*, and *KIF3B*, however not all cases result in an expressed fusion gene [31].

Genetically unclassified BCP-ALL (B-other)

Despite the well-defined cytogenetic classification in BCP-ALL, up to 30% of the BCP-ALL cases remain unclassified, and the underlying genetic alterations are either unknown or non-recurrent. This group of ALL patients is collectively referred to as B-other. With the aid of several high-throughput techniques for genome-wide profiling, such as whole genome sequencing (WGS), transcriptome sequencing and array-based copy number analysis, several alterations have recently been uncovered in B-other and have reshaped our notion about this subgroup [32]. In general these findings comprise key alterations defined by distinct gene expression patterns, novel fusion genes and recurrent micro-deletions in genes affecting key pathways, as will be outlined later in this thesis.

T-ALL

Opposite to the well-defined cytogenetic classification of BCP-ALL, T-ALL is characterized by few recurrent chromosomal abnormalities and a more heterogeneous genetic landscape [33]. The most frequent alterations in T-ALL are activating *NOTCH1* mutations and 9p deletions resulting in inactivation of the *CDKN2A/B* gene [34]. Together, these two types of alterations occur in more than 50% of the T-ALL cases. Chromosomal rearrangements giving rise to expressed fusion genes are also frequent in T-ALL, commonly involving the T-cell receptor gene [35]. Although the molecular basis of T-ALL has radically been redefined in recent years, most of the described genetic alterations have had a low prognostic impact [33].

Risk-associated markers and clinical management in ALL

As the multiple ALL subtypes are associated with different clinical outcome the discovery of recurrent genetic abnormalities constitute yet another landmark in the clinical management of ALL (**Figure 2**). Thus, the presence of genetic alterations is an important factor for risk-based stratification and directed treatments in ALL [36]. Genetic markers for “standard” risk comprise the two largest subtypes, HeH and t(12;21)*ETV6-RUNX1*, that are associated with favorable outcome. Three markers that confer intermediate risk are the (1;19)*TCF3-PBX*, dic(9;20) and the iAMP21 subtypes. Subtypes that are considered high-risk genetic markers and are associated with inferior prognosis include the less frequent subtypes t(9;22)*BCR-ABL1*, 11q23/*MLL* and hypodiploid subtypes. These subtypes are associated with higher rates of relapse, therapy resistance or death as compared to other subtypes [5].

Apart from molecular-based stratification, clinical parameters are essential for risk-associated treatment stratification of ALL [37]. An age less than one year or more than 10 years at diagnosis are considered high-risk features, which can be attributed to the age distribution amongst the ALL subtypes [5]. For instance, the 11q23/*MLL* subtype which confers poor prognosis is frequently observed in cases younger than 1 year whereas iAMP21 cases tend to be older, with a median age of 9. Other features such as high white blood cell count and involvement of the central nervous system are also considered to be high-risk features. Additionally, establishment of routine monitoring of minimal residual disease (MRD) to assess overall therapy response has proven to be a highly reliable prognostic indicator [38, 39]. Evaluation of MRD by flow cytometry or real-time quantitative PCR to monitor treatment response at a submicroscopic level enables detection of residual leukemic cells during or after induction therapy. Detection of MRD can be performed by screening of leukemia-specific immunophenotypes or clonal rearrangements of the T-cell receptor (TCR) or immunoglobulin (Ig) [40]. Additionally, studies have shown that monitoring of treatment response is also possible by screening of risk-associated markers, such as the *ETV6-RUNX1* or *TCF3-PBX1* fusion genes, which complements current MRD methods [41, 42]. These findings highlight the importance of establishing the underlying alteration in the leukemic cells at diagnosis, which may become fundamental markers for MRD.

Another important aspect for the clinical management of ALL has been the establishment of several international collaborative working groups, including the Nordic Society of Pediatric Hematology and Oncology group (NOPHO) across the Nordic countries that aim for unifying diagnostics and common treatment protocols and strategies [43]. As a result, intensification of treatment protocols and addition of new drugs, such as tyrosine kinase inhibitors for treatment of t(9;22)*BCR-ABL1* positive ALL, have also had an impact on clinical management of ALL.

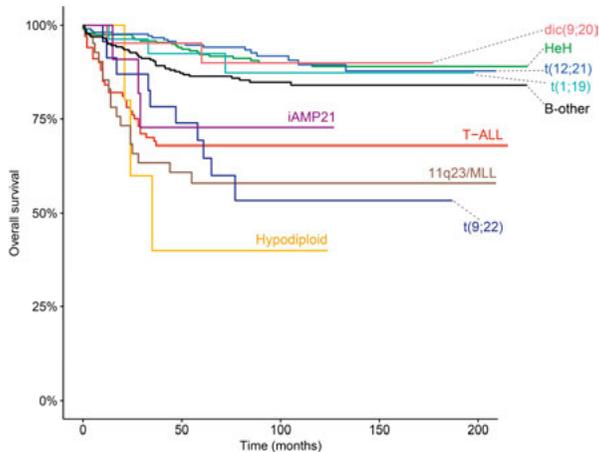


Figure 2. Kaplan-Meier curves showing the overall survival of >900 samples from the different subtypes of ALL treated according to the NOPHO protocols between 1996 and 2011.

ALL in the “omics”-era and discovery of new subtypes

The low-resolution methods such as karyotyping and FISH that are routinely used for clinical diagnostics in ALL are limited in their ability to detect novel or discreet changes of the genome, including changes at base pair resolution. Microarray-based technologies and next generation sequencing (NGS) have greatly increased our ability to identify novel submicroscopic alterations in a genome-wide context. As a consequence of such technological advancements, new molecularly distinct subgroups have been described in ALL, predominantly in the heterogeneous B-other group [32, 44, 45]. For this reason, the classification system of BCP-ALL is presently changing as several new and recurrent putative subgroups have emerged that are transforming the conventional view of subtype classification in ALL.

Genomic characterization in ALL

Array-based technologies

The advent of microarray-based technologies provided a leap from low- to high-throughput analysis that transformed the settings for molecular biology studies [46]. Microarray technologies comprise numerous applications including genome-wide profiling of single nucleotide polymorphisms (SNP), copy number alterations (CNA), gene expression patterns, and transcription factor binding site profiling [47]. Although microarrays are limited in their genomic coverage compared to NGS technologies, they are particularly advantageous for large-scale studies involving numerous samples and provide

a straightforward approach for analyzing hundreds of thousands of genomic sites at a reasonable cost.

SNP and gene expression arrays in ALL

Since 2007, SNP arrays have been used to uncover a spectrum of secondary CNAs (DNA deletions or gains), that are important cooperating lesions for disease pathogenesis, especially in BCP-ALL [48-50]. Transcription factors that are essential for hematopoiesis, including *PAX5*, *IKZF1*, *ETV6* and *EBF1* are among the most frequently altered genes in addition to the tumor suppressors and cell cycle regulators *CDKN2A/B*, *PTEN* and *RBI* [51-53]. In fact, the crucial B-cell lineage regulating transcription factor *PAX5* is recurrently mutated, deleted or rearranged in over one-third of BCP-ALL cases [52]. The type and frequency of *PAX5* alterations vary between the ALL subtypes, and *PAX5* deletions occur in up to 36% of B-other cases [54]. Deletions of the lymphoid transcription factor *IKZF1* are reported in 15% of BCP-ALL cases, but is more common in the t(9;22)*BCR-ABL1* subtype where *IKZF1* deletions are found in >60% of the patients [55, 56]. Unlike the gross chromosomal rearrangements and aneuploidies that are usually initiating lesions acquired early during the leukemic transformation, the submicroscopic DNA changes represent subsequently acquired alterations [51, 57-60]. Some of the CNAs, such as deletions of *PAX5*, *IKZF1* or *CDKN2A/B*, have been shown to promote the onset of leukemia in mouse models, supporting their role as central cooperative lesions for leukemogenesis [61-63].

Before the advent of NGS, gene expression arrays represented a unique platform to study differential expression of a defined set of target genes. Early microarray-based expression studies in ALL could establish that the different cytogenetic subtypes of ALL have unique gene expression signatures, which led to a considerable better molecular understanding of the subtypes of ALL [64-66]. In 2009, a new subtype termed *BCR-ABL1*-like (also known as Ph-like) was described in the B-other group with the aid of gene expression microarrays [56, 67]. This group of patients was characterized by an expression profile similar to that of the t(9;22)*BCR-ABL1* subtype, yet they lack the canonical *BCR-ABL1* fusion gene detectable by gene expression arrays. Similarly to the t(9;22)*BCR-ABL1*, the *BCR-ABL1*-like subtype also harbors high frequencies of *IKZF1* deletions, and are associated with a poor outcome [56, 67, 68]. Today, *BCR-ABL1*-like is recognized as a new provisional entity of leukemia that comprises 10-15% of the ALL patients and is classified as a high-risk marker [69].

RNA-sequencing

The advancements in NGS techniques, coupled with the reduction in sequencing costs, have enabled new opportunities for comprehensive analysis of the transcriptome. Understanding the abundance and structure of RNA is

essential for interpreting molecular processes involved in maintaining cell functions and for disease pathogenesis. For this purpose, RNA-sequencing offers unbiased determination of gene expression levels, expressed variants, alternative splicing and detection of fusion genes. Multiple protocols for RNA-sequencing exist, typically involving enrichment or depletion of “total” RNA for specific RNA molecules, i.e. polyadenylated RNA and ribosomal RNA, respectively, followed by conversion of the remaining RNA of interest to cDNA prior to sequencing library preparation [70].

RNA-sequencing in ALL

RNA-sequencing is a powerful tool for detection of known and novel fusion genes. In ALL, RNA-sequencing is particularly advantageous due to the high frequency of expressed fusion genes as a consequence of recurrent translocations. In light of this, recent studies taking advantage of RNA-sequencing in BCP-ALL has uncovered the presence of several key fusion genes in the B-other subtype that are associated with clinical outcome and are thus of potential clinical relevance. The most notable findings have been the recurrent fusion gene rearrangements involving *DUX4*, *ZNF384* and *MEF2D* [32, 44, 45].

The *DUX4*-rearranged subtype is characterized by a fusion gene resulting from the insertion of the *DUX4* gene into the *IGH* locus on chr14 and recurrent *ERG* deletions [71-74]. Before the discovery of expressed *DUX4*-related fusion genes, prior studies had reported recurrent deletions of the *ERG* gene in approximately 5% of ALL patients and this putative subgroup of ALL patients was thus referred to as “*ERG*-deleted” [75-77]. Today, it has been established that *ERG* deletions are restricted to the *DUX4*-rearranged subtype. The *DUX4*-rearranged subtype has so far been reported in up to 7% of the BCP-ALL cases and appears to have a favorable prognosis under current clinical protocols in multiple studies [44, 71, 72, 78].

Rearrangements involving *ZNF384* and *MEF2D* are less frequent, with estimated frequencies of 1-6% and 1-4% in BCP-ALL, respectively [32, 45]. Both *ZNF384* and *MEF2D* rearrangements have been reported to involve multiple gene fusion partners [72, 73, 79-82]. Patients with *ZNF384* have been reported to have varying clinical outcome depending on its 3' fusion partner (denoted as intermediate prognosis), whilst cases harboring *MEF2D* rearrangements are associated with poor clinical outcome [45, 79]. Due to the small sample sizes analyzed to date, larger studies will be required in order to fully establish the prognostic relevance of these two novel subtypes.

The *BCR-ABL1*-like subtype, as described above, is a heterogeneous subgroup involving different genetic alterations that activate kinase and cytokine receptor signaling and is associated with a dismal outcome [68]. Apart from its aforementioned molecular resemblance to the t(9;22)*BCR-ABL1* subtype, RNA-sequencing has revealed a plethora of expressed fusion gene within the *BCR-ABL1*-like subtype, involving tyrosine kinase receptor genes

such as *PDGFRB*, *ABL1*, *ABL2* or *CSF1R*, and JAK/STAT activating rearrangements involving *CRLF2*, *JAK2* and *EPOR* [83-86]. A subset of the *BCR-ABL1*-like patients lack these rearrangements, but harbor other less frequent JAK/STAT activating lesions or rare kinase related fusion genes [68]. Gene expression profiling is helpful for discerning B-other patients with putative *BCR-ABL1*-like expression patterns and a common approach relies on microarray-based gene expression signatures from a predefined probe set [87, 88]. For cases where gene-expression microarray data is not available, RNA-sequencing has also been successfully applied for identification of these patients based on gene expression signatures that are similar to t(9;22)*BCR-ABL1* patients [71, 73]. Importantly, as a consequence of lesions involving tyrosine kinases, the use of tyrosine kinase inhibitors could improve outcome for the patients in this subgroup [84, 89].

Whole genome sequencing

Whole genome sequencing (WGS) is a powerful technique for high-resolution profiling of genetic variation and has revealed numerous somatic single nucleotide variants (SNV) and mutational signatures that are important contributors for cancer progression. By genome-wide profiling, WGS enables de-novo detection of somatically acquired mutations in non-coding regions including introns and regulatory regions such as enhancers in an unbiased manner [90]. Despite its strength, WGS is limited in identifying the full spectrum of structural variants such as translocations, insertions and deletions [90]. Furthermore, complex genomic regions such as repetitive sequences are still challenging to fully resolve in detail by standard short-read WGS data.

Recently, a new technology termed “linked-read” WGS (10x Genomics) was introduced [91]. Linked-read WGS enables improved resolution across complex genomic regions and detection of structural variants, owing to its capacity to generate the long-range phased genomic regions that is beyond the capacity of standard short-read sequencing. The principle of long-read sequencing is to utilize massive barcoding strategies and microfluidics for partitioning single DNA molecules to generate barcoded sequencing libraries that are subjected to standard short-read WGS. After WGS the sequencing reads are aligned and assembled into synthetic “long-read” genome sequences with the aid of unique barcodes for the WGS reads generated from the same DNA molecule. This enables computational reconstruction of long haplotype blocks and thus phasing of genetic variants including, chromosomal rearrangements, single nucleotide variants, and small insertion and deletions (**Figure 3**). Linked-read WGS offers a promising approach for assignment of complex variants, such as SNVs, insertion-deletions, chromosome breaks, gene fusions and other rearrangements to individual chromosomes in germline and cancer genomes [91-93]

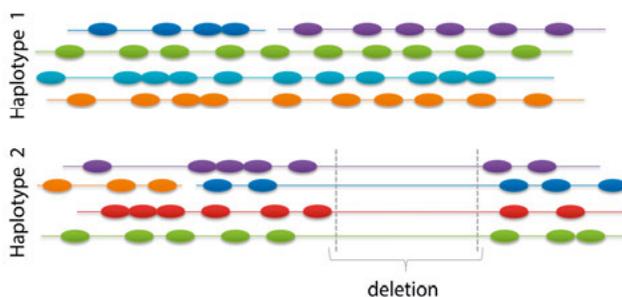


Figure 3. Reconstruction of phased blocks by linked-read whole genome sequencing. The colored circles represent short-read sequences that are linked together based on their common barcodes to generate unambiguous haplotype blocks for the homologous chromosome pairs. A deletion on haplotype 2 is illustrated by the absence of short reads aligning to that region.

WGS in ALL

The mutational landscape of ALL has been extensively studied with the aid of high-throughput sequencing such as short-read WGS [51]. Opposite to gene deletions, the frequency of somatic mutations is low in ALL, and the 11q23/*MLL* subtype has been found to have amongst the lowest mutation rates of all cancers [94, 95]. Indeed, a spectrum of mutations has been detected in ALL, including mutations in members of the RAS signaling pathway (*NRAS*, *KRAS*) and hypodiploid-associated *TP53* mutations [96-99]. The most notable findings has been generated by studies of the clonal evolution in relapsed ALL that have identified genes that are enriched at relapse such as activating *NT5C2*, that drives chemotherapy resistance in ALL and is detected in up to 20% of relapsed cases [100, 101].

Characterization of the DNA methylome in ALL

DNA methylation is an epigenetic modification of the DNA that occurs predominantly at CpG dinucleotides (a cytosine followed by a guanine), which is mediated by DNA methyltransferases (DNMTs). The modification involves the addition of a methyl group to cytosine; giving rise to a “5th base” i.e. 5-methylcytosine. DNA methylation is a robust and stable epigenetic mark that can be faithfully analyzed with various techniques, including microarray-based genotyping, sequencing-based approaches and mass-spectrometry [102].

Infinium microarrays provide a straightforward assessment of the DNA methylation status of many thousands of CpG sites in large sample sets, at a relatively low cost. Infinium microarrays for DNA methylation analysis include predetermined CpG sites distributed across the genome, which cover 99% of the genes in the Reference Sequence database [103]. The number of target probes included on the arrays have increased over the years, starting

with the HumanMethylation 27k array (targeting ~27000 CpG sites), followed by the 450k array and the EPIC array (targeting over 850 000 CpG sites) that covers both genes and enhancer regions [104-106]. It is to be noted that DNA methylation arrays target only up to 3% of the 28 million CpG sites present in the human genome and alternative methods such as whole-genome bisulfite sequencing are required for complete genome-wide DNA methylome assessment at single-base pair resolution [107].

In healthy cells, it has been estimated that 60-80% of the 28 million CpG sites present in the human genome can become methylated. Unmethylated CpGs are typically clustered together in about 200 bp long CpG-rich regions (“CpG islands”) and are enriched in gene promoters [108, 109]. DNA methylation plays an important role in orchestrating cell-type specific gene expression, giving each cell type its specific identity. In general, but not always, a high DNA methylation level (hypermethylation) in promoter regions silences gene expression [108].

DNA methylation in ALL

In ALL, the genome-wide DNA methylation landscape is altered and the methylation pattern of the ALL cells are clearly distinct from that of normal healthy blood cells [64, 110, 111]. Changes in DNA methylation in cancer cells are mainly seen as hypermethylated CpG islands compared to normal controls and has been observed across all genetic and immunophenotypic subtypes of ALL [112]. These differentially methylated CpG sites (DMCs) that are shared between the genetic subtypes of ALL suggests a common mechanism for leukemic transformation in ALL. Further dissection of the DNA methylation landscape across the different subtypes of BCP-ALL has identified strong subtype-specific DMCs that are enriched in promoter and enhancer regions [64, 110-112].

In the context of ALL prognosis, it has been shown that promotor-associated DNA methylation holds prognostic information that can predict relapse in T-ALL [113]. In this study, patients were stratified based on a CpG island methylator phenotype (CIMP) profile, where patients with lower DNA methylation levels at diagnosis (CIMP-) were more likely to relapse, compared to patients with high methylation levels (CIMP+) where relapses were less frequent. Recently, it was shown that CIMP classification is also important for relapsed BCP-ALL and correlates with overall survival after relapse [114].

DNA methylation classification in ALL

Today, it is well-acknowledged that each genetic subtype of BCP-ALL is associated with a distinct set of differentially methylated CpG sites that can be used to distinguish between the subtypes [110]. This feature can be utilized to design DNA methylation-based classifiers to predict the cytogenetic subtype of an ALL sample, solely based on DNA methylation status of a

limited set of CpG sites [115]. Classifiers for the most common and established subtypes of BCP-ALL (HeH, t(12;21)*ETV6-RUNX1*, t(1;19)*TCF3-PBX1*, 11q23/*MLL*, t(9;22)*BCR-ABL1*, dic(9;20) and iAMP21) in addition to T-ALL, have been developed [115]. Interestingly, when the DNA methylation classifiers were applied to a large set of previously undefined B-other samples, nearly half of the samples displayed highly similar DNA methylation patterns as samples with a known subtype at diagnosis. Verification of the predicted subtype in B-other cases led to the discovery of several unexpected fusion genes, including a small number of patients with similar DNA methylation profiles as positive t(12;21)*ETV6-RUNX1* cases, however alternative fusion genes involving either *ETV6* or *RUNX1* were detected, instead of the canonical *ETV6-RUNX1*. Recently, additional patients with *ETV6-RUNX1*-like profiles have been reported based on gene expression or immunophenotypic patterns that are similar to positive t(12;21)*ETV6-RUNX1* cases, and are now referred to as the “*ETV6-RUNX1*-like” group [71, 116]. *ETV6-RUNX1*-like cases frequently harbor co-occurring fusion genes and/or deletions involving *ETV6* and *IKZF1* [44, 71, 116]. Very few relapses have been reported in the *ETV6-RUNX1*-like subtype, suggesting a good prognostic outcome for this putative subtype, similarly to the canonical t(12;21)*ETV6-RUNX1* subtype. Larger studies are required to fully discern the prognostic significance of this emerging subtype.

Not restricted to ALL, DNA methylation as a marker for classifying cancers has been successfully applied to a wide spectrum of malignancies to discriminate between different molecular and outcome groups [117-120]. Together, this reflects a potential use of DNA methylation as a future biomarker.

Present investigations

Thesis aims

The overall aim of this thesis was to study genetic alterations in pediatric acute lymphoblastic leukemia (ALL) patients using multiple genome-wide techniques. More specifically the aims were to:

- Decipher the genetic alterations in patients with ALL focusing on patients lacking known risk-stratifying markers (B-other).
- Explore novel genomic strategies for detection of ALL subtypes based on in-depth analyses of structural rearrangements and fusion genes.
- Gain further understanding of the frequency and prognostic relevance of the identified genetic alterations.

Materials and methods

Several methods for genome-wide analysis were applied throughout the work presented in this thesis. The following section highlights the most important methods used in this work and detailed information about each of these different methods is presented in the individual papers I-IV.

Patient samples and clinical data

The samples from patients with pediatric ALL analyzed in this thesis belong to a unique collection of bone marrow and peripheral blood samples obtained from >1000 children diagnosed with ALL during 1996-2012 within the Nordic countries. The samples were collected at collaborating institutes within the Nordic Society of Pediatric Hematology and Oncology (NOPHO). Phenotypic data has been thoroughly recorded since sample collection in the NOPHO patient registry, which enables follow-up analysis the clinical outcome of the patients over an extended period of time.

Primary leukemic cells collected at initial diagnosis were isolated by Ficoll-Paque centrifugation prior to cryopreservation. Extraction of DNA and RNA was performed using column-based extraction methods including a

DNase treatment step of the RNA. High molecular weight DNA (HMW DNA) was extracted from two samples (Paper IV). Isolated material was quantified using fluorescent assays and quality of the RNA was assessed using the RNA integrity number (RIN) algorithm which is based on the ratio of measured ribosomal RNA 28S and 18S.

Clinical outcome of the patients was assessed by calculating the probabilities of event free survival and overall survival using the Kaplan-Meier method. For event-free survival estimation, an event was defined as resistant disease, induction failure, relapse, death in remission or second malignant neoplasm.

DNA methylation analysis

For genome-wide interrogation of the DNA methylation levels of >450,000 CpG sites genome-wide in the leukemic samples, the Human Methylation 450k BeadChip (450k array, Illumina) was utilized. Sodium bisulfite treatment was performed on 250-500 ng of genomic DNA using the EZ DNA Methylation-Gold Kit (Zymo Research). The bisulfite step converts unmethylated cytosines (C) to uracils, which become thymines (T) after PCR, whilst methylated cytosines remain unaffected. By genotyping the C/T polymorphisms using fluorescent probes, the methylation levels across the pre-defined CpG sites included on the 450k array can be determined.

Raw fluorescent signal intensities from the 450k array were detected and quality controlled with the Minfi R package [121]. The level of DNA methylation for each CpG site was transformed into β (beta) values, which corresponds to the intensity signal from the methylated C-nucleotides over the sum of the signals from the C- and T-nucleotides using GenomeStudio (Illumina). The β -values range from 0-1 which corresponds to 0-100% methylation for each CpG site included on the 450k array. After normalization of the β -values using the peak-based correction algorithm [122], probes with a detection p-value >0.01 or with SNPs or indels with minor allele frequencies >5% within the last 3 bp of the 3' end of the primers were excluded from downstream analyses [110].

RNA-sequencing and fusion gene detection

Strand-specific next generation RNA-sequencing libraries were constructed from 100-300ng total RNA following ribosomal RNA depletion. High-quality RNA samples (RIN values >7) were used for library preparation using the ScriptSeq V2 Kit (EpiCentre) or the TruSeq stranded Total RNA kit (Illumina). A subset of the sequencing libraries from RNA samples of lower quality (RIN <7) were prepared using the TruSeq RNA Access kit (Illumina). The libraries were sequenced using Illumina systems (HiSeq2000, 2500, Miseq or NovaSeq) with 50-150bp paired-end chemistry, generating

20-40 million read pairs per sample. In paper III, Cutadapt [123] was used to trim the sequencing reads to 65 bp to obtain a uniform sample set as different sequencing platforms had been used. Sequencing reads were quality controlled using FastQC [124] prior to alignment against the human reference genome (hg19) with the TopHat2 or STAR software [125, 126]. RNA-sequencing data was quality-controlled using either the RNA-SeQC or RSeQC software [127, 128].

Fusion genes were detected using the FusionCatcher software [129]. Targeted screening of several well-established fusion genes in ALL was performed by counting reads supporting fusion transcripts in the TopHat2-aligned data normalized to the total sequencing depth in the sample. A multi-step filtering process of the candidate fusion genes detected by FusionCatcher was applied to reduce the number of false positives. Briefly, these steps included filtering of blacklisted fusion genes based on reference data provided by FusionCatcher followed by removal of fusion genes based on; (i) common mapping of fusion-supporting reads to multiple positions, which is indicative of sequence homology; (ii) highly promiscuous genes where the 5' or the 3'-fusion gene partner maps to several other genes; (iii) few sequencing reads supporting the fusion junction; or, (iv) expression of the fusion gene in normal CD19+ B-cells and CD3+ T-cells from healthy blood donors.

Gene expression profiling

Quantification of gene expression levels from the RNA-sequencing data was performed using Cufflinks (version 2.2.0) or featureCounts (Subread version 1.5.0) [130, 131]. FeatureCounts was used to summarize the counts of aligned reads and normalized using variance stabilizing transformation with DESeq2 [132]. Differentially expressed genes were analyzed within the Cufflinks pipeline by determining the fragments per kilobase per million mapped reads (FPKM) levels. Cuffdiff was subsequently used to detect differentially expressed genes by pair-wise comparison of the expression levels between two ALL subgroups [133]. Selection of differentially expressed genes in paper III was based on a false discovery rate corrected p-value <0.1 and a two-fold difference in mean expression detected between three or more compared ALL groups. Genes with low expression levels (FPKM <5) were excluded from the analysis.

Copy number analysis

Illumina Infinium microarrays were used to screen for copy number alterations. Several Illumina Infinium arrays were utilized, including the 450k array for DNA-methylation analysis, Omni2.5 array, OmniExpressExome array and the Multi-Ethnic Global array.

Notably, the 450k array is based on the same biochemical principles as the regular SNP genotyping arrays, i.e. using two fluorophores to interrogate the loci included on the bead arrays. Thus, the total raw fluorescence intensity from the 450k array can also be used to detect copy number variants in genomic DNA. The raw signal intensities generated from the 450k array were normalized and transformed to “Log R ratio”, which corresponds to the observed intensity signal from the studied sample over the intensity signal obtained from a reference sample (baseline). To determine the reference baseline, 450k data from CD19+ B-cells and CD3+ T-cells from healthy blood donors representing normal diploid genomes, was used. Log R ratios were analyzed using the CopyNumber450kCancer R package [134].

Log R ratios from the high-density SNP arrays were calculated in GenomeStudio (Illumina) to assess copy number status using an in-house panel of diploid samples or with Illumina standards for reference intensities based on HapMap samples. In addition to analyzing Log R ratios for copy number status, genotyping of A/C/T/G polymorphisms for detection of copy neutral allelic imbalances by estimation of the B-allele frequency was performed in GenomeStudio. Copy number alterations were called from the SNP array data using the Tumor Aberration Prediction Suite (TAPS) analysis software [135].

Linked-read whole genome sequencing

Linked-read whole genome sequencing (linked-read WGS) technology for haplotyping (phasing) of genetic variations was evaluated in study IV. For this purpose, the GemCode and/or the Chromium technology from 10x Genomics was applied. This technology uses a combination of microfluidics and molecular barcoding to generate WGS libraries [91]. The resulting linked-read sequencing libraries enable computational reconstruction of haplotype and structural variant information within a single WGS experiment. Briefly, long DNA molecules are partitioned into single oil drops together with gel beads loaded with enzymes, primers and barcodes. As little as 1 ng of genomic DNA is sufficient as starting material, which is equivalent to approximately 300 copies of the human genome. The resulting barcoded DNA fragments are subsequently sheared, pooled, and finally a standard Illumina short-read sequencing library is generated. GemCode libraries were sequenced on an Illumina HiSeq2500 and Chromium libraries were sequenced on the HiSeqX instrument. A customized sequencing protocol was used for the GemCode libraries whereas standard 150bp paired-end chemistry was used for the Chromium libraries. The linked-read data was processed using a DNA analysis pipeline available from 10x Genomics (Long Ranger) that performs sample de-multiplexing, barcode processing, alignment, quality control and phasing. Structural variant calling was performed using Long

Ranger and CNVnator [136]. Identified structural variants by Long Ranger were visualized in the Loupe genome browser (10x Genomics).

Results

Identification of the *PAX5-ESRRB* fusion gene (Study I)

In study I, we describe a novel and recurrent fusion gene using a combination of RNA-sequencing, reverse transcription PCR (RT-PCR) and copy number detection in B-other. These patients were characterized by the expression of *PAX5-ESRRB*, a fusion gene stemming from an unbalanced translocation between chr9 and chr14 (**Figure 4**). This fusion gene remained undetected at diagnosis using traditional karyotyping in each of the four patients. Owing to the unbalanced nature of the translocation, we utilized the Log R ratio from the 450k array data to screen for copy number alterations involving chr9 and chr14 in 664 pediatric BCP-ALL cases. These patients had previously been analyzed using the 450k DNA methylation array, and the frequency of this fusion was estimated to be <1% in BCP-ALL (4/664). The four ALL cases carrying the *PAX5-ESRRB* fusion gene were at clinical remission after follow-up for 30-173 months. At a molecular level, patients carrying the *PAX5-ESRRB* fusion gene showed distinct gene expression patterns characterized by high expression of *ESRRB*, in addition to a predominant hypomethylation of DNA, as compared to other BCP-ALL subtypes.

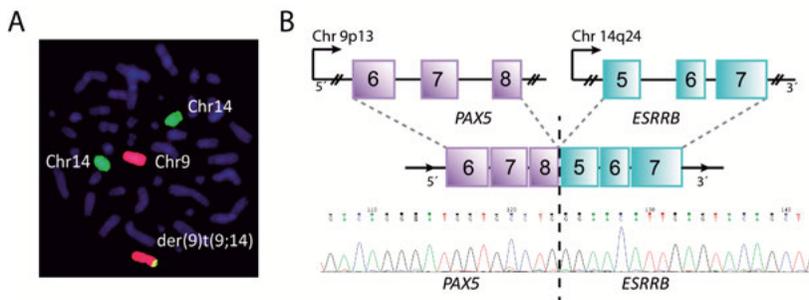


Figure 4. Chromosomal rearrangements giving rise to the expressed *PAX5-ESRRB* fusion gene. (A) FISH analysis revealed the presence of a der(9)t(9;14) rearrangement resulting from a translocation between chromosome 14q to chromosome 9p. (B) Sanger sequencing revealed the expression of an in-frame *PAX5-ESRRB* fusion gene in all cases. Exons involved in the fusion transcript are represented by the numbered squares in the upper panels and color coded to indicate the specific genes involved. Sanger sequencing traces spanning across the fusion junction (vertical dashed line) are illustrated in the lower panel.

Fusion genes with distinct DNA methylation profiles (Study II)

In study II, we further explored the fusion gene landscape of ALL in an unbiased manner by performing total RNA-sequencing on 134 ALL patient samples assigned to various subgroups, including B-other. After rigorous filtering of candidate fusion genes called by FusionCatcher, a total of 64 unique fusion genes were identified in 80 of the 134 (60%) patients included in the study. Subtype-defining fusion genes including *ETV6-RUNX1*, *BCR-ABL1* and *KMT2A*-chimeras, were unambiguously detected in 29 of the 31 BCP-ALL cases with the established subtypes t(12;21), t(9;22) or 11q23/*MLL*.

Although most fusion genes (43/64; 67%) were only detected in a single patient, we detected several previously undetected and recurrent fusion gene families in over 80% (35/42) of the patients in the B-other group. Several of these fusion genes arose due to cryptic or subtle structural rearrangements that were not detectable by traditional diagnostic methods. The most recurrent fusion genes that were detected in 50% of the analyzed B-other transcriptomes, involved the *DUX4*, *ZNF384* or *PAX5* genes. Whereas multiple gene partners were found to fuse with the *ZNF384* and *PAX5* genes, *DUX4* fusion genes involved the *IGH* locus located on chromosome 14 in each case. Furthermore, similar to the well-established subtypes of ALL, the putative subtypes harboring *DUX4* and *ZNF384* displayed distinct subtype-specific DNA methylation and gene expression profiles (**Figure 5**). Using rapid amplification of cDNA ends (RACE), two additional patients harboring *DUX4* and *ZNF384* fusion genes were identified. These patients were discovered based on similar DNA methylation and gene expression profiles as the verified *DUX4*-rearranged and *ZNF384*-rearranged cases, however our screening strategy was not able to detect a fusion gene involving *DUX4* or *ZNF384* in the RNA-sequencing data. Further characterization of the *DUX4*-rearranged and *ZNF384*-rearranged cases revealed a high prevalence of *NRAS* mutation in the *DUX4* subgroup (56%) in addition to frequent *ERG* deletions (78%). Assessment of prognostic relevance for patients with *DUX4* fusion genes confirmed a good prognosis, and an intermediate prognosis in patients harboring the *ZNF384* fusion genes, as recently reported by other studies [72, 73, 79].

Amongst the sequenced B-other cases, three patients previously described to have identical DNA methylation patterns and classified to have the t(12;21)*ETV6-RUNX1* subtype were included in this study. These patients had previously been found to harbor novel, alternative *ETV6* or *RUNX1* fusion genes instead of the canonical *ETV6-RUNX1* fusion gene [115]. In agreement with the DNA methylation profile, these three patients also clustered with t(12;21)*ETV6-RUNX1* samples based on their gene expression profiles.

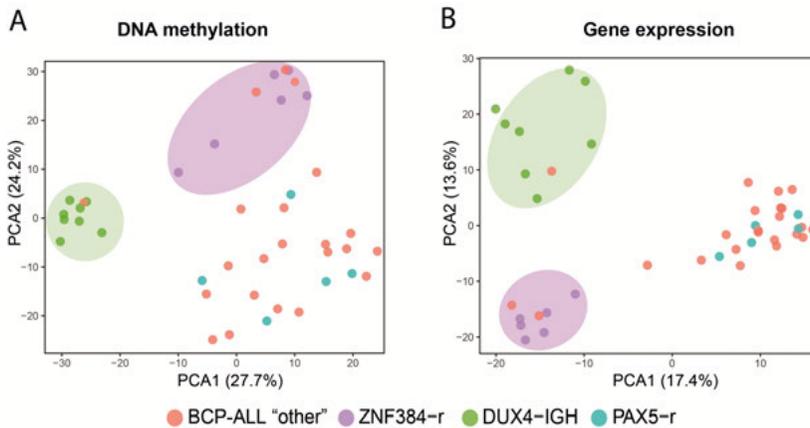


Figure 5. Principle component analysis (PCA) of B-other samples using (A) DNA methylation and (B) gene expression profiles. Samples are color coded based on recurrent fusion gene families as detected by RNA-sequencing comprising the *DUX4-IGH* (green), *ZNF384*-rearranged (purple) and *PAX5*-rearranged (blue) subgroups. B-other samples with other fusion genes or without any detectable fusion genes by RNA-sequencing are colored in red.

DNA methylation classifiers for novel ALL subtypes (Study III)

In study III, subtype-specific DNA methylation patterns in the common subtypes of ALL were utilized [115], together with new insights on the putative *DUX4* and *ZNF384* subtypes reported in study II, to design DNA methylation-based classifiers for ALL. For this purpose, we further refined our previously designed supervised learning procedure based on nearest shrunken centroids, to expand our original DNA methylation classifier by inclusion of the novel subtypes reported in study II. Classifiers were built for 10 ALL subtypes: T-ALL, HeH, *t(12;21)ETV6-RUNX1*, *t(1;19)TCF3-PBX1*, *11q23/MLL*-rearrangements, *t(9;22)BCR-ABL1*, *dic(9;20)*, *iAMP21*, *DUX4*- and *ZNF384*-rearrangements.

An external cohort of 243 blinded ALL samples with an established subtype at ALL diagnosis, was used to assess the performance of the classifier, which resulted in unequivocal assignment of 87% of the ALL cases to the correct subtype. For validation of the subset of samples for which the classifier failed to assign the correct subtype ($n=31$), a lower than expected blast count was observed, which could have confounded the classifier and account for the discordant results. One sample diagnosed as HeH was classified as a *DUX4*-rearranged subtype. This finding was validated by RNA-sequencing, which confirmed the expression of the *DUX4-IGH* fusion gene.

Next, the designed classifiers were applied as a tool to screen for subtype membership in 284 B-other cases. Over 50% of the B-other patients were assigned to an existing ALL subtype using this DNA methylation-based classification method (**Figure 6**). Interestingly, the most commonly predict-

ed subtype was the *DUX4*-rearranged subtype which together with the *ZNF384*-rearranged subtype accounted for 15% (10% and 5%, respectively) of the B-other samples. To confirm the predicted subtypes, based on their DNA methylation signature, copy number analysis was performed in addition to RNA-sequencing for cases with RNA available. RNA-sequencing identified several novel and known fusion genes in the B-other patients, which confirmed the new subtype class including *DUX4*-rearranged, *ZNF384*-rearranged, *BCR-ABL1*-like and *ETV6-RUNX1*-like.

The emerging *ETV6-RUNX1*-like subtype was characterized fusion genes involving the *ETV6* or *IKZF1* genes together with frequent co-occurring *ETV6* deletions, and with no detectable expression of the canonical *ETV6-RUNX1* fusion gene. Of note, several of these samples had in fact been negative for *t(12;21)ETV6-RUNX1* by FISH or PCR at diagnosis.

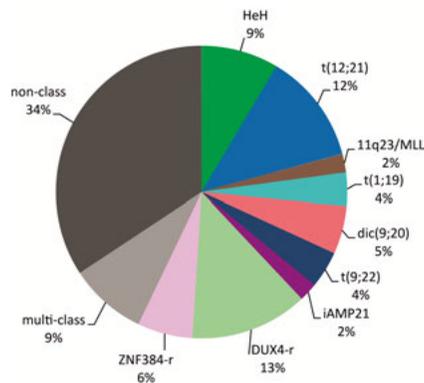


Figure 6. DNA methylation classification in B-other. Distribution of predicted subtypes in B-other including samples that were assigned to more than one class (multi-class) and samples that did not get a subtype score (non-class) for any of the subtypes included in the classifier.

Digital karyotyping by linked-read WGS (Study IV)

In study IV, we evaluated novel linked-read WGS that enables increased resolution compared to standard WGS and phasing of structural variants on each chromosome. Linked-read sequencing was applied for analysis of DNA samples from 12 ALL patients and one ALL cell line (REH) of varying quality. To ensure optimal phasing and computational reconstruction of long haplotypes, isolation of high-molecular weight DNA (HMW DNA) is crucial. We assessed the effect of the quality of input DNA by comparing HMW DNA to that of DNA extracted with standard column-based DNA preparations and had undergone repeated freeze-thawing without selection for long DNA fragments. Although shorter phasing blocks were assigned using standard DNA samples for linked-read WGS, comparable results were

obtained for detection of structural variants in two control samples harboring the $t(12;21)ETV6-RUNX1$ and $t(9;22)BCR-ABL1$ fusion genes.

Linked-read WGS enabled precise digital karyotyping of a wide range of structural variants such as balanced and unbalanced translocations, deletions and aneuploidies. Complex rearrangements involving up to five different chromosomes that were either missed or incorrectly annotated by traditional karyotype analysis were resolved at single-base pair resolution by linked-read WGS (**Figure 7**). As a proof of principle, digital karyotyping detected all subtype-defining events reported at diagnosis in the HeH, $t(12;21)$ and $t(9;22)$ subtypes analyzed in this study. Co-occurring *ETV6* deletions were found in all $t(12;21)ETV6-RUNX1$ patients and by utilizing phasing information, we were able to better resolve the deletion of *ETV6* on the allele that was not affected by the translocation in one patient.

Furthermore, the spectrum of genetic alterations in B-other patients harboring fusion genes involving *DUX4*, *ZNF384* or *PAX5* were delineated at a resolution beyond that of traditional karyotyping or short-read WGS. Owing to the phasing ability of the technology, a previously unobserved bi-allelic *ERG* deletion was discovered in a patient with the *DUX4-IGH* fusion gene (**Figure 7**). Furthermore, we detected heterozygous *IKZF1* deletions in both *DUX4-IGH* cases subjected to linked-read WGS.

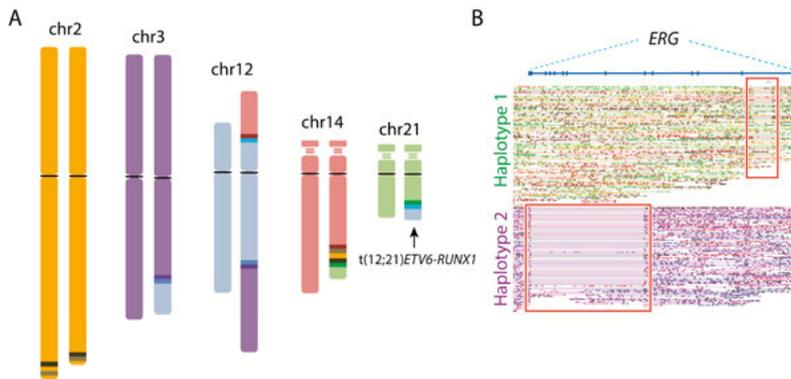


Figure 7. Genomic view of structural rearrangements in ALL by linked-read WGS. (A) Digital karyotyping by linked-read WGS resolved a complex translocations involving five chromosomes (chr 2, 3, 12, 14 and 21) in addition to the subtype defining $t(12;21)ETV6-RUNX1$ aberration indicated with an arrow. (B) Genomic view of the linked-reads mapped to each of the two chromosomes at the *ERG* locus on chr21 in a BCP-ALL patient with the *DUX4-IGH* fusion gene. Reads are color-coded by chromosome (haplotype 1 in green and haplotype 2 in purple) and deletions are marked by red squares.

Summarized discussion

B-other represents a large subgroup of ALL patients where the chromosomal aberrations remain largely undescribed based on currently used diagnostic methods. Identification of the key alterations that define new oncogenic subtypes within the B-other group is imperative, as these patients may benefit from alternative treatment regimens such as more or less intensive treatment protocols or in some cases addition of targeted therapies.

The studies summarized in this thesis have together resolved several recurrent alterations in the B-other group using different genome-wide techniques for molecular profiling in >1000 pediatric ALL patients that were diagnosed and treated according to NOPHO protocols. The large collection of genome-wide DNA methylation data represents a unique backbone for this thesis, while RNA-sequencing and linked-read WGS performed on a subset of patients supports that modern technologies improves the resolution of the chromosomal alterations, especially in the B-other subgroup, by offering precise base-pair delineation of the genetic alterations. Due to the unique nature of the ALL cohort described in this study and the availability of clinical and follow-up data for these patients, we have shown that several of the newly identified alterations detected in B-other [32, 44, 45] are recurrent and prognostically relevant in ALL.

One of the most recurrent alterations discovered in the B-other group is the *DUX4-IGH* fusion gene, which we estimate to be present in approximately 4% of BCP-ALL (14% in B-other). Similar to what has been seen in other studies, NOPHO BCP-ALL patients with *DUX4*-rearrangements have a favorable outcome and *ERG* deletions are frequent in this group. Owing to the phasing ability of linked-read WGS we were able to better resolve a submicroscopic *ERG* deletion in one *DUX4-IGH* patient and identified a compound heterozygous *ERG* deletion, that could not be resolved by RNA-sequencing or copy number analysis, and was undetected in a previous short-read WGS study where this patient was sequenced [98].

Opposite to the *DUX4*-rearranged subtype, *BCR-ABL1*-like is considered a high-risk subtype with dismal prognosis and comprises various kinase-activating lesions [87]. The NOPHO patients harboring tyrosine kinase fusion genes identified by RNA-sequencing in the present studies suffered from frequent relapse and fatal outcomes. Thus, these results strengthen pre-

vious observations that patients with the *BCR-ABL1*-like phenotype may benefit from alternative treatment strategies, such as the addition of targeted tyrosine kinase inhibitors [84].

Moreover, this is the largest study addressing the putative *ETV6-RUNX1*-like subtype to date, in which we confirm the favorable prognosis associated with this emerging subtype and frequent co-occurring *ETV6* deletions [71, 116]. Our findings strengthens that *ETV6-RUNX1*-like is a recurrent putative subtype in BCP-ALL and comprises 1.4-2% of the cases (5-7% in B-other). Additional functional studies will be required in order to understand the molecular mechanism for the *ETV6-RUNX1*-like subtype and its molecular and clinical similarity to the classical *t(12;21)ETV6-RUNX1* subtype, although one might speculate that alterations of the *ETV6* gene is a common denominator between these analogous subtypes.

Interestingly, co-occurring *ETV6* deletions are also common in positive *t(12;21)ETV6-RUNX1* patients and detected in up to 60% of the cases [54, 137]. In each of the three *t(12;21)ETV6-RUNX1* patients we analyzed by linked-read WGS, each case harbored a co-occurring *ETV6* deletion on the non-translocated allele, resulting in a bi-allelic disruption of *ETV6*. Using haplotype information from the linked-read WGS data, the entire *ETV6* locus was phased in one *t(12;21)ETV6-RUNX1* patient that resolved an intergenic *ETV6* deletion on one allele and the translocation forming the *ETV6-RUNX1* fusion gene originating from the other allele.

Fusion genes involving *PAX5* occur in approximately 2.5% of BCP-ALL and is a known heterogeneous fusion gene partner [138]. In our RNA-sequencing studies, we found that *PAX5* forms chimeric transcripts with up to 10 different genes in the patients included herein. Together, fusion genes involving *PAX5* were detected in approximately 2% of the BCP-ALL cases and is thus in line with previous findings [138].

Unlike the established and new putative subtypes of ALL such as those with *DUX4*- and *ZNF384*-rearrangements, the heterogeneous group of patients with *PAX5* fusion genes does not appear to have a strong and distinctive DNA methylation or gene expression pattern (**Figure 5**). Functional studies of *PAX5* chimeras have established both common and distinct properties associated with the different fusion gene partners that fuse to *PAX5* [139]. Notably, the different fusion gene partners may affect the DNA-binding affinity and regulatory potential of the *PAX5* fusion gene in repressing or activating its target genes, consequently affecting the molecular features of the leukemic cells [139].

In particular, we described for the first time a recurrent subgroup characterized by the *PAX5-ESRRB* fusion gene from an original cohort of 664 Nordic BCP-ALL samples diagnosed before 2008. The four patients with *PAX5-ESRRB* in our study did display distinct DNA methylation profiles that dif-

ferred from that of other *PAX5* fusion genes, suggesting that there is some degree of unique molecular signatures, at least in this small group of patients. Nevertheless, this analysis was performed in a small sample set and warrants further investigation in order to understand the diversity of *PAX5* chimeras and to establish whether other recurrent *PAX5* chimeras share distinct and unique molecular features. Interestingly, although we performed copy number analysis and RNA-sequencing in additional recently diagnosed B-other patients, no other patients with *PAX5-ESRRB* were detected in our studies nor have other studies on BCP-ALL reported this *PAX5* fusion gene, which suggests that this fusion gene is extremely rare in BCP-ALL.

Although the studies presented in this thesis resolved the genomic complexity for over half of the patients in the B-other group thus enabling an improved delineation of the genetic alterations in this subgroup, the underlying genetic alteration still remains unknown for a subset of the patients. Consequently, further analysis is required for these patients in order to fully resolve the lesions driving the leukemia, and potentially identifying new subtypes.

From a technical perspective, we demonstrate the power of DNA methylation as a marker for ALL subtype prediction that can be used to retrospectively revise the subtype membership of a large proportion of previously undefined B-other cases. A main strength with this approach is that subtype membership can be given to patients diagnosed over 20 years ago. This reflects the stability of DNA methylation as a biomarker even for DNA samples that were extracted from cryopreserved patient samples that had been biobanked for a long time. Nonetheless, the DNA methylation profile alone is currently not sufficient to confidently establish a subtype membership, and we must still rely on complementary methods such as copy number analysis or verification of expressed fusion genes by RT-PCR, RNA-sequencing or RACE, to unequivocally assign a new subtype to an ALL patient.

We found that our DNA methylation classifiers are sensitive to leukemic cell composition. Samples with lower leukemic blast count could not be accurately assigned to the expected subtype, whereas RNA-sequencing from the same sample with low blast count successfully identified the expected fusion gene. The restricted dynamic range of detection by DNA methylation as compared to RNA-sequencing, which offers a greater dynamic range and is less sensitive to tumor content, is an important aspect when considering a robust test for clinical diagnostics.

Although DNA methylation classification has been a valuable tool for screening of ALL subtypes in the large retrospective cohort described in this thesis, it is unlikely that this approach can be implemented as a diagnostic test for ALL in a clinical setting. Instead, methods that are based on fusion gene detection, such as RNA-sequencing or similar, would be more conven-

ient as they provide a direct answer of expressed fusion transcripts and its involved gene partners. Knowing the exact fusion gene may be beneficial for MRD screening [41] and also for subtypes such as *ZNF384*-rearranged and 11q23/*MLL* as they may confer different clinical outcomes depending on the fusion gene partner [79, 140].

Lastly, we evaluated the utility of linked-read WGS technology for detection and phasing of structural variants in 13 ALL genomes. With this technology we were able to resolve nearly all of the aberrations detected by standard diagnostic analysis, and in the majority of the cases we were able to provide an improved delineation of the true structure of the chromosomal rearrangements. To achieve long haplotype blocks and thus optimal phasing of the diploid chromosomes, HMW DNA is recommended as source for library preparation. We found that, in our hands, standard column-based DNA extraction methods implemented in this project provided DNA molecules of sufficient length for accurate detection of structural variants that performed comparably to that of HMW DNA for detection of the known, subtype-defining translocations, albeit at the cost of shorter haplotype blocks. Consequently, this suggests that older biobanked DNA that has been extracted with standard columns may be used for genomic profiling by linked-read WGS. Although shorter haplotype blocks were obtained, we show that phasing information was not necessarily required to resolve several complex chromosomal rearrangements using standard DNA. Nonetheless, information about phased structural variants owing to longer haplotype blocks enabled detection of allele-specific alterations such as the abovementioned *ERG* and *ETV6* deletions, thus enabling a more comprehensive delineation of the oncogenic lesions in ALL, consequently it would be highly recommended that HMW DNA is used when possible.

The rearrangement involving *DUX4-IGH* presents a particularly difficult aberration to resolve, and possibly why this fusion gene has not been identified until recently with RNA-sequencing studies. A likely explanation is the complex genomic region where *DUX4* is located i.e. within the tandem repeat region D4Z4 consisting of 11-100 of repeat segments. The D4Z4 macrosatellite array is found within the subtelomeric region of chr4q and chr10q [141]. With the computational approaches used in this thesis, the insertion of *DUX4* into the *IGH* locus was not be adequately resolved by linked-read WGS. Thus complementary assays, such as RNA-sequencing, were required to detect this aberration.

A particular challenge with linked-read WGS lies in the lack of tools that can accurately call structural variants with high confidence. In this thesis, we specifically focused on applying linked-read WGS to identify and characterize alterations that were already known to be present in the 13 ALL genome from diagnostic karyotyping or previous studies. In some instances we resolved previously undetected structural rearrangements; however they were

all related to a previously defined aberration. It is to be noted that many additional variants were called in each sample that we did not describe nor validate. Additional filtering or refined detection algorithms would be required to reduce the number of called structural variants in these data sets and to enrich for true somatic events. This may have been mitigated by including matched germline DNA in our study design, however this was not feasible for all of the patient samples included in the study. Also, all samples subjected to linked-read WGS had a high tumor purity (>90% leukemic blasts) and detection of aberrations in samples of lower tumor purity may be more challenging. Nonetheless, the linked-read WGS technology enabled simultaneous detection of aneuploidies, translocations and deletions in a single reaction, resulting in precise digital karyotypes, and thus holds great promise for future clinical incorporation.

Concluding remarks

In this thesis, by adopting a combined approach comprising DNA methylation classification, array-based copy number detection, fusion gene detection by RNA-sequencing and linked-read WGS, we could comprehensively dissect the genomic landscape of a large retrospective cohort of Nordic pediatric ALL patients. Collectively, these studies provide a more detailed view of the chromosomal abnormalities and expressed fusion genes in >1000 pediatric ALL patients than was possible based on the routine methods implemented at the time of ALL diagnosis. Many aberrations in ALL arise as a consequence of cryptic or complex rearrangements of the genome and are not detected at diagnosis by current diagnostic methods, such as karyotyping, SNP-arrays and FISH, which do not have the resolution necessary to comprehensively detect these complex rearrangements. Thus, new sequencing-based approaches are promising techniques to overcome this problem. Technical advancements in molecular profiling coupled with the reduction in sequencing costs are pivotal factors that make genome-wide sequencing more feasible for clinical diagnosis. Indeed, the simplicity of targeted panels for screening of a subset of specific aberrations may be enticing from a clinical context. However, owing to their limited genome-wide coverage, additional acquired alterations that are rare or even private, may not be fully resolved as compared to genome-wide approaches. Together, increased molecular understanding of the emerging subgroups is of great importance for further improved risk stratification and may be considered for future incorporation into clinical diagnostics of ALL.

Acknowledgements

I'm truly grateful to all the patients and their families without whom this work would not have been possible. Numerous people have contributed to this thesis in many different ways, and I want to express my sincere gratitude to all of you for your support along this journey. In particular, I would like to thank:

My main supervisor, **Ann-Christine Syvänen**, for welcoming me into your group and for giving me the opportunity to undertake such challenging yet very rewarding projects. I've learned a lot during these years and I'm glad you gave me the opportunity to learn bioinformatics.

My co-supervisors **Sören Lehmann** and **Jessica Nordlund** for taking me on as a student. **Jessica**, this would not have been possible without you. I am forever grateful for everything you have done, for always being there to answer my several doubts and questions and of course, proof-reading of manuscripts. I admire your enthusiasm for research and you've been invaluable for driving things forward with new and exciting ideas. Thank you for always motivating me and for feeding me with fruit, candy and cookies during these thesis writing weeks 😊. Not only my co-supervisor, you've been a wonderful and very thoughtful friend throughout all these years!

To all clinicians, researchers and associated collaborators within Nordic Society of Pediatric Hematology and Oncology (NOPHO). In particular **Gudmar Lönnérholm** and **Erik Forestier** for sharing your vast knowledge about ALL and all your work enabling this thesis. To **Lucia Cavalier** for your positive spirit and for taking the time to perform those final FISH experiments to boost our papers! To **Vasilios Zachariadis** for enlightening discussions and successful collaborations.

To the Molecular Medicine group including past and present members **Amanda, Christofer, Ekaterina, Elin, Erika, Eva, Frida, Lina, Märten, Mathias, Michelle, Per** and **Tom**. For creating a fun and inspirational environment that made these years very much enjoyable. To **Katarina** for all the administrative work. **Anders, Jonas** and **Johan**, for your bioinformatic guidance and patience during my "data scientist-maturation". Special thanks to my former office mates **Juliana, Nour** and **Sara Nilsson** with whom I

have shared many good times, pizzas and laughter's with **Sara Nystedt**, for your contribution to all our projects, especially for your help with the RACE experiments, and for the great office company together with Jonas and Anders my last year. **Shumaila**, for always being such a positive spirit and for the company during late lunches.

To everyone at the **SNP&SEQ** platform whom carried out an excellent work of generating high quality sequencing and array data for this thesis. For creating a friendly and inspirational atmosphere during lab meetings and coffee breaks.

To all Senior PhDs I befriended during my time at the Rudbeck laboratory that opened my eyes to the world of science. **Ammar, Demet, Dijana** (soon to be PhD), **Fiona, Nicki, Paul, Sara** and others. Thank you for all the encouraging pep talks and for being a brilliant source of information after I decided to embark this PhD-expedition. My sweet **Leli-Peli**, you are that funny-Irish-adopted-sister we always wanted in my family! Thank you for always putting a smile on my face and for proofreading my kappa.

My wonderful girl squads: Jansson (**Anna**), Felin (**Elin**), Kakan Karlsson (**Hanna**), Fidu (**Ida**), Kriss (**Kristina**), Mia-Maria (**Maria**), Merry-Berry (**Merima**), Mims (**Midja**), Monique (**Mona**), Vero (**Veronica**). My memories are forever colorfully painted with all the amazing moments and adventures we have shared. For all the times you've listened to my naggings about life (specially the horrifying PhD life) over pints of beer. For simply making life much more fun! Thank you!

To past and present PhD comrades and others with whom I shared many gym classes to release some PhD frustration: **Fred, Costas, Panos, Cristina** and **Claudia**. Although I've mostly been at the gym in "spirit" lately, I will try to go more regularly from now on. I hope.

My Latin circle and extended family: **Irma** "grande", **Irma** "chica", **Dayana, Sonia, Alvaro, Cecilia, Mary, Marcela, Evert, Ernesto, Carlos, Mario, Paulo** and many more. You've followed me throughout the milestones of my life, always surrounding me with love and support. Gracias por todo el amor y apoyo, especialmente durante este último año. Los quiero muchísimo!

My beloved **Viktor**, for being such a caring, understanding and AMAZING person. For your curiosity of my work and actually wanting to understand what I do. There is not enough space to write all the things you've done to facilitate this journey for me and how truly grateful I am for that. Especially

all the lunch boxes you've made me! Tack för att du stått ut med mig de här senaste månaderna, min klippa!

My **family**, my world. There are no words to express how fortunate I am to have you in my life. You have been a constant source of love, happiness, comfort and encouragement from day one. My dear **dad Antonio**, for always reminding me to never give up and for the loyal morning/evening calls just to say hello (don't ever stop with that!). My **sister Millaray**, my absolute best friend through thick and thin. Thank you for all your help and comfort during this chaotic year, and for pushing me all the way to the finish line. I know we never really "cut our umbilical cords" to mom and I don't understand how we are still alive, but I know having you by my side has helped me through the hard times. My precious nephews, **Filip** and **Oliver**, for the pure joy you spread wherever you go. I love you so much my heart literally explodes. **Alex**, for everything you do and have done for our family. Gracias por ser como son, los amo con toda mi alma.

Mami, te extraño cada día. Gracias por todo lo que me enseñaste y por todo el amor que dejaste en este mundo. Por los recuerdos que llevaré en mi corazón. Fuiste una fuente inagotable de amor, apoyo y ánimo. Aunque no estés, siempre estamos juntos, má.

Esto es para ti, te amo infinito.

Yanara Marincevic-Zuniga
Uppsala, 2018

References

1. Hjalgrim, L.L., et al., Age- and sex-specific incidence of childhood leukemia by immunophenotype in the Nordic countries. *J Natl Cancer Inst*, 2003. **95**(20): p. 1539-44.
2. Hunger, S.P., et al., Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005: a report from the children's oncology group. *J Clin Oncol*, 2012. **30**(14): p. 1663-9.
3. Toft, N., et al., Results of NOPHO ALL2008 treatment for patients aged 1-45 years with acute lymphoblastic leukemia. *Leukemia*, 2018. **32**(3): p. 606-615.
4. Pui, C.H. and W.E. Evans, A 50-year journey to cure childhood acute lymphoblastic leukemia. *Semin Hematol*, 2013. **50**(3): p. 185-96.
5. Cooper, S.L. and P.A. Brown, Treatment of pediatric acute lymphoblastic leukemia. *Pediatr Clin North Am*, 2015. **62**(1): p. 61-73.
6. Hunger, S.P. and C.G. Mullighan, Acute Lymphoblastic Leukemia in Children. *N Engl J Med*, 2015. **373**(16): p. 1541-52.
7. Oskarsson, T., et al., Relapsed childhood acute lymphoblastic leukemia in the Nordic countries: prognostic factors, treatment and outcome. *Haematologica*, 2016. **101**(1): p. 68-76.
8. Paulsson, K. and B. Johansson, High hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosomes Cancer*, 2009. **48**(8): p. 637-60.
9. Paulsson, K., et al., Genetic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A*, 2010. **107**(50): p. 21719-24.
10. Okuda, T., et al., AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell*, 1996. **84**(2): p. 321-330.
11. Wang, L.C., et al., The TEL/ETV6 gene is required specifically for hematopoiesis in the bone marrow. *Genes & Development*, 1998. **12**(15): p. 2392-2402.
12. Lopez, R., et al., TEL is a sequence-specific transcriptional repressor. *Journal of Biological Chemistry*, 1999. **274**(42): p. 30132-30138.
13. Sun, C.C., L.X. Chang, and X.F. Zhu, Pathogenesis of ETV6/RUNX1-positive childhood acute lymphoblastic leukemia and mechanisms underlying its relapse. *Oncotarget*, 2017. **8**(21): p. 35445-35459.
14. Ford, A.M. and M. Greaves, ETV6-RUNX1 (+) Acute Lymphoblastic Leukaemia in Identical Twins. *Adv Exp Med Biol*, 2017. **962**: p. 217-228.
15. Mori, H., et al., Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proc Natl Acad Sci U S A*, 2002. **99**(12): p. 8242-7.
16. Wiemels, J.L., et al., Protracted and variable latency of acute lymphoblastic leukemia after TEL-AML1 gene fusion in utero. *Blood*, 1999. **94**(3): p. 1057-62.

17. Schafer, D., et al., Five percent of healthy newborns have an ETV6-RUNX1 fusion as revealed by DNA-based GIPFEL screening. *Blood*, 2018. **131**(7): p. 821-826.
18. Bernt, K.M. and S.P. Hunger, Current concepts in pediatric Philadelphia chromosome-positive acute lymphoblastic leukemia. *Front Oncol*, 2014. **4**: p. 54.
19. Sanyal, M., et al., B-cell development fails in the absence of the Pbx1 proto-oncogene. *Blood*, 2007. **109**(10): p. 4191-9.
20. Bain, G., et al., E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. *Cell*, 1994. **79**(5): p. 885-92.
21. Bhojwani, D., J.J. Yang, and C.H. Pui, Biology of childhood acute lymphoblastic leukemia. *Pediatr Clin North Am*, 2015. **62**(1): p. 47-60.
22. Yu, G., et al., Upfront treatment with the first and second-generation tyrosine kinase inhibitors in Ph-positive acute lymphoblastic leukemia. *Oncotarget*, 2017. **8**(63): p. 107022-107032.
23. Meyer, C., et al., The MLL recombinome of acute leukemias in 2017. *Leukemia*, 2018. **32**(2): p. 273-284.
24. Hyun, K., et al., Writing, erasing and reading histone lysine methylations. *Exp Mol Med*, 2017. **49**(4): p. e324.
25. Greaves, M., When one mutation is all it takes. *Cancer Cell*, 2015. **27**(4): p. 433-4.
26. Harrison, C.J., Blood Spotlight on iAMP21 acute lymphoblastic leukemia (ALL), a high-risk pediatric disease. *Blood*, 2015. **125**(9): p. 1383-6.
27. Soulier, J., et al., Amplification of band q22 of chromosome 21, including AML1, in older children with acute lymphoblastic leukemia: an emerging molecular cytogenetic subgroup. *Leukemia*, 2003. **17**(8): p. 1679-82.
28. Harrison, C.J., et al., An international study of intrachromosomal amplification of chromosome 21 (iAMP21): cytogenetic characterization and outcome. *Leukemia*, 2014. **28**(5): p. 1015-21.
29. Zachariadis, V., et al., The frequency and prognostic impact of dic(9;20)(p13.2;q11.2) in childhood B-cell precursor acute lymphoblastic leukemia: results from the NOPHO ALL-2000 trial. *Leukemia*, 2011. **25**(4): p. 622-8.
30. Schoumans, J., et al., Characterisation of dic(9;20)(p11-13;q11) in childhood B-cell precursor acute lymphoblastic leukaemia by tiling resolution array-based comparative genomic hybridisation reveals clustered breakpoints at 9p13.2 and 20q11.2. *Br J Haematol*, 2006. **135**(4): p. 492-9.
31. An, Q., et al., Heterogeneous breakpoints in patients with acute lymphoblastic leukemia and the dic(9;20)(p11-13;q11) show recurrent involvement of genes at 20q11.21. *Haematologica*, 2009. **94**(8): p. 1164-9.
32. Iacobucci, I. and C.G. Mullighan, Genetic Basis of Acute Lymphoblastic Leukemia. *J Clin Oncol*, 2017. **35**(9): p. 975-983.
33. Karrman, K. and B. Johansson, Pediatric T-cell acute lymphoblastic leukemia. *Genes Chromosomes Cancer*, 2017. **56**(2): p. 89-116.
34. Belver, L. and A. Ferrando, The genetics and mechanisms of T cell acute lymphoblastic leukaemia. *Nat Rev Cancer*, 2016. **16**(8): p. 494-507.
35. Chen, B., et al., Identification of fusion genes and characterization of transcriptome features in T-cell acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A*, 2018. **115**(2): p. 373-378.

36. Schmiegelow, K., et al., Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. *Leukemia*, 2010. **24**(2): p. 345-54.
37. Stanulla, M. and M. Schrappe, Treatment of childhood acute lymphoblastic leukemia. *Semin Hematol*, 2009. **46**(1): p. 52-63.
38. Pui, C.H., et al., Clinical impact of minimal residual disease in children with different subtypes of acute lymphoblastic leukemia treated with Response-Adapted therapy. *Leukemia*, 2017. **31**(2): p. 333-339.
39. Cave, H., et al., Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia. European Organization for Research and Treatment of Cancer--Childhood Leukemia Cooperative Group. *N Engl J Med*, 1998. **339**(9): p. 591-8.
40. Campana, D., Minimal residual disease in acute lymphoblastic leukemia. *Hematology Am Soc Hematol Educ Program*, 2010. **2010**: p. 7-12.
41. Alm, S.J., et al., Minimal residual disease monitoring in childhood B lymphoblastic leukemia with t(12;21)(p13;q22); ETV6-RUNX1: concordant results using quantitation of fusion transcript and flow cytometry. *Int J Lab Hematol*, 2017. **39**(2): p. 121-128.
42. Huang, Y.J., et al., Concordance of two approaches in monitoring of minimal residual disease in B-precursor acute lymphoblastic leukemia: Fusion transcripts and leukemia-associated immunophenotypes. *J Formos Med Assoc*, 2017. **116**(10): p. 774-781.
43. Pui, C.H., et al., Childhood Acute Lymphoblastic Leukemia: Progress Through Collaboration. *J Clin Oncol*, 2015. **33**(27): p. 2938-48.
44. Lilljebjorn, H. and T. Fioretos, New oncogenic subtypes in pediatric B-cell precursor acute lymphoblastic leukemia. *Blood*, 2017. **130**(12): p. 1395-1401.
45. Tasian, S.K. and S.P. Hunger, Genomic characterization of paediatric acute lymphoblastic leukaemia: an opportunity for precision medicine therapeutics. *Br J Haematol*, 2017. **176**(6): p. 867-882.
46. Hoheisel, J.D., Microarray technology: beyond transcript profiling and genotype analysis. *Nat Rev Genet*, 2006. **7**(3): p. 200-10.
47. Bumgarner, R., Overview of DNA microarrays: types, applications, and their future. *Curr Protoc Mol Biol*, 2013. **Chapter 22**: p. Unit 22 1.
48. Moorman, A.V., The clinical relevance of chromosomal and genomic abnormalities in B-cell precursor acute lymphoblastic leukaemia. *Blood Rev*, 2012. **26**(3): p. 123-35.
49. Wang, Y. and S.A. Armstrong, Genome-wide SNP analysis in cancer: leukemia shows the way. *Cancer Cell*, 2007. **11**(4): p. 308-9.
50. Mullighan, C.G. and J.R. Downing, Global genomic characterization of acute lymphoblastic leukemia. *Semin Hematol*, 2009. **46**(1): p. 3-15.
51. Mullighan, C.G., The molecular genetic makeup of acute lymphoblastic leukemia. *Hematology Am Soc Hematol Educ Program*, 2012. **2012**: p. 389-96.
52. Mullighan, C.G., et al., Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*, 2007. **446**(7137): p. 758-64.
53. Kuiper, R.P., 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): p. 1258-66.

54. Schwab, C.J., et al., Genes commonly deleted in childhood B-cell precursor acute lymphoblastic leukemia: association with cytogenetics and clinical features. *Haematologica*, 2013. **98**(7): p. 1081-8.
55. van der Veer, A., et al., IKZF1 status as a prognostic feature in BCR-ABL1-positive childhood ALL. *Blood*, 2014. **123**(11): p. 1691-8.
56. Mullighan, C.G., et al., Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med*, 2009. **360**(5): p. 470-80.
57. Moorman, A.V., New and emerging prognostic and predictive genetic biomarkers in B-cell precursor acute lymphoblastic leukemia. *Haematologica*, 2016. **101**(4): p. 407-416.
58. Rand, V., et al., Genomic characterization implicates iAMP21 as a likely primary genetic event in childhood B-cell precursor acute lymphoblastic leukemia. *Blood*, 2011. **117**(25): p. 6848-55.
59. Cazzaniga, G., et al., Developmental origins and impact of BCR-ABL1 fusion and IKZF1 deletions in monozygotic twins with Ph+ acute lymphoblastic leukemia. *Blood*, 2011. **118**(20): p. 5559-64.
60. Paulsson, K., High hyperdiploid childhood acute lymphoblastic leukemia: Chromosomal gains as the main driver event. *Mol Cell Oncol*, 2016. **3**(1): p. e1064555.
61. Heltemes-Harris, L.M., et al., Ebf1 or Pax5 haploinsufficiency synergizes with STAT5 activation to initiate acute lymphoblastic leukemia. *J Exp Med*, 2011. **208**(6): p. 1135-49.
62. Virely, C., et al., Haploinsufficiency of the IKZF1 (IKAROS) tumor suppressor gene cooperates with BCR-ABL in a transgenic model of acute lymphoblastic leukemia. *Leukemia*, 2010. **24**(6): p. 1200-4.
63. Smeenk, L., et al., Molecular role of the PAX5-ETV6 oncoprotein in promoting B-cell acute lymphoblastic leukemia. *EMBO J*, 2017. **36**(6): p. 718-735.
64. Figueroa, M.E., et al., Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia. *J Clin Invest*, 2013. **123**(7): p. 3099-111.
65. Ross, M.E., et al., Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. *Blood*, 2003. **102**(8): p. 2951-9.
66. Yeoh, E.J., et al., Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell*, 2002. **1**(2): p. 133-43.
67. Den Boer, M.L., et al., A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. *Lancet Oncol*, 2009. **10**(2): p. 125-34.
68. Tasian, S.K., M.L. Loh, and S.P. Hunger, Philadelphia chromosome-like acute lymphoblastic leukemia. *Blood*, 2017. **130**(19): p. 2064-2072.
69. Arber, D.A., et al., The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*, 2016. **127**(20): p. 2391-405.
70. Kukurba, K.R. and S.B. Montgomery, RNA Sequencing and Analysis. *Cold Spring Harb Protoc*, 2015. **2015**(11): p. 951-69.
71. Lilljebjorn, H., et al., Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia. *Nat Commun*, 2016. **7**: p. 11790.
72. Yasuda, T., et al., Recurrent DUX4 fusions in B cell acute lymphoblastic leukemia of adolescents and young adults. *Nat Genet*, 2016. **48**(5): p. 569-74.
73. Liu, Y.F., et al., Genomic Profiling of Adult and Pediatric B-cell Acute Lymphoblastic Leukemia. *EBioMedicine*, 2016. **8**: p. 173-83.

74. Zhang, J., et al., Deregulation of DUX4 and ERG in acute lymphoblastic leukemia. *Nat Genet*, 2016. **48**(12): p. 1481-1489.
75. Mullighan, C.G., et al., ERG deletions define a novel subtype of B-progenitor acute lymphoblastic leukemia. *Blood*, 2007. **110**(11): p. 212a-213a.
76. Clappier, E., et al., An intragenic ERG deletion is a marker of an oncogenic subtype of B-cell precursor acute lymphoblastic leukemia with a favorable outcome despite frequent IKZF1 deletions. *Leukemia*, 2014. **28**(1): p. 70-7.
77. Harvey, R.C., et al., Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome. *Blood*, 2010. **116**(23): p. 4874-84.
78. Marincevic-Zuniga, Y., et al., Transcriptome sequencing in pediatric acute lymphoblastic leukemia identifies fusion genes associated with distinct DNA methylation profiles. *J Hematol Oncol*, 2017. **10**(1): p. 148.
79. Hirabayashi, S., et al., ZNF384-related fusion genes define a subgroup of childhood B-cell precursor acute lymphoblastic leukemia with a characteristic immunotype. *Haematologica*, 2017. **102**(1): p. 118-129.
80. Qian, M., et al., Whole-transcriptome sequencing identifies a distinct subtype of acute lymphoblastic leukemia with predominant genomic abnormalities of EP300 and CREBBP. *Genome Res*, 2017. **27**(2): p. 185-195.
81. Suzuki, K., et al., MEF2D-BCL9 Fusion Gene Is Associated With High-Risk Acute B-Cell Precursor Lymphoblastic Leukemia in Adolescents. *J Clin Oncol*, 2016. **34**(28): p. 3451-9.
82. Gu, Z., et al., Genomic analyses identify recurrent MEF2D fusions in acute lymphoblastic leukaemia. *Nat Commun*, 2016. **7**: p. 13331.
83. Roberts, K.G., et al., Genetic Alterations Activating Kinase and Cytokine Receptor Signaling in High-Risk Acute Lymphoblastic Leukemia. *Cancer Cell*, 2012. **22**(2): p. 153-166.
84. Roberts, K.G., et al., Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. *N Engl J Med*, 2014. **371**(11): p. 1005-15.
85. Imamura, T., et al., Characterization of pediatric Philadelphia-negative B-cell precursor acute lymphoblastic leukemia with kinase fusions in Japan. *Blood Cancer J*, 2016. **6**: p. e419.
86. Boer, J.M., et al., Tyrosine kinase fusion genes in pediatric BCR-ABL1-like acute lymphoblastic leukemia. *Oncotarget*, 2017. **8**(3): p. 4618-4628.
87. Boer, J.M. and M.L. den Boer, BCR-ABL1-like acute lymphoblastic leukaemia: From bench to bedside. *Eur J Cancer*, 2017. **82**: p. 203-218.
88. Boer, J.M., 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): p. E354-E357.
89. Weston, B.W., et al., Tyrosine kinase inhibitor therapy induces remission in a patient with refractory EBF1-PDGFRB-positive acute lymphoblastic leukemia. *J Clin Oncol*, 2013. **31**(25): p. e413-6.
90. Levy, S.E. and R.M. Myers, Advancements in Next-Generation Sequencing. *Annu Rev Genomics Hum Genet*, 2016. **17**: p. 95-115.
91. Zheng, G.X., et al., Haplotyping germline and cancer genomes with high-throughput linked-read sequencing. *Nat Biotechnol*, 2016. **34**(3): p. 303-11.
92. Greer, S.U., et al., Linked read sequencing resolves complex genomic rearrangements in gastric cancer metastases. *Genome Med*, 2017. **9**(1): p. 57.
93. Garcia, S., et al., Linked-Read sequencing resolves complex structural variants. *bioRxiv*. doi: <http://dx.doi.org/10.1101/011650>, 2017.

94. Andersson, A.K., et al., The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nat Genet*, 2015. **47**(4): p. 330-7.
95. Alexandrov, L.B., et al., Signatures of mutational processes in human cancer. *Nature*, 2013. **500**(7463): p. 415-21.
96. Holmfeldt, L., et al., The genomic landscape of hypodiploid acute lymphoblastic leukemia. *Nat Genet*, 2013. **45**(3): p. 242-52.
97. Lindqvist, C.M., et al., Deep targeted sequencing in pediatric acute lymphoblastic leukemia unveils distinct mutational patterns between genetic subtypes and novel relapse-associated genes. *Oncotarget*, 2016. **7**(39): p. 64071-64088.
98. Lindqvist, C.M., et al., The mutational landscape in pediatric acute lymphoblastic leukemia deciphered by whole genome sequencing. *Hum Mutat*, 2015. **36**(1): p. 118-28.
99. Roberts, K.G. and C.G. Mullighan, Genomics in acute lymphoblastic leukaemia: insights and treatment implications. *Nat Rev Clin Oncol*, 2015. **12**(6): p. 344-57.
100. Tzoneva, G., et al., Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. *Nat Med*, 2013. **19**(3): p. 368-71.
101. Meyer, J.A., et al., Relapse-specific mutations in NT5C2 in childhood acute lymphoblastic leukemia. *Nat Genet*, 2013. **45**(3): p. 290-4.
102. Laird, P.W., Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet*, 2010. **11**(3): p. 191-203.
103. Bibikova, M., et al., High density DNA methylation array with single CpG site resolution. *Genomics*, 2011. **98**(4): p. 288-95.
104. Bibikova, M., et al., Genome-wide DNA methylation profiling using Infinium(R) assay. *Epigenomics*, 2009. **1**(1): p. 177-200.
105. Sandoval, J., et al., Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics*, 2011. **6**(6): p. 692-702.
106. Pidsley, R., et al., Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol*, 2016. **17**(1): p. 208.
107. Soto, J., et al., The impact of next-generation sequencing on the DNA methylation-based translational cancer research. *Transl Res*, 2016. **169**: p. 1-18 e1.
108. Deaton, A.M. and A. Bird, CpG islands and the regulation of transcription. *Genes Dev*, 2011. **25**(10): p. 1010-22.
109. Jones, P.A., Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*, 2012. **13**(7): p. 484-92.
110. Nordlund, J., et al., Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. *Genome Biol*, 2013. **14**(9): p. r105.
111. Chatterton, Z., et al., Epigenetic deregulation in pediatric acute lymphoblastic leukemia. *Epigenetics*, 2014. **9**(3): p. 459-67.
112. Nordlund, J. and A.C. Syvanen, Epigenetics in pediatric acute lymphoblastic leukemia. *Semin Cancer Biol*, 2017.
113. Borssen, M., et al., Promoter DNA methylation pattern identifies prognostic subgroups in childhood T-cell acute lymphoblastic leukemia. *PLoS One*, 2013. **8**(6): p. e65373.
114. Borssen, M., et al., DNA methylation holds prognostic information in relapsed precursor B-cell acute lymphoblastic leukemia. *Clin Epigenetics*, 2018. **10**: p. 31.

115. Nordlund, J., et al., DNA methylation-based subtype prediction for pediatric acute lymphoblastic leukemia. *Clin Epigenetics*, 2015. **7**(1): p. 11.
116. Zaliouva, M., et al., ETV6/RUNX1-like acute lymphoblastic leukemia: A novel B-cell precursor leukemia subtype associated with the CD27/CD44 immunophenotype. *Genes Chromosomes Cancer*, 2017. **56**(8): p. 608-616.
117. Bhoi, S., et al., Prognostic impact of epigenetic classification in chronic lymphocytic leukemia: The case of subset #2. *Epigenetics*, 2016. **11**(6): p. 449-55.
118. Capper, D., et al., DNA methylation-based classification of central nervous system tumours. *Nature*, 2018. **555**(7697): p. 469-474.
119. Danielsson, A., et al., MethPed: a DNA methylation classifier tool for the identification of pediatric brain tumor subtypes. *Clin Epigenetics*, 2015. **7**: p. 62.
120. Hao, X., et al., DNA methylation markers for diagnosis and prognosis of common cancers. *Proc Natl Acad Sci U S A*, 2017. **114**(28): p. 7414-7419.
121. Aryee, M.J., et al., Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, 2014. **30**(10): p. 1363-9.
122. Dedeurwaerder, S., et al., Evaluation of the Infinium Methylation 450K technology. *Epigenomics*, 2011. **3**(6): p. 771-84.
123. Martin, M., Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. doi: <http://dx.doi.org/10.14806/ej.17.1.200>, 2011.
124. Andrews, S., FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>, 2010.
125. Dobin, A., et al., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 2013. **29**(1): p. 15-21.
126. Kim, D., et al., TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*, 2013. **14**(4): p. R36.
127. Wang, L., S. Wang, and W. Li, RSeQC: quality control of RNA-seq experiments. *Bioinformatics*, 2012. **28**(16): p. 2184-5.
128. DeLuca, D.S., et al., RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics*, 2012. **28**(11): p. 1530-2.
129. Nicorici, D.S., M.; Edgren, H.; Kangaspeska, S.; Murumagi, A.; Kallioniemi, O.; Virtanen, S.; Kilkku, O., FusionCatcher - a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *bioRxiv*. doi: <http://dx.doi.org/10.1101/011650>, 2014.
130. Trapnell, C., et al., Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*, 2010. **28**(5): p. 511-5.
131. Liao, Y., G.K. Smyth, and W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 2014. **30**(7): p. 923-30.
132. Love, M.I., W. Huber, and S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 2014. **15**(12): p. 550.
133. Trapnell, C., et al., Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol*, 2013. **31**(1): p. 46-53.

134. Marzouka, N.A., et al., CopyNumber450kCancer: baseline correction for accurate copy number calling from the 450k methylation array. *Bioinformatics*, 2016. **32**(7): p. 1080-2.
135. Rasmussen, M., et al., Allele-specific copy number analysis of tumor samples with aneuploidy and tumor heterogeneity. *Genome Biol*, 2011. **12**(10): p. R108.
136. Abyzov, A., et al., CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res*, 2011. **21**(6): p. 974-84.
137. Lilljebjorn, H., et al., The correlation pattern of acquired copy number changes in 164 ETV6/RUNX1-positive childhood acute lymphoblastic leukemias. *Hum Mol Genet*, 2010. **19**(16): p. 3150-8.
138. Nebral, K., et al., Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia. *Leukemia*, 2009. **23**(1): p. 134-43.
139. Fortschegger, K., et al., Functional heterogeneity of PAX5 chimeras reveals insight for leukemia development. *Mol Cancer Res*, 2014. **12**(4): p. 595-606.
140. Pui, C.H., et al., Outcome of treatment in childhood acute lymphoblastic leukaemia with rearrangements of the 11q23 chromosomal region. *Lancet*, 2002. **359**(9321): p. 1909-15.
141. van Geel, M., et al., Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. *Genomics*, 2002. **79**(2): p. 210-7.

Acta Universitatis Upsaliensis

*Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Medicine 1452*

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-345085



ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2018