IKAROS and LEUKEMIA

Olsson, Linda

2014

Link to publication

Citation for published version (APA):
Olsson, L. (2014). IKAROS and LEUKEMIA. Division of Clinical Genetics, Lund University.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
IKAROS and LEUKEMIA

The images on the cover show the p-arm of chromosome 7 with IKZF1 deletions. The gene is denoted in red in two pediatric B-cell precursor acute lymphoblastic leukemia cases that subsequently relapsed. On the front page, this deletion is focal (involving only IKZF1); it is non-focal in the case below. Here, the deletion was present in a minor subclone at diagnosis (top two panels) that became the major clone at relapse (lower two panels); this indicates that this clone was resistant to therapy and hence important to detect already at the time of diagnosis.
IKAROS and LEUKEMIA

Linda Olsson

LUND UNIVERSITY

DOCTORAL DISSERTATION

With the approval of the Faculty of Medicine, Lund University, this thesis will be defended on November 22, 2014, at 10.00 in Belfragesalen, Lund, Sweden.

Faculty opponent

Professor Nicholas Cross, PhD
Wessex Regional Genetics Laboratory, Salisbury
Faculty of Medicine, University of Southampton
United Kingdom
Acute lymphoblastic leukemia (ALL) is characterized by an accumulation of immature lymphoid cells in the bone marrow and is the most common cancer type in children. It is an immunophenotypically, morphologically, clinically, and genetically heterogeneous disorder that comprises several distinct subtypes. Proper classification is important because determining the correct subtype plays a vital role for prognostication and treatment strategy.

During the last decades, the use of polychemotherapy and implementation of risk stratification based on the presence of certain acquired genetic changes in pediatric B-cell precursor (BCP) ALL have increased the overall survival rates substantially; they are now approaching 90%. This notwithstanding, 20% of patients still relapse and only half of these survive. A considerable proportion of all relapses lacks the high risk-stratifying genetic changes included in most current ALL treatment protocols. Hence, it is important to identify novel genetic features associated with treatment failure to ensure proper therapy intensity and to detect genes and pathways that in the future can be targeted by specific drugs.

To identify relapse-associated genetic aberrations in pediatric BCP ALL, single nucleotide polymorphism array analyses were performed on uniformly treated patients accrued between 1992 and 2011 from the Lund and Linköping University Hospitals (Article I). In the 191 successfully analyzed cases, deletions of IKZF1 (ΔIKZF1) and SPRED1 were shown to be associated with a poor prognosis, with ΔIKZF1 being an independent risk factor for relapse. To ascertain whether ΔIKZF1 is an independent risk factor also in the context of minimal residual disease (MRD) findings, an extended cohort including all 334 Swedish pediatric BCP ALL cases with known IKZF1 status was investigated (Article II). That study confirmed that ΔIKZF1 confers a poor prognosis, revealed that such deletions are particularly common in cases with uninformative cytogenetics, and showed that the prognostic impact of ΔIKZF1 is independent of MRD stratification. However, coexisting genetic changes may play a role in modifying the pathogenetic and/or clinical impact of ΔIKZF1. Therefore, ΔIKZF1-positive cases were investigated further in Article III in order to identify additional, recurrent changes.

Furthermore, targeted deep sequencing of all IKZF1 exons in 140 BCP ALL cases was performed, identifying sequence mutations (mutIKZF1) in 5.7%. Of the mutIKZF1-positive cases, one-fourth also harbored ΔIKZF1. In total, 35 cases with IKZF1 abnormalities (abnIKZF1), comprising ΔIKZF1 and/or mutIKZF1, could be analyzed with regard to other genetic anomalies. These analyses showed that CRLF2 rearrangements, caused by deletions of the pseudoautosomal region 1 (PAR1), and JAK2 mutations were significantly overrepresented in abnIKZF1-positive cases and that the presence of PAR1 deletions conferred a poor prognostic impact. Thus, in order to ascertain correctly the clinical ramifications of abnIKZF1 in pediatric BCP ALL, PAR1 deletions should possibly also be screened for.

Key words: pediatric acute lymphoblastic leukemia, relapse, risk-stratifying factors, IKZF1, cooperative changes

Supplementary bibliographical information

ISSN and key title 1652-8220

Recipient’s notes Number of pages 250

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature Date 15/10-14
IKAROS and LEUKEMIA

Linda Olsson

Division of Clinical Genetics
Department of Laboratory Medicine
Faculty of Medicine

Lund University
2014
Cover: The SNP array images were extracted by me from the GenomeStudio software 2011.1 (Illumina, San Diego, CA).

© Linda Olsson

LUND UNIVERSITY
Faculty of Medicine Doctoral Dissertation Series 2014:121
ISBN 978-91-7619-050-0
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2014
“When I was a young investigator I was researching for a transcriptional regulator that would function as a master builder for the lymphoid system. I did not like acronyms such as NF1 or ZnF1(!!!) as I felt were quite anonymous, boring and not the least inspiring. I thought a better fitting name for this gene would be IKAROS, the son of the greatest master builder in the ancient Greek world Dedalos. So IKAROS once removed from a master regulator of the hematopoietic system confers lymphoid cell fate and regulates homeostasis of immature and mature lymphocytes. Unlike Dedalos, IKAROS was playful and perhaps foolish and his demise/fall costed him as well as us (mice and men) with a failing immune system and the rise of leukemias.”

Professor Katia Georgopoulos,

Harvard Medical School, who was the one identifying and characterizing the role of IKAROS in hematopoiesis in the early 1990s, kindly responding to my e-mail asking her why she named the protein IKAROS (not mentioned in any of the publications on this gene).
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contents</td>
<td>7</td>
</tr>
<tr>
<td>Original Articles</td>
<td>9</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>10</td>
</tr>
<tr>
<td>Preface</td>
<td>11</td>
</tr>
<tr>
<td>Introduction</td>
<td>12</td>
</tr>
<tr>
<td>Leukemia</td>
<td>12</td>
</tr>
<tr>
<td>Pediatric BCP ALL</td>
<td>13</td>
</tr>
<tr>
<td>Epidemiology</td>
<td>13</td>
</tr>
<tr>
<td>Etiology</td>
<td>14</td>
</tr>
<tr>
<td>Diagnosis and Therapy</td>
<td>16</td>
</tr>
<tr>
<td>Cytogenetic Aberrations</td>
<td>17</td>
</tr>
<tr>
<td>Numerical Abnormalities</td>
<td>17</td>
</tr>
<tr>
<td>Structural Abnormalities</td>
<td>20</td>
</tr>
<tr>
<td>Risk-Stratifying Abnormalities</td>
<td>25</td>
</tr>
<tr>
<td>Submicroscopic Genetic Changes</td>
<td>25</td>
</tr>
<tr>
<td>Copy Number Alterations and Uniparental Isodisomies</td>
<td>26</td>
</tr>
<tr>
<td>Sequence Mutations</td>
<td>27</td>
</tr>
<tr>
<td>Gene Expression Profiling</td>
<td>28</td>
</tr>
<tr>
<td>DNA Methylation Profiling</td>
<td>29</td>
</tr>
<tr>
<td>Aberrant MicroRNAs</td>
<td>29</td>
</tr>
<tr>
<td>Relapse and Treatment Failure</td>
<td>30</td>
</tr>
<tr>
<td>Genetic Evolution Patterns at Relapse</td>
<td>30</td>
</tr>
<tr>
<td>Genetic Aberrations Associated with Relapse</td>
<td>31</td>
</tr>
<tr>
<td>Gene Expression Signatures Associated with Relapse</td>
<td>33</td>
</tr>
<tr>
<td>Therapeutic Targets</td>
<td>34</td>
</tr>
<tr>
<td>My PhD Studies in Brief</td>
<td>34</td>
</tr>
<tr>
<td>IKAROS and LEUKEMIA</td>
<td>36</td>
</tr>
<tr>
<td>The Transcription Factor IKAROS</td>
<td>36</td>
</tr>
<tr>
<td>$IKZF1$ in Human Disease</td>
<td>39</td>
</tr>
<tr>
<td>Constitutional $IKZF1$ Abnormalities</td>
<td>39</td>
</tr>
<tr>
<td>Acquired $IKZF1$ Abnormalities</td>
<td>40</td>
</tr>
</tbody>
</table>
**IKZF1 in BCP ALL**

- The IK6 Isoform 42
- ΔIKZF1 42
- *IKZF1* Sequence Mutations 43
- Clinical Features and Significance of *IKZF1* changes 44
- How to detect *IKZF1* Abnormalities 45
- Pathogenetic Impact 45
- Gene Expression Signature 46

**Cooperative Changes in BCP ALL**

- Abnormal Kinase Signaling Pathways and Deletions of *ERG* 47

**Controversies** 48

**Conclusions and Future Directions** 49

Svensk Sammanfattning 50

Acknowledgments 53

References 55

Articles I-III 97
Original Articles

This thesis is based on the following articles, which will be referred to in the text by their Roman numerals.


## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL</td>
<td>Acute lymphoblastic leukemia</td>
</tr>
<tr>
<td>AML</td>
<td>Acute myeloid leukemia</td>
</tr>
<tr>
<td>BCP</td>
<td>B-cell precursor</td>
</tr>
<tr>
<td>BM</td>
<td>Bone marrow</td>
</tr>
<tr>
<td>CLL</td>
<td>Chronic lymphocytic leukemia</td>
</tr>
<tr>
<td>CML</td>
<td>Chronic myeloid leukemia</td>
</tr>
<tr>
<td>CNA</td>
<td>Copy number alteration</td>
</tr>
<tr>
<td>DLBCL</td>
<td>Diffuse large B-cell lymphoma</td>
</tr>
<tr>
<td>DN</td>
<td>Dominant negative</td>
</tr>
<tr>
<td>ΔIKZF1</td>
<td>Deletion of IKZF1</td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescence <em>in situ</em> hybridization</td>
</tr>
<tr>
<td>GCPS</td>
<td>Greig cephalopolysyndactyly syndrome</td>
</tr>
<tr>
<td>HeH</td>
<td>High hyperdiploidy</td>
</tr>
<tr>
<td>HR</td>
<td>High risk</td>
</tr>
<tr>
<td>Indel</td>
<td>Insertion/deletion</td>
</tr>
<tr>
<td>IR</td>
<td>Intermediate risk</td>
</tr>
<tr>
<td>MDS</td>
<td>Myelodysplastic syndrome</td>
</tr>
<tr>
<td>MiRNA</td>
<td>MicroRNA</td>
</tr>
<tr>
<td>MRD</td>
<td>Minimal residual disease</td>
</tr>
<tr>
<td>NOPHO</td>
<td>Nordic Society of Pediatric Hematology and Oncology</td>
</tr>
<tr>
<td>OS</td>
<td>Overall survival</td>
</tr>
<tr>
<td>PB</td>
<td>Peripheral blood</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>RAG</td>
<td>Recombination activating gene</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SR</td>
<td>Standard risk</td>
</tr>
<tr>
<td>UPID</td>
<td>Uniparental isodisomy</td>
</tr>
<tr>
<td>WBC</td>
<td>White blood cell</td>
</tr>
<tr>
<td>wt</td>
<td>Wild type</td>
</tr>
</tbody>
</table>
In the late 1950s and early 1960s, cytogenetic analyses of acute lymphoblastic leukemias (ALLs) revealed that they harbored acquired chromosome changes. During the following decades, several ALL-associated abnormalities were detected, many of which were shown to correlate with particular immunophenotypic and clinical characteristics, including outcome. In fact, genetic aberrations have successfully been implemented as risk-stratifying factors modifying the intensity of the therapy given, something that together with improvements in polychemotherapy has resulted in survival rates of approximately 90% in most contemporary treatment protocols for pediatric B-cell precursor (BCP) ALL. However, relapses, which occur in one-fifth of the children, are still associated with a dismal prognosis. Furthermore, although ALL is one of the genetically best characterized malignant disorders, much remains to be known about the underlying molecular mechanisms for treatment failure and subsequent relapse. The aim of the present thesis was to identify genetic aberrations that influence treatment response, that may be used as novel risk-stratifying markers, and that, in the future, may be directly targeted by therapy.

The thesis is structured into three sections. Part one provides a general background to leukemia, with an emphasis on genetic abnormalities in pediatric BCP ALL, including also the findings in Articles I-III. The second part is a review of the \textit{IKZF1} gene, shown to be associated with poor outcome in Article I and further investigated in Articles II and III, and its role in leukemia. The final and third section comprises the original Articles I-III.

Lund, October 2014
Introduction

Leukemia

In 1844, the French doctor Alfred François Donné was the first to describe a pathological condition in which an unusually high quantity of white blood cells (WBC) was observed in a peripheral blood (PB) sample from a patient with splenomegaly who subsequently died. He then noted the same phenomenon in additional patients. Dr. Donné believed that the excess of WBC should not be interpreted as pus or something associated with inflammation (as previously believed) but that it rather represented a maturation arrest of blood cells (Donné, 1844; Thorburn, 1974; Piller, 2001). This seminal observation notwithstanding, the discovery and understanding of disorders associated with increased number of immature WBC in the PB is largely accredited to the Scottish pathologist John Hughes Bennett, who published one case report (Bennett, 1845; Piller, 2001), and the German pathologist Rudolph Virchow, who first introduced the term “Weisses Blut” and then named the disease “Leukämie”, derived from Greek, meaning white blood, as does also, of course, Weisses Blut (Virchow, 1847; Piller, 2001). These two scientists showed that leukemia is a clinically distinct entity that, in those days, was a uniformly lethal disease.

Today, leukemia is known to be a heterogeneous malignant disorder that comprises several subgroups, albeit with overlapping symptoms, that affect blood formation in the bone marrow (BM) and that are caused by acquired genetic aberrations in hematopoietic stem cells or precursor cells (Pui, 2012). The abnormal, malignant cells accumulate in the BM, often also in the PB, and suppress the normal blood cells, resulting in anemia, thrombocytopenia, and lack of mature functional WBC. This leads to fatigue, bleedings, and infections. Extramedullary involvement (leukemic cells outside the BM/PB) of, for example, the central nervous system, spleen, liver, and testes may also be present. Based on which lineage of the normal hematopoiesis the abnormal blood cells resemble morphologically and immunophenotypically, the leukemias can be dichotomized into myeloid or lymphoid malignancies. Furthermore, leukemias can be “acute” or “chronic”, historically referring to a short or long anamnesis of symptoms and (without modern treatment) short or long survival. Acute myeloid leukemia (AML) and ALL are characterized by a rapid expansion of immature cells, i.e., blasts, whereas chronic myeloid leukemia (CML) and chronic lymphocytic
leukemia (CLL) are associated with a slow accumulation of more mature cells. Current treatment protocols also include clinical findings, e.g., age and WBC count, and the presence of certain leukemia-associated genetic aberrations, such as translocations/fusion genes and some gene mutations, in the malignant cells to subgroup the cases into relevant diagnostic and prognostic entities. Together with polychemotherapy and, in some instances, allogeneic stem cell transplantation such risk stratification has greatly contributed to the increased likelihood of remission and long-term survival of acute leukemias during the last few decades (Grimwade et al, 2010; Harrison et al, 2010; Radhi et al, 2010; Schmiegelow et al, 2010; Pui, 2012; Estey, 2013; Inaba et al, 2013; Ofran & Rowe, 2013). The goal of my PhD study has been to improve further the genetic risk stratification of pediatric BCP ALL.

Pediatric BCP ALL

Epidemiology

Approximately 1000 cases of leukemia are diagnosed annually in Sweden (www.cancerfonden.se). The vast majority, in particular AML, CLL, and CML, mainly occurs in adults, whereas ALL predominates in children/adolescents, representing 25% of the 300 pediatric cancer cases diagnosed each year in Sweden and 85% of all pediatric acute leukemias, of which BCP ALL is the most common subtype (Swerdlow et al, 2008; Pui, 2012; www.barcancerfonden.se). Because identification of novel risk-stratifying genetic changes in childhood BCP ALL is the focus of my thesis work, this leukemia type is discussed further below.

The relatively high prevalence of ALL in children compared with adults, in whom ALL constitutes only 2-3% of all hematologic malignancies which in turn only represent 5-8% of all adult malignant disorders (Downing & Shannon, 2002), was first recognized in 1917 through an epidemiological survey of 1457 children with leukemia (Ward, 1917). The incidence of pediatric BCP ALL in Sweden is 3.59/100 000 children/year (Hjalgrim et al, 2003); this is one of the highest incidences in the world. As a comparison, the global incidence is 3.0/100 000/year and in the US it is only 2.7/100 000/year (McNeil et al, 2002; Howard et al, 2008; Grigoropoulos et al, 2013). However, despite this variation in incidence, BCP ALL is nevertheless the major pediatric malignancy in all developed countries (Bunin et al, 1996; Linet et al, 1999). Although BCP ALL affects children /adolescents of all ages, there is a pronounced incidence peak at 2-5 years (Hjalgrim et al, 2003; Forestier & Schmiegelow, 2006; Pui et al, 2008). The sex ratio is approximately 1.0; thus, girls and boys are at equal risk of developing BCP ALL (Vaitkevičienė et al, 2011).
Etiology

The etiology of childhood BCP ALL is multifactorial and still to a large extent poorly understood. Therefore, one challenge today is to identify external agents and internal genetic variants that contribute to, and that may predict, the occurrence of BCP ALL. Indeed, some factors have been associated with increased risk of BCP ALL, including a few syndromes, inherited polymorphisms and mutations, and environmental exposures (Belson et al., 2007; Pui, 2012; Inaba et al., 2013). Because there is ample and convincing evidence for an in utero origin of many pediatric BCP ALL cases (Gale et al., 1997), pregnancy-related risk factors may be etiologically more important for childhood BCP ALL than lifestyle-related factors (which are clearly more relevant for adult cancer). Although many types of prenatal exposure, such as radiation, smoking, and alcohol consumption, as well as high maternal age and high birth weight have been associated with increased risk of BCP ALL, the estimates have often been imprecise, with several associations not having been confirmed in subsequent studies; these “risk factors” hence remain controversial (Hjalgrim et al., 2004; Greaves, 2006; Pui, 2012; Inaba et al., 2013; Kendal et al., 2013). However, it has been proposed that the above-mentioned factors nevertheless play a role in utero, leading to mutations that are important for initiation but not sufficient for overt leukemia; additional changes are required (Greaves, 2006). A possibly causal factor (“second hit”) is an aberrant immune response to infection(s) or an abnormal pattern of infection due to a high degree of “population mixing” (Kinlen & Doll, 2004). Other studies, however, have shown the opposite, i.e., that exposure to common infections actually may decrease the risk of ALL and that delayed exposure, due to geographic or social isolation resulting in low population mixing, actually increases the probability of second hits (Gilham et al., 2005; Greaves, 2006; Schmiegelow et al., 2008; Marcotte et al., 2014).

In general, the genetic aberrations that give rise to ALL are acquired, i.e., they are only present in the leukemic cells. Inherited genetic susceptibility has, until quite recently, been considered of less importance, not least because studies have failed to detect highly penetrant germline mutations and because familial aggregations of ALL are rare (Inaba et al., 2013; Urayama et al., 2013). Although monozygotic twins often may be concordant for ALL, suggesting heritability, this concordance is due to “metastasis” of leukemic (or pre-leukemic) cells from one twin to the other via the shared placenta rather than to inherited predisposition (Greaves et al., 2003). However, recurrent cases of ALL within families have been known for a long time (Anderson, 1955; Gunz et al., 1978) and, furthermore, advances in genome-wide profiling methods, such as single nucleotide polymorphism (SNP) arrays and next generation sequencing (described further in Articles I-III), have made it possible to identify low-penetrance predisposition alleles that contribute to leukemogenesis (Christensen & Murray, 2007). Thus, genome-wide association studies have shown that certain constitutional SNPs in
ARID5B, CDKN2A, CEBPE, GSTM1, and IKZF1 (Figure 1), resulting in absent or altered mRNA levels of these genes, are associated with increased risk of pediatric BCP ALL (Papaemmanuil et al, 2009; Treviño et al, 2009; Prasad et al, 2010; Sherborne et al, 2010; Yang et al, 2010; Ross et al, 2013; Xu et al, 2013; Ma et al, 2014). In addition, some genetic disorders, such as Down syndrome, Bloom syndrome, and ataxia telangiectasia, are clearly associated with increased risk of ALL (Krivit & Good, 1957; German et al, 1979; Hecht & Hecht, 1990). More recently, germline mutations of TP53 (causing Li-Fraumeni syndrome characterized by a predisposition to a wide variety of malignant disorders, including adrenal cortical carcinoma, breast cancer, osteosarcoma, and soft tissue sarcomas), PAX5, and SH2B3 have also been associated with increased risk of BCP ALL (Holmfeldt et al, 2013; Perez-Garcia et al, 2013; Powell et al, 2013; Shah et al, 2013). Still, despite these examples of genetic predisposition, it should be emphasized that they account for only a small fraction of cases and that they are not prevalent or predictive enough to justify (ethically or clinically) presymptomatic screening of children (Kharazmi et al, 2012; Pui, 2012; Inaba et al, 2013; Urayama et al, 2013).

Figure 1. A schematic model depicting the role of genetic alterations in the pathogenesis of pediatric BCP ALL. The initiating event, commonly a chromosomal rearrangement that occasionally may be acquired in utero, contributes in association with cooperative genetic aberrations to an inappropriate proliferation of immature cells. The clonal heterogeneity seen in diagnostic ALL samples is illustrated by the black and white cores in the leukemic cells, indicating that some genetic alterations, only present in minor subclones (white core) or acquired during treatment, may confer resistance to therapy and promote relapse. Reprinted from Seminars in Hematology, Volume 50, Mullighan, Genomic characterization of childhood acute lymphoblastic leukemia, 314-324, 2013, with permission from Elsevier.
Diagnosis and Therapy

To diagnose ALL, a BM sample is essential because as many as 10% of the patients lack circulating leukemic cells in the PB (Gajjar *et al.*, 1995). The diagnosis is based on the morphology of the PB and BM blasts under the microscope and, in particular and most importantly, on their immunophenotypic features. By the use of different monoclonal antibodies detecting various cell surface and cytoplasmic antigens, so called CD markers, ALL can be broadly subdivided into BCP ALL (approximately 85% of all pediatric ALL cases), T-cell ALL (15%), and mature B-cell or biphenotypic ALL (together accounting for less than 2%). This classification is vital for proper treatment assignment (Swerdlow *et al.*, 2008; Pui, 2012).

Ever since the discovery in the late 1940s of different chemotherapeutic agents active against BCP ALL, the treatment outcome has gradually improved, from being a uniformly fatal disease before 1950 to the current overall survival (OS) rates of approximately 90% (Farber & Diamond, 1948; Schmiegelow *et al.*, 2010; Hunger *et al.*, 2012; Pui & Evans, 2013; Vora *et al.*, 2013). This dramatic increase in OS is mainly due to the development and introduction of polychemotherapy, i.e., the use of several cytostatic drugs acting in concert against the leukemic blasts. Improved understanding of the biology of ALL and the realization that certain clinical and genetic features affect treatment response and thus can be used for risk stratification (see section “Risk-Stratifying Abnormalities” below), which influences the treatment intensity, has further added to this success story. Furthermore, “personalized medicine dosing” of some of the chemotherapeutic agents given, taking into account individual differences in metabolic pathways affecting drug response (pharmacogenetics), and improved supportive care have also contributed greatly to the increased OS (Krynetski *et al.*, 1995; Evans *et al.*, 1998; Gustafsson *et al.*, 1998; Anderer *et al.*, 2000; Rocha *et al.*, 2005; Schultz *et al.*, 2007; Davies *et al.*, 2008; Pui *et al.*, 2008; Schmiegelow *et al.*, 2010). In parallel with the improvements in risk stratification based on diagnostic findings, it has also become apparent that treatment response, as determined by quantification, by the use of polymerase chain reaction (PCR) or flow cytometry analyses, of remaining leukemic cells (minimal residual disease, MRD), at certain time points during therapy is essential for proper modification of the subsequent treatment given. In fact, MRD is considered the most powerful indicator for outcome today (van Dongen *et al.*, 1998; van der Velden *et al.*, 2007; Björklund *et al.*, 2009; Eckert *et al.*, 2013; Toft *et al.*, 2013; Vora *et al.*, 2013; Article II).

Despite all the above-mentioned advances, a significant proportion of the patients still relapses. Furthermore, the toxic burden of the aggressive treatment is high, resulting in unfavorable short-term, in the worst case death of the child, and long-term side effects, such as cardiac dysfunction, cognitive impairment, decreased fertility as well as increased risk of secondary malignancies and
premature death (Green et al., 2009; Mulrooney et al., 2009; Olsen et al., 2009; Stanulla & Schrappe, 2009; Pui et al., 2011; Garwicz et al., 2012).

Thus, in order to identify the leukemia- and host-related mechanisms underlying differences in treatment response and outcome, to develop targeted therapies, and to accomplish the ultimate goal of curing all children and adolescents with ALL, further studies are needed. One of several possible avenues to follow is to identify novel risk-stratifying genetic changes that also may be “druggable targets”. This is the path I have been following in my PhD study.

Cytogenetic Aberrations

Like cancer in general, BCP ALL arises as a direct result of mutations that alter the genetic control systems and that lead to uncontrolled self-renewal and proliferation, differentiation block, and decreased apoptosis of the leukemic cells (Hanahan & Weinberg, 2011). Since the initial cytogenetic studies of ALL cases in the late 1950s, revealing that they harbor acquired chromosomal abnormalities (Ford et al., 1958), a great number of numerical and structural chromosomal changes in BCP ALL has been shown to provide diagnostically and/or prognostically important information; in fact, several aberrations are today used as risk-stratifying factors (Figure 2) and are hence essential to detect at diagnosis (Forestier et al., 2000b; Johansson et al., 2004; Schultz et al., 2007; Harrison, 2009; Moorman et al., 2010).

Numerical Abnormalities

Losses and, in particular, gains of whole chromosomes are common in BCP ALL (Mitelman et al., 2014). The numerical changes may be single abnormalities, such as trisomy 5 (Harris et al., 2004) and trisomy 21 (Karrman et al., 2006), secondary changes to well-known primary translocations (Johansson et al., 1996), or comprise characteristic ploidy groups.

Pediatric BCP ALL can be divided, based on modal chromosome numbers, into the following subgroups: near-haploidy (25-29 chromosomes), low hypodiploidy (31-39), high hypodiploidy/hypodiploidy (40-44/<45), diploidy/pseudodiploidy (46), low hyperdiploidy (47-50), high hyperdiploidy (HeH; 51-67), near-triplody (66-79), and near-tetraploidy (84-100) (Harrison & Johansson, 2014).
Figure 2. Cytogenetic subgroups, as ascertained by conventional chromosome banding and targeted molecular genetic analyses, of pediatric BCP ALL. These subtypes are characterized by genetic alterations that have been shown to be relevant risk-stratifying factors associated with diverse immunophenotypic and clinical features. *Other translocations comprise rearrangements detected in low frequency and that are currently not used as risk-stratifying abnormalities.

*HeH*

HeH is the most common cytogenetic subgroup in pediatric BCP ALL, comprising 25% of all cases (Figure 2), and is characterized by favorable clinical characteristics, such as age 2–5 years and low WBC counts (Paulsson & Johansson, 2009). Although associated with a good prognosis, HeH-positive cases are not by default grouped as standard risk (SR); risk stratification is used to identify non-SR patients that need more treatment. Indeed, only approximately 50% of the patients are grouped as SR based on clinical features (Paulsson et al., 2013). The gained chromosomes in HeH are clearly nonrandom, with trisomies of chromosomes X, 4, 6, 10, 14, 17, and 18 and tetrasomy 21 being characteristic (Paulsson & Johansson, 2009; Paulsson et al., 2010, 2013). It is presently unknown how the trisomies and tetrasomies contribute to the leukemic process. Most likely, the pathogenetic impact may be ascribed to dosage effects of certain gene loci on the gained chromosomes (Gruszka-Westwood et al., 2004; Paulsson & Johansson, 2009). There is direct as well as indirect evidence in support for the HeH pattern (at least in many instances) arising before birth through a simultaneous gain of multiple chromosomes in a single abnormal division (Maia et al., 2004; Paulsson et al., 2005). Thus, HeH is not sufficient for leukemogenesis. It has been shown that approximately 80% of the cases harbor additional structural genetic aberrations; these secondary genetic events undoubtedly play an important role for overt...
leukemia to arise (Paulsson et al., 2008, 2010). Recently, a high frequency of subclones in HeH-positive BCP ALL has been reported, suggesting an underlying chromosomal instability (Alpár et al., 2014). Furthermore, two recent studies (Hussin et al., 2013; Woodward et al., 2014) have reported that rare allelic variants of the PRDM9 gene, which encodes a major regulator of meiotic recombination, are associated with HeH and that they may be inherited predisposing factors promoting chromosomal instability.

**Hypodiploidy**

Near-haploid and low hypodiploid cases, constituting approximately 1% of pediatric BCP ALLs (Figure 2; Article I), are characterized by multiple monosomies and are associated with a very poor outcome (Raimondi et al., 2003; Harrison et al., 2004; Nachman et al., 2007). In many near-haploid cases, the chromosome complement is duplicated, resulting in a doubled clone that can be mistaken for HeH in cytogenetic analyses. However, instead of trisomies, these “HeH” clones mainly harbor disomies and tetrasomies, where all the disomies are uniparental isodisomies (UPIDs), i.e., both copies of the chromosome are derived from only one of the parents (Karow et al., 2007; Carroll et al., 2009; Mitelman et al., 2014); such UPIDs are easily identified by SNP array analyses (Article I). Similarly, the near-triploid group, which is quite infrequent in pediatric BCP ALL, most often, or perhaps always, represents a duplication of a hidden low hypodiploid clone and should hence be grouped together with such cases (Charrin et al., 2004). As for the HeH cases, the pathogenetic impact of the massive chromosome losses in the near-haploid and low hypodiploid cases is probably related to gene dosage effects. In contrast to HeH, near-haploid cases very rarely harbor additional structural aberrations, whereas low hypodiploid cases often carry such changes (Safavi et al., 2013). However, one recent study, using massively parallel deep sequencing, identified multiple gene mutations in both near-haploid and low hypodiploid cases (Holmfeldt et al., 2013); such mutations in BCP ALL are discussed further below (see section *Sequence Mutations*).

High hypodiploid cases constitute a heterogeneous group, with some studies separating cases with 40-44 chromosomes from those with 45 chromosomes (Raimondi et al., 2003; Harrison et al., 2004; Nachman et al., 2007). Many high hypodiploid cases carry well-known BCP ALL-associated abnormalities, such as dic(9;20)(p13;q11), t(9;22)(q34;q11), and t(12;21)(p13;q22), and should be grouped together with these specific changes (see section *Structural Abnormalities* below) rather than according to modal chromosome number; this is also the case for the pseudodiploid, near-tetraploid, and low hyperdiploid BCP ALLs (Attarbaschi et al., 2006; Harrison & Johansson, 2014).
Structural Abnormalities

Structural aberrations in pediatric BCP ALL, as ascertained by chromosome banding analyses, are manifold, heterogeneous, and comprise both unbalanced and balanced abnormalities. Frequent large unbalanced changes include duplications of chromosome arm 1q [dup(1q)] and deletions involving 6q [del(6q)], 9p [del(9p)], and 13q [del(13q)]. Of these, del(9p) is known to target the cell cycle-related gene CDKN2A, resulting in hemi- or homozygous loss of this gene (see section Copy Number Alterations and Uniparental Isodisomies below); the molecular genetic consequences of the other changes are unknown. Furthermore, these changes are mostly regarded as secondary events, with dup(1q) being associated with HeH, del(6q) with t(12;21)(p13;q22), and del(9p) and del(13q) being detected across all cytogenetic subtypes (Davidsson et al, 2007; Paulsson & Johansson, 2009; Lilljebjörn et al, 2010; Moorman et al, 2010; Harrison & Johansson, 2014).

Balanced changes are rearrangements within or between chromosomes that are not associated with any net gain or loss of chromosomal material, such as translocations and inversions. These rearrangements may result in chimeric genes with novel functions, deregulated gene expression because of promoter swapping, and, less commonly, in loss of function (Rabbitts & Stocks, 2003; Mitelman et al, 2007; Wang, 2012; Das & Tan, 2013; Byrne et al. 2014).

t(12;21)(p13;q22)

Most translocations in BCP ALL result in gene fusions. One prime example is the ETV6-RUNX1 chimera – the most frequent fusion gene in pediatric BPC ALL – found in 25% of cases (Figure 2). It is generated through the translocation t(12;21)(p13;q22) (Golub et al, 1995; Swerdlow et al, 2008; Linka et al, 2013). The t(12;21) cannot be detected using conventional chromosome banding techniques because the translocation rearranges two similarly sized and banded terminal chromosomal regions; therefore, fluorescence in situ hybridization (FISH) or reverse-transcriptase PCR analyses are required for its detection (Harrison et al, 2005; Forestier et al, 2007). How the encoded fusion protein is involved in leukemogenesis is not fully understood. However, both ETV6 and RUNX1 are transcription factors of importance in normal hematopoiesis and it has been suggested that the main result of the fusion is repression of the normal activity of RUNX1 (Hiebert et al, 1996; Zelent et al, 2004; Palmi et al, 2014). As for HeH, t(12;21) often arises prenatally (Mori et al, 2002; Greaves & Wiemels, 2003) (Figure 1). Thus, the ETV6-RUNX1 fusion is not sufficient for overt leukemia; additional changes are needed. Frequent secondary aberrations include del(6q) and trisomy 16 and 21 as well as submicroscopic deletions of the wild type (wt) ETV6, CDKN2A/B, PAX5, and TBL1XR1; some of the latter deletions may be caused by an aberrant recombination activating gene (RAG) activity (Attarbaschi et al, 2004; Parker et al, 2008; Lilljebjörn et al, 2010; Papaemmanuil et al, 2014).
The t(12;21) is associated with an excellent outcome; however, some studies have indicated a relatively high risk of late relapses (Forestier et al., 2007, 2008b; Moorman et al., 2010; Pui et al., 2011).

t(1;19)(q23;p13)

The balanced t(1;19)(q23;p13) and its unbalanced counterpart der(19)t(1;19)(q23;p13) are present in 5-6% of pediatric BCP ALL (Figure 2). The translocation fuses the transcription factor-encoding TCF3 gene and the homeobox gene PBX1, giving rise to a potent chimeric oncoprotein (Smith et al., 1997; Paulsson et al., 2007; Schmiegelow et al., 2010; Uckun et al., 2013b). In a review of pediatric BCP ALL cases with t(1;19)/der(19)t(1;19) from the Nordic countries, the median age was 7 years and the median WBC count was 16 x 10^9/l, both slightly higher than in HeH- and t(12;21)-positive cases. Furthermore, the OS was worse in patients with t(1;19)/der(19)t(1;19) compared with those with t(12;21) or HeH (Andersen et al., 2011). However, clinical trials optimizing the treatment have effectively reversed the poor prognostic impact previously afforded by this translocation. Thus, it is important to screen for the TCF3-PBX1 fusion for proper risk stratification (Jeha et al., 2009; Schmiegelow et al., 2010). Variant 19p13/TCF3 rearrangements have also been detected, such as the t(17;19)(q22;p13)[TCF3-HLF] that is associated with a dismal prognosis (Inukai et al., 2007; Pui et al., 2008; Moorman, 2012).

dic(9;20)(p13;q11)

Another relatively common structural change in pediatric BCP ALL is the dicentric abnormality dic(9;20)(p13;q11), shown by interphase FISH analyses to be present in up to 5% of cases (Zachariadis et al., 2011). Analyses of dic(9;20)-positive BCP ALLs in the Nordic countries have revealed that they have a pronounced age peak at three years, similar to HeH and t(12;21), that girls are clearly overrepresented, and that the outcome, at least when treated as SR, is suboptimal (Schoumans et al., 2006; Forestier et al., 2008a; Zachariadis et al., 2011). The dic(9;20) results in loss of 9p and, most often, also 20q; although the breakpoints are clustered they are heterogeneous and the functional outcome is unclear. However, the PAX5 gene at 9p13.2 has been shown to be rearranged in several cases. Thus, haploinsufficiency of PAX5 may be the main functional outcome, together with homozygous loss of CDKN2A, of this dicentric translocation (An et al., 2008; Kawamata et al., 2008b; Pichler et al., 2010; Zachariadis et al., 2012).

t(9;22)(q34;q11)

The Philadelphia chromosome, representing one of the derivatives, namely der(22)t(9;22)(q34;q11), of the t(9;22)(q34;q11), is characteristic of CML and is probably the most studied chromosome rearrangement in human neoplasia (Nowell & Hungerford, 1960; Rowley, 1973; Szczepański et al., 2010). Although
strongly associated with CML, it has been known since 1970 that the t(9;22) also occurs in ALL (Propp & Lizzi, 1970). The breakpoints in chromosome 9 are clustered within intron 1 of the \textit{ABL1} gene whereas the breakpoints in chromosome 22 are more heterogeneous within the \textit{BCR} gene, giving rise to differently sized \textit{BCR-ABL1} fusion products, most commonly p190 and p210. The variant transcripts all encode a chimeric tyrosine kinase with enhanced activity that has been shown to transform hematopoietic cells both \textit{in vitro} and in mice (Heisterkamp et al, 1985; Chan et al, 1987; van Etten et al, 1989; Lugo et al, 1990; Cortez et al, 1997; van Etten, 2002; Quintás-Cardama & Cortes J, 2009). This oncoprotein is found in 2-3\% (Figure 2) of pediatric BCP ALL and is associated with older age (the median age in children is 7-10 years) and somewhat increased WBC counts (median ~30 x 10^9/l) at diagnosis (Crist et al, 1990; Uckun et al, 1998; Aricò et al, 2010; Schultz et al, 2014). The t(9;22) is much more prevalent in adult BCP ALL, occurring in 25\% cases (Pui & Evans, 2006). Both CML and BCP ALL with t(9;22) used to have a dire outcome. However, the development of specific tyrosine kinase inhibitors, initially imatinib (Gleevec) (Druker, 2008) has changed the prognosis dramatically, with recent studies having shown that the survival rates of pediatric t(9;22)-positive BCP ALL have more than doubled compared with historical controls (Rives et al, 2011; Schultz et al, 2014).

\textbf{11q23/MLL rearrangements}

Chromosome abnormalities involving 11q23 that result in \textit{MLL} gene rearrangements are particularly common in infant ALL cases, seen in 75-80\% of such cases, but are also relatively frequent in older children; in total, approximately 5-6\% (Figure 2) of all pediatric BCP ALL cases harbor \textit{MLL} rearrangements (Rubnitz et al, 1994; Forestier et al, 2000a; Pieters et al, 2007). The \textit{MLL} gene has so far been shown to fuse to 79 different partner genes in BCP ALL (Meyer et al, 2013a) with the most common translocations/fusion genes being t(4;11)(q21;q23) [\textit{MLL-AFF1}] and t(11;19)(q23;p13.3) [\textit{MLL-MLLT1}]. The normal function of MLL is to regulate the transcription of \textit{HOX} genes during hematopoiesis. These genes code for critical histone regulators, and studies of MLL chimeras have shown inappropriate expression of \textit{HOX} genes, with the MLL fusion partner contributing either through transactivation or by dimerization of MLL (Rubnitz et al, 1994; Armstrong et al, 2002; Andersson et al, 2005; Meyer et al, 2009; Marschalek, 2011, Wächter et al, 2014).

\textbf{iAMP21}

A relatively recent addition to the plethora of structural changes in BCP ALL is the intrachromosomal amplification of chromosome 21 (iAMP21). This abnormality is present in \~2\% of cases (Figure 2). The structurally abnormal chromosome 21 in iAMP21-positive cases is morphologically extremely heterogeneous and may well be misinterpreted as a marker and missed by chromosome 22
banding analysis. Hence, FISH or SNP array analyses are needed to identify iAMP21 with certainty (Moorman et al, 2010; Heerema et al, 2013, Li et al, 2014; Article I). This cytogenetic subgroup was first described in 2003 in a FISH study of RUNX1 (at 21q22.12) in BCP ALL (Soulier et al, 2003). Although cases with iAMP21 always harbor amplification of RUNX1, it should be emphasized that it has not been shown that this is the pathogenetically important gene. In fact, SNP array analyses reveal a complex rearrangement pattern on chromosome 21, with both amplifications and deletions throughout chromosome arm 21q (Figure 3). Several studies have provided evidence that iAMP21 is an initiating event, mutually exclusive of other “primary” BCP ALL-associated abnormalities. Additional, secondary changes are common, including gain of chromosome X, loss or deletion of chromosome 7, and ETV6 and RB1 deletions (Harrison et al, 2014). The iAMP21 has been shown to arise through a breakage-fusion-bridge mechanism coupled with chromothripsis (Rand et al, 2011; Sinclair et al, 2011; Li et al, 2014). Interestingly, children with the rare constitutional Robertsonian translocation der(15;21)(q10;q10) have a pronounced (2,700-fold increased!) risk of developing iAMP21-positive BCP ALL (Li et al, 2014). Patients with iAMP21 have been shown to be older (median 9 years versus 5 years) and have a lower WBC count (median 3.9 x 10^9/l versus 12.4 x 10^9/l) compared with patients without this abnormality (Soulier et al, 2003; Heerema et al, 2013; Moorman et al, 2013). A dismal outcome, at least after SR treatment, has been emphasized (Moorman et al, 2010).

Figure 3. SNP array image showing intrachromosomal amplification of chromosome 21 (iAMP21). The two top panels display B allele frequencies (genotypes) and log2 ratios (copy numbers), respectively, at the time of diagnosis of a pediatric iAMP21-positive BCP ALL (included in Article I). The two lower panels show the same for the remission sample. The panels were extracted from the genome studio v2011.1 software (Illumina). Chromosome 21 is schematically depicted below the panels. The iAMP21 is a complex rearrangement involving both amplifications and deletions throughout 21q. The remission sample confirms that the iAMP21 is a somatic change, only present in the leukemic cells.
IGH Translocations

Apart from translocations leading to gene chimeras in BCP ALL, some result in deregulated gene expression because of illegitimate recombination between one of the immunoglobulin genes (IGH at 14q32, IGK at 2p11, and IGL at 22q11) and an “oncogene”. Such translocations were previously only thought to be present in mature B-cell ALL (Burkitt lymphoma/leukemia), resulting in deregulation of the MYC gene at 8q24.21, and in T-cell ALL, where translocations involving the T-cell receptor loci and various partner genes are common (Staal et al., 2007; Boerma et al., 2009; Karrman et al., 2009; Larmonie et al., 2013). However, in recent years, an increasing number of IGH translocations has been identified also in pediatric BCP ALL (Russel et al., 2014). One relatively common group of IGH translocations comprises those that target various members of the CEBP gene family, such as t(8;14)(q11;q32) [IGH-CEBPD] and t(14;19)(q32;q13) [IGH-CEBPA]. Patients with IGH-CEBP-positive BCP ALL do not differ clinically compared with those without such translocations and CEBP deregulation does not seem be an independent risk factor; however, it has recently been reported that older patients (>10 years) with IGH translocations have a worse outcome (Akasaka et al., 2007; Moorman et al., 2012; Russel et al., 2014).

Another recurrent type of IGH translocation is the t(X;14)(p22;q32)/t(Y;14)(p11;q32). The t(X;14) and t(Y;14), which result in juxtaposition of the IGH and CRLF2 genes in the pseudoautosomal region 1 (PAR1) on Xp22/Yp11, have been shown to be frequent in BCP ALL in Down syndrome patients; these translocations also occur, but more rarely, in other pediatric BCP ALL cases (Mullighan et al., 2009a; Article III). The rearrangements result in deregulated expression of the CRLF2 gene, encoding a cytokine receptor-like factor (Russell et al., 2009a); the CRLF2 gene is further discussed in the section “Gene Expression Profiling” below. Other translocations involving the IGH locus in BCP ALL include t(5;14)(q31;q32) [IGH-IL3; associated with eosinophilia], t(6;14)(p22;q32) [IGH-ID4], and t(14;19)(q32;p13) [IGH-EPOR] (Tono-oka et al., 1984; Grimaldi & Meeker, 1989; Russell et al., 2008, 2009b; George et al., 2012).

Rare Translocations

There are several recurrent although rare translocations in BCP ALL, such as dic(9;12)(p13;p13) [PAX5-ETV6] and dic(7;9)(p11-13;p11-13) [unknown molecular genetic consequences]; their prognostic impact is currently unclear because of their rarity (Figure 2) (Strehl et al., 2003; Lundin et al., 2007; Nebral et al., 2009). With the advent of massively parallel deep sequencing, including transcriptome (mRNA) sequencing, several novel, as well as clinically/therapeutically relevant, fusion genes have also been detected in BCP ALL, such as NUP214-ABL1 [t(9;9)(q34;q34)/del(9)(q34q34); previously only reported in T-cell ALL], EBF1-PDGFRB [del(5)(q32q33.3)], and MEF2D-CSF1R [t(1;5)(q21;q33)]; patients with these gene fusions have been reported to respond to tyrosine kinase inhibitors and are hence important to identify (De Braekeleer et al., 2014).

In conclusion, approximately 70% of pediatric BCP ALL cases can now be categorized based on specific cytogenetic alterations associated with characteristic immunophenotypic and clinical features (Figure 2); indeed, some changes – HeH, ETV6-RUNX1, and IGH-IL3 – delineate specific BCP ALL subtypes in the current WHO classification (Forestier et al, 2000b; Johansson et al, 2004; Schultz et al, 2007; Swerdlow et al, 2008; Harrison et al, 2010; Moorman et al, 2010, 2012; Schmiegelow et al, 2010; Pui et al, 2011; Toft et al, 2013; Harrison & Johansson, 2014). The challenge now is to address the remaining 30% with uninformative cytogenetics, i.e., non-characteristic abnormalities, normal karyotype, and karyotypic failure; a group I refer to as “other” in this thesis and in Articles I-III.

Risk-Stratifying Abnormalities

In the current Nordic Society of Pediatric Hematology and Oncology (NOPHO) ALL-2008 treatment protocol, some of the above-mentioned genetic abnormalities are used to stratify the patients into non-SR treatment, i.e., intermediate risk (IR) or high risk (HR), at diagnosis (Schmiegelow et al, 2010; Frandsen et al, 2014).

If any of the following three abnormalities are identified at diagnosis, the patient will initially be grouped as IR (but can be regrouped based on the MRD findings later on): t(1;19)/der(19)t(1;19), dic(9;20), and iAMP21. Cases grouped as HR based on cytogenetics comprise those with hypodiploidy (<45 chromosomes) or 11q23/MLL rearrangements. Patients with Philadelphia chromosome-positive BCP ALL are not treated according to the NOPHO ALL-2008 protocol; instead, they receive therapy, including tyrosine kinase inhibitors, according to the European Study on Philadelphia-Positive Acute Lymphoblastic Leukemia (Schmiegelow et al, 2010; Biondi et al, 2012; van der Veer et al, 2014).

Submicroscopic Genetic Changes

In addition to the cytogenetically detectable abnormalities described above, many of which (at least the balanced ones) are considered initiating and primary events, driving leukemogenesis, numerous submicroscopic copy number alterations (CNAs) and sequence mutations have been identified in BCP ALL by the use of SNP array and massively parallel sequencing analyses; these additional changes undoubtedly cooperate in the leukemogenic process and may well be required for the development of overt BCP ALL (Figures 2 and 4). These studies have expanded our knowledge about the types and patterns of genomic change that
contribute to leukemia and have resulted in the identification of new risk-stratifying factors and therapeutic targets. In addition, they have led to the discovery of novel subtypes in BCP ALL, previously lumped together in the quite large and clinically very heterogeneous cytogenetic group “other” (Figure 4). Thus, almost all pediatric BCP ALL cases can now, or at least in the near future, be classified genetically (Kuiper et al., 2007; Mullighan et al., 2007; Kawamata et al., 2008a; Harrison et al., 2009; Pui et al., 2011, 2012; Zhang et al., 2011; Downing et al., 2012).

Figure 4. Genetic subtypes and cooperative aberrations as ascertained by genome-wide profiling, extending the cytogenetic subgroups of BCP ALL. This is of course not a static model but rather an overview of our current knowledge. The distribution of the novel subtypes, denoted in red, is not exact — firm frequency data will be available first when next generation sequencing has become more widely implemented in clinical routine. It should also be stressed that the majority of the submicroscopic changes are observed across all cytogenetic subgroups. Hence, further studies are needed before the complete picture can be drawn.

**Copy Number Alterations and Uniparental Isodisomies**

SNP array analyses of pediatric BCP ALL have shown an average of approximately six CNAs per case, with virtually all regions of the genome having been implicated, and have revealed that microdeletions are clearly more common than amplifications/duplications (ratio 2:1) (Kuiper et al., 2007; Mullighan et al., 2007, 2008a; Kawamata et al., 2008a; Bungaro et al., 2009; Article I and unpublished data). Whole chromosome as well as segmental UPIDs have also
been shown to be frequent, occurring in 20-30% of cases across cytogenetic subgroups but being particularly common in near-haploid, low hypodiploid, and high hyperdiploid cases (Kawamata et al, 2008a, 2009; Mullighan et al, 2008a; Paulsson et al, 2010; Safavi et al, 2013; Article I). One frequent UPID, detected in many cancer types besides BCP ALL, involves 9p; such UPIDs are associated with prior heterozygous deletions of CDKN2A and hence result in a homozygous deletion of this gene (Pei et al, 2006; Safavi et al, 2013). The pathogenetic consequences, if any, of the other UPIDs in BCP ALL remain to be elucidated.

In total, more than 50 recurrent regions have been shown to be involved in CNAs. Deletions, often focal ones, of CDKN2A, ETV6, IKZF1, and PAX5 (Figure 1) that code for cell cycle regulators or transcription factors are especially frequent in pediatric BCP ALL (Kuiper et al, 2007; Mullighan et al, 2007; Kawamata et al, 2008a; Paulsson et al, 2010; Schwab et al, 2013; Moorman et al, 2014; Articles I-III). Some of these, for example, deletions of Pax5 and Ikzf1, have been shown to accelerate the onset of leukemia in mouse models, revealing that they are important cooperative changes in the leukemogenic process (Heltemes-Harris et al, 2008; Virely et al, 2010). Interestingly, the frequencies of CNAs vary among the different cytogenetic subgroups. For example, MLL-positive BCP ALLs have been reported to harbor less than one CNA per case (Bardini et al, 2011) strongly indicating that MLL fusions are potent oncogenes that hardly require any additional events to induce leukemia, whereas both the ETV6-RUNXI- and BCR-ABL1-positive subgroups have a higher incidence of CNAs: 7.5 and 8.8/per case based on Kuster et al (2011) and Mullighan et al (2008b), respectively. Also the patterns of CNAs vary to some extent. Although the majority of the most prevalent CNAs are observed in all cytogenetic subgroups, there are some notable exceptions (Figure 4): ETV6 deletions are overrepresented in cases with ETV6-RUNXI, CDKN2A is homozygously deleted in (almost) all dic(9;20)-positive cases, and IKZF1 is deleted in approximately 80% of BCR-ABL1-positive BCP ALL and in ~30% of cases cytogenetically grouped as “other” (Mullighan et al, 2007, 2008b; Kawamata et al, 2008a; Iacobucci et al, 2009; Sulong et al, 2009; Zachariadis et al, 2012; Dobbins et al, 2013; Schwab et al, 2013; Articles I-III). Some of the deletions may well confer a clinically important impact, either by themselves or in certain combinations. In fact, very recently, Moorman et al (2014) reported that deletions of, for example, ETV6, EBF1, and IKZF1 can be used to risk-stratify the majority of pediatric BCP ALLs.

Sequence Mutations

Since the introduction of massively parallel sequencing or “next generation sequencing”, as it was called in the beginning (it is really my generation’s sequencing method!), there has been a virtual explosion of studies identifying various types of mutation – small deletions and insertions (“indels”), missense
mutations, and nonsense mutations – in a wide variety of human malignant disorders (Stransky et al, 2011; Stratton, 2011; Downing et al, 2012; Seo et al, 2012; Zhang et al, 2012b; Hansen & Bedard, 2013; Ley, 2013; Vogelstein et al, 2013; Cowley et al, 2014). These comprehensive sequencing efforts have revealed that leukemias in general and pediatric BCP ALL in particular harbor relatively few sequence mutations, at least as compared with carcinomas and other solid tumors. To date, however, only eight studies applying whole genome/exome/transcriptome sequencing of pediatric BCP ALL have been reported and they have all focused on specific cytogenetic subgroups or on relapse samples rather than on the entire BCP ALL group (Lilljebjörn et al, 2012; Roberts et al, 2012, 2014a; Chen et al, 2013; Holmfeldt et al, 2013; Meyer et al, 2013b; Tzoneva et al, 2013; Papaemmanuil et al, 2014). Despite this, based on the results generated by these large-scale analyses as well as on findings from targeted sequencing, several recurrent somatic sequence mutations of genes encoding transcription factors, such as ETV6, IKZF1, and PAX5, and genes coding for proteins/kinases involved in cell signaling, for example, JAK1, JAK2, KRAS, NRAS, PTPN11, and SH2B3, have emerged (Figures 1 and 4). Thus, sequence mutations play an important role in the leukemogenic process and are hence important to detect in order to understand fully the genomic landscape of pediatric BCP ALL (Case et al, 2008; Paulsson et al, 2008; Mullighan et al, 2009c, 2011; Zhang et al, 2011; Iacobucci et al, 2012a; Loh et al, 2013; Woo et al, 2014; Article III).

Gene Expression Profiling

By combining data obtained from genetic analyses of deletions, sequence mutations, and gene rearrangements with those from gene expression profiling studies, a novel BCP ALL subgroup has emerged, namely cases associated with deregulated CRLF2 expression. This CRLF2-positive group comprises approximately 6% of pediatric BCP ALL (Figure 4). The underlying mechanisms for the aberrant CRLF2 activity vary but mainly involve translocation with IGH (described above) and deletions of PAR1 at Xp22/Yp11 that result in the P2RY8-CRLF2 fusion; more rarely, activation occurs through gain-of-function sequence mutations in the CRLF2 gene (Mullighan et al, 2009a; Russell et al, 2009; Chapiro et al, 2010; Ensor et al, 2011). Approximately 50% of these cases also harbor activating mutations in JAK1 or JAK2 and/or deletions of IKZF1 (∆IKZF1). The CRLF2-positive group has been associated with an inferior outcome in some, but not all, studies addressing this issue (Mullighan et al, 2009a,c; Cario et al, 2010; Harvey et al, 2010a; Ensor et al, 2011; Chen et al, 2012; Loh et al, 2013; van der Veer et al, 2013; Yamashita et al, 2013).

Global gene expression studies have also discovered novel subgroups of BCP ALL, such as the BCR-ABL1-like and the ERG-positive subgroups (Bungaro et al, 2009; Den Boer et al, 2009; Harvey et al, 2010b). BCR-ABL1-like cases, which
may comprise up to 10% of pediatric BCP ALL (Figure 4), are characterized by an expression profile similar to BCR-ABL1-positive cases, but lack this gene fusion, and are associated with a poor outcome (Den Boer et al, 2009; Harvey et al, 2010b; van der Veer et al, 2013; Roberts et al, 2014b). It has been shown that there is a high frequency of A|KZ|F1 in this subgroup and that BCR-ABL1-like cases often also harbor rearrangements of genes encoding tyrosine kinase or cytokine receptors, such as ABL1 (by default other than fusion to BCR), CRLF2, JAK2, and PDGFRB, resulting in improper activity of kinase signaling and suggesting that the use of kinase inhibitors may be a valid therapeutic strategy (Mullighan et al, 2009b; Harvey et al, 2010a; Maude et al, 2012; Roberts et al, 2012, 2014a; Weston et al, 2013).

BCP ALL-characteristic translocations have been associated with, and shown to define, characteristic gene expression patterns, whereas submicroscopic genetic changes in general do not correlate with such patterns (Andersson et al, 2007; Harvey et al, 2010b). However, BCP ALLs with deletions of the transcription factor-encoding gene ERG at 21q22.2, which do not (seem to) carry any known “primary” rearrangements, display a distinct gene expression profile. The ERG-positive group has variably been reported to comprise between 2 and 7% of pediatric BCP ALL (Figure 4) and has been associated with a superior outcome (Harvey et al, 2010b; Pui et al, 2011; Clappier et al, 2014; Zaliova et al, 2014; Article III).

**DNA Methylation Profiling**

In addition to genetic abnormalities, changes in DNA methylation have also been shown to play an important role in BCP ALL, with distinct signatures correlating with different cytogenetic subtypes (Gutierrez et al, 2003; Yang et al, 2006; Figueroa et al, 2013; Nordlund et al, 2013; Chatterton et al, 2014). Furthermore, there is a growing number of leukemia-associated genes that contributes to this epigenetic dysregulation, for example, the EZH2, histone gene cluster-1 (HIST1), HOX, MLL, and NSD2 genes. In fact, targeted deep sequencing of 472 known epigenetic regulators has revealed that approximately 25% of all pediatric BCP ALLs harbor mutations in such genes (Starkova et al, 2010; Loudin et al, 2011; Chung et al, 2012; Zhang et al, 2012a; Jaffe et al, 2013; Mar et al, 2014).

**Aberrant MicroRNAs**

During recent years, microRNAs (miRNAs; non-coding RNA), which negatively regulate the expression of genes by targeting their mRNA for cleavage or translational repression, have also been implicated in the development of BCP
ALL (Caldas & Brenton, 2005). Indeed, certain abnormal expression levels of different miRNAs, caused by deletions, amplifications, abnormal methylation patterns, or point mutations, have been suggested to be ALL subtype-specific (Zanette et al., 2007; Schotte et al., 2009, 2011; Duyu et al., 2014; Gutierrez-Camino et al., 2014). Recently, constitutional polymorphisms in some miRNAs, e.g., Has-miR-196a2, have also been associated with increased susceptibility to childhood ALL (Tong et al., 2014).

**Relapse and Treatment Failure**

Approximately 20% of all pediatric BCP ALL patients, across all cytogenetic subtypes and risk groups, relapse (Einsiedel et al., 2005; Bhojwani & Pui, 2013). The leukemic clones emerging at relapse are characterized by drug resistance (Klumper et al., 1995). This, to a major extent, explains why relapses are associated with inferior outcome; the cure rates after relapse are less than 40% (Ko et al., 2010; Tallen et al., 2010; van den Berg et al., 2011). In contrast to the clinical and genetic risk stratification performed at diagnosis, shown to increase treatment response and hence survival as discussed above, renewed stratification based on such features at the time of relapse has been of less value in several relapse trials (Schroeder et al., 1995; Einsiedel et al., 2005; Nguyen et al., 2008). Today, mainly time to relapse, site of relapse, and MRD are used in order to decide whether allogeneic stem cell transplantation should be performed or not (Bhojwani & Pui, 2013; Krentz et al., 2013).

**Genetic Evolution Patterns at Relapse**

The biological mechanisms that underlie and mediate the lack of responsiveness to presently available therapeutic agents are, in most instances, unknown. However, the resistance seems to be associated with acquisitions of new genetic changes in the leukemic clone, or rather subclones, over time; a process akin to the one “used by” bacteria to develop resistance to various antibiotics. Numerous studies have attempted to identify relapse-associated changes in pediatric ALL by genetically comparing paired diagnostic/relapse samples using techniques available at the time of study, such as conventional chromosome banding (Shikano et al., 1990; Heerema et al., 1992; Raimondi et al., 1993), FISH (Peham et al., 2004; Zuna et al., 2004), clone-specific PCR (Choi et al., 2007; Henderson et al., 2008), SNP array (Mullighan et al., 2008b; Yang et al., 2008; Kawamata et al., 2009; Davidsson et al., 2010; van Delft et al., 2011; Article I), and, most recently, massively parallel sequencing analyses (Mullighan et al., 2011; Meyer et al., 2013b; Tzoneva et al., 2013). Combined, these studies have identified
various clonal evolution patterns from the time of diagnosis to relapse. The relapse clone may 1) be genetically identical to the one seen at diagnosis; 2) harbor additional changes compared with the major clone or a subclone at diagnosis (“clonal evolution”); 3) have both additional and fewer changes compared with the diagnostic clone(s), corresponding to evolution from an ancestral/preleukemic clone; or 4) be a “new” (perhaps treatment-induced) leukemia genetically distinct from the first leukemia (Figure 5). In pediatric BCP ALL, the majority of relapse cases belong to group 2 or 3 (Mullighan et al, 2008b; Yang et al, 2008; Kawamata et al, 2009; Article I).

Figure 5. A schematic model depicting genomic heterogeneity and clonal relationships between diagnostic and relapse samples. Many cases of BCP ALL are genomically heterogeneous at diagnosis because of acquisition of different additional mutations in subclones from the time of initiation to overt leukemia; however, there is often one major clone, illustrated by a larger cell among the surrounding minor clones. The relapse clone may be genetically distinct from the one(s) found at diagnosis, corresponding to a “new” leukemia, or represent an evolution from a clone seen at diagnosis or from an ancestral/preleukemic clone. The figure is based on Figure 2 in Mullighan et al (2008b).

Genetic Aberrations Associated with Relapse

Apart from identifying the above-mentioned clonal evolution patterns at relapse, several studies have revealed specific genetic aberrations that are associated with, and contribute to, treatment failure (Klumper et al, 1995; Chessells et al, 2003; Einsiedel et al, 2005; Pui & Evans, 2006; Malempati et al,
2007; Mullighan et al., 2008b, 2011; Raetz et al., 2008; Yang et al., 2008; Kawamata et al., 2009; Möricke et al., 2008; Yang et al., 2008; Kawamata et al., 2010; Tallen et al., 2010; Meyer et al., 2013b; Article I). For example, genome-wide association and whole exome sequencing studies of high risk BCP ALL have revealed an allele variant (SNP) in the transcription factor-encoding gene GATA3 in BCR-ABL1-like cases and germline TP53 mutations in low hypodiploid cases that may contribute to the dismal outcome associated with such cases. Furthermore, SNPs in genes involved in pharmacogenetics, such as ASNS, FPGS, and PDE4B, have been shown to add significant information, independent of known risk factors, of prognostic importance. Such data provide insights into host genome variants that may be involved in relapse (Yang et al., 2009, 2012; Holmfeldt et al., 2013; Liu et al., 2013; Perez-Andreu et al., 2013; Walsh et al., 2013; Pastorczak et al., 2014; Wesołowska-Andersen et al., 2014).

SNP array analyses have also been used to detect somatic CNAs associated with relapse. For example, EBF1, IKZF1, RAG1, and TBL1XR1 deletions are more frequent in diagnostic samples from patients who subsequently relapse than in samples from those who remain in complete remission (Mullighan et al., 2008b, 2009c; Yang et al., 2008; Kawamata et al., 2009; Kuiper et al., 2010; Article I). In my SNP array studies, only ΔIKZF1 was consistently shown to have an independent impact on outcome, irrespective of MRD (Articles I and II and further discussed in the section “IKAROS and LEUKEMIA” below). In addition, some deletions have been shown either to only occur or to be enriched (outgrowth of a minor diagnostic subclone) at the time of relapse, such as those involving BTG1, CREBBP, MSH6, NR3C1, NT5C2, SETD2, and TP53 (Mullighan et al., 2008b, 2011a; Hof et al., 2011; Zang et al., 2011; Meyer et al., 2013b; Tzoneva et al., 2013; Mar et al., 2014; Article I). Deletions of MSH6, coding for a critical component of the DNA mismatch repair system, BTG1, encoding a transcription cofactor, and NR3C1, coding for a glucocorticoid receptor, have all been associated with drug insensitivity and glucocorticoid resistance, indicating that loss of function of these genes influence treatment response (Irving et al., 2005; Yang et al., 2008; van Galen et al., 2010; Kuster et al., 2011).

**CREBBP**

Hogan et al (2011) showed that there is a global hypermethylation of CpG islands in promoter regions of several genes, such as BGT4, CDKN2A, HIST3H2BB, and PAX9, at relapse of BCP ALL; a pattern distinct from the one observed at diagnosis. This strongly suggested that epigenetic deregulation contributed to treatment failure. Indeed, Mar et al (2014) recently reported that the majority of relapse samples harbors sequence mutations in epigenetic regulators, most frequently in the transcriptional coactivator/acetyltransferase gene CREBBP and in the methyltransferase gene SETD2. In fact, somatic mutations of the CREBBP gene have been reported in up to 20% of relapsed ALLs, especially in HeH-positive cases, and to always be preserved at relapse if present already at the
time of diagnosis (Mullighan et al., 2011; Inthal et al., 2012). In conclusion, the epigenetic changes incurred by the above-mentioned abnormalities are likely to be involved in clonal evolution and chemotherapy resistance.

**TP53**

TP53 is a well-known tumor suppressor gene that is frequently mutated or deleted in a wide variety of malignancies – solid tumors as well as hematologic disorders (Leroy et al., 2014). Although TP53 mutations are relatively rare in pediatric BCP ALL (Moorman et al., 2010), it has been reported that low hypodiploid BCP ALLs often carry acquired TP53 mutations; interestingly and importantly, germline mutations, corresponding to Li-Fraumeni syndrome, have been identified in some patients with low hypodiploid ALL (Holmfeldt et al., 2013; Mühlbacher et al., 2014). Furthermore, TP53 abnormalities have been shown to be enriched at relapse (Figure 1) and suggested to be a relevant prognostic factor at relapse (Hof et al., 2011; Krentz et al., 2013).

**NT5C2**

Acquired sequence mutations in the cytosolic 5'-nucleotidase II gene NT5C2, resulting in increased enzymatic activity with subsequent resistance to chemotherapy with 6-mercaptopurine and 6-thioguanine – drugs used in maintenance therapy of childhood ALL – were recently identified in pediatric BCP ALL. The NT5C2 mutations were mainly detected in cases cytogenetically grouped as “other” and seemed to be acquired at relapse (Meyer et al., 2013b; Tzoneva et al., 2013). However, by applying sensitive backtracking techniques, such as deep sequencing, it could be shown that the clones at relapse that harbored the NT5C2 mutations were present, albeit at very low frequencies (<1%), already at the time of diagnosis in some instances. This may well be true for all the relapse-associated aberrations mentioned above (Figure 5). If so, diagnostic clones with mutations that confer resistance to therapy would have a selective edge and would be the main reason for treatment failure (Irving et al., 2005; Mullighan et al., 2008b; Waanders et al., 2012; Paulsson, 2013).

**Gene Expression Signatures Associated with Relapse**

Gene expression profiling has also identified signatures that correlate with relapse, revealing, for example, upregulation of genes associated with proliferation, such as the BIRC5 gene coding for survivin. In addition, different gene expression classifier sets may predict early response to therapy and outcome, with aberrant expression of, e.g., EMP1, SMYD2, and CASP8AP2 being an independent risk factor for relapse (Yeoh et al., 2002; Staal et al., 2003; Beesley et al., 2005; Bhojwani et al., 2006, 2008; Flotho et al., 2006; Andersson et al., 2007; Kang et al., 2010; Li et al., 2013; Ariës et al., 2014; Sakamoto et al., 2014).
Moreover, changes in the expression of a few miRNAs, such as miR-16, miR335, miR-210, and Hsa-miR-124a, have been associated with drug resistance (Agirre et al., 2009; Kaddar et al., 2009; Yan et al., 2013; Mei et al., 2014).

**Therapeutic Targets**

The above-mentioned genetic and epigenetic aberrations associated with relapse and hence a dismal outcome are potential targets for novel and personalized therapies, at least for treatment of high risk ALL and resistant ALL. For example, CREBBP-positive cases may respond to treatment with epigenetic inhibitors (Bolden et al., 2006; Tsapis et al., 2007; Mullighan et al., 2011). Furthermore, there are ongoing phase 1 clinical trials of JAK inhibitors in BCP ALL with JAK abnormalities, such as those belonging to the BCR-ABL1-like subgroup (Maude et al., 2012; Pui et al., 2012; Roberts et al., 2012, 2014b; Tasian et al., 2012), and inhibitors of 5′-nucleotidases may be useful in NT5C2-mutated cases (Gallier et al., 2011; Jordheim et al., 2013; Meyer et al., 2013b).

The promises raised by new “targeting drugs” of relapsed ALL notwithstanding, it is still vital to identify the level of clonal heterogeneity, including coexisting and possibly cooperative genetic abnormalities seen by high-resolution genome-wide profiling methods, and to detect treatment-resistant subclones already at the time of diagnosis if we are to achieve the goal of preventing all treatment failures (Anderson et al., 2011; Greaves & Maley, 2012; Pui et al., 2012; Harrison, 2013; Ma et al., 2013; Paulsson, 2013).

**My PhD Studies in Brief**

The general aim of my thesis work has been to characterize genetically a large cohort of uniformly treated and diagnosed pediatric BCP ALL patients using SNP array and targeted deep sequencing analyses (described in Articles I-III) in order to identify relapse-associated genetic aberrations, novel changes, and cooperative events of biological and/or clinical importance and to further advance our understanding of leukemogenesis. When my PhD project started in 2010, only a few large-scale studies using these techniques had been performed. The SNP array study (Article I) revealed that somatic heterozygous (focal as well as non-focal) ΔIKZF1 is an important risk-stratifying factor, conferring a poor prognostic impact, especially in BCP ALL cases cytogenetically grouped as “other”. Thus, I suggest that analyses of IKZF1 provide clinically significant information and that the IKZF1 status should be implemented in the future NOPHO ALL treatment protocol. However, before implementing IKZF1 analyses in clinical routine, it is important to ascertain whether additional genetic changes and/or clinical findings
can modify the impact of ΔIKZF1, for better or worse. Hence, I validated the impact of ΔIKZF1 in the context of MRD in a larger Swedish patient cohort treated with NOPHO ALL protocols (Article II) and, furthermore, I explored coexisting genetic abnormalities in ΔIKZF1-positive cases (Article III). In the next section, our present knowledge of IKZF1 is summarized, including its association with pediatric BCP ALL, together with the results presented in Articles I-III and the conclusions drawn.
IKAROS and LEUKEMIA

The Transcription Factor IKAROS

The zinc finger protein IKAROS, encoded by the \textit{IKZF1} gene at 7p12.2 (chr7:50,304,124-50,405,101; http://www.ensembl.org/index.html), was first identified in 1992 as an important transcription factor for the maturation of T lymphocytes in mice (Georgopoulos \textit{et al}, 1992). However, an isoform of this protein, LyF-1, had previously been reported as an essential player in the regulation of both T- and B-cell lineage specification (Lo \textit{et al}, 1991).

Strangely, in none of the more than 600 articles published on IKAROS is it made clear why the protein was so named. Perhaps it had something to do with “high-flying ambitions”? I thus contacted Professor Georgopoulos, who was the one identifying and characterizing the role of IKAROS in hematopoiesis in the early 1990s, asking her why she named the protein IKAROS. She promptly and kindly replied that she had thought that a fitting name for this essential transcription factor “would be IKAROS, the son of the greatest master builder in the ancient Greek world Dedalos. So IKAROS once removed from a master regulator of the hematopoietic system confers lymphoid cell fate and regulates homeostasis of immature and mature lymphocytes. Unlike Dedalos, IKAROS was playful and perhaps foolish and his demise/fall costed him as well as us (mice and men) with a failing immune system and the rise of leukemias.” So, after having read more than 600 articles on IKAROS, I finally know how its name came about.

\textit{IKZF1} consists of eight exons (Figure 6). Exon 1 is not translated but may, together with the promoter region, regulate the transcription of the gene, and little is known about the function of exons 2, 3, and 7. Exons 4-6 encode the four N-terminal Zinc fingers that are required for DNA binding and exon 8 codes for the two C-terminal zinc fingers that are used by IKAROS to dimerize either with itself or with other members of the IKAROS family. In total, the family consists of five Greek siblings: IKAROS (\textit{IKZF1}), HELIOS (\textit{IKZF2} at 2q34), AILOS (\textit{IKZF3}; 17q12-21.1), EOS (\textit{IKZF4}; 12q13.2), and PEGASUS (\textit{IKZF5}; 10q26.13) (Morgan \textit{et al}, 1997; Kelley \textit{et al}, 1998; Kaufmann \textit{et al}, 2003; Iacobucci \textit{et al}, 2012b; Schjerven \textit{et al}, 2013).

Besides full-length IKAROS (IK1), there are at least nine isoforms (IK2, IK3, IK4, IK5, IK6, IK7, IK8, IK9, and IKX), generated through alternative splicing of \textit{IKZF1} (Figure 6). Of these, IK5-9 are dominant negative (DN), with a
disrupted DNA-binding domain but with a retained dimerization domain; the remaining isoforms (IK1-4 and IKX) contain functional DNA-binding domains (Francis et al., 2011; Iacobucci et al., 2012b; Joshi et al., 2014). The isoforms are differentially expressed during myeloid (mostly IKX) and lymphoid (all isoforms) development, erythropoiesis (IK1-3), and in hematopoietic stem cells (mainly IK1). Apart from its role in hematopoiesis, IKAROS is abundantly expressed in the hypothalamus and the pituitary gland (hypophysis) and IKAROS has been shown to be essential for proper development of the neuroendocrine system (Ezzat et al., 2006a). Thus, the function of IKAROS is complex and multifaceted (Hahm et al., 1994; Molnár & Georgopoulos, 1994; Klug et al., 1998; Payne et al., 2003).

IKAROS, through its various isoforms, can, after binding to regulatory regions of key genes that have motifs for IKAROS in their promoter regions, such as PCAM1, PAX5, and CD79B, both activate and inhibit gene expression by modifying epigenetic patterns via recruitment of e.g., chromatin remodeling complexes (Almendro et al., 1996; Mahmoud & Kawano, 1996; Thompson et al., 1996; Kim et al., 1999, Koipally et al., 2002; Gurel et al., 2008; Joshi et al., 2014). Ikaros null mice are characterized by hematopoietic stem cell defects and lack of various hematologic cell types, such as B-cells and dendritic cells as well as their progenitors; this strongly indicates that IKAROS is a master regulator of normal hematopoiesis (Georgopoulos et al., 1994; Wu et al., 1997; Lopez et al., 2002; Allman et al., 2006). Hence, the normal function of IKAROS is to promote differentiation, inhibit proliferation, and to control migration-related properties in in hematopoietic cells (Cobb et al., 2000; Koipally & Georgopoulos, 2000; Christopherson et al., 2001; Sabbattini et al., 2001; Yoshida et al., 2013; Joshi et al., 2014). When this process goes awry, leukemia ensues.

The first indication that IKZF1, or rather its mouse ortholog Ikzf1, may play a role in leukemogenesis came from studies showing that heterozygous Ikzf1 mutations, leading to decreased expression of DNA-binding isoforms or increased expression of DN isoforms, result in T-cell leukemia/lymphoma (Winandy et al., 1995; Karlsson et al., 2002; Papathanasiou et al., 2003). Since then, numerous studies have addressed the pathogenetic impact of IKZF1 abnormalities in a variety of human disorders, mainly BCP ALL. Interestingly, mice with only Ikzf1 mutations do not develop BCP ALL, whereas haploinsufficiency of Ikaros seems to be sufficient to accelerate the development of BCP ALL in BCR-ABL1-positive mice (Virely et al., 2010). Thus, IKZF1 abnormalities are not a primary change in leukemogenesis; they represent secondary events where haploinsufficiency of IKZF1 may play a vital role during the evolution of the leukemic clone.

Apart from expression of certain DN isoforms and ΔIKZF1, loss of function of IKAROS may also occur through hyperphosphorylation. Indeed, it has been shown that the activity of IKAROS can be regulated through phosphorylation by the casein kinase II (CK2), which inhibits the function of IKAROS; in contrast, the protein phosphatase 1 (PP1) restores its activity in hematopoietic cells (Gómez-del Arco et al., 2004; Gurel et al., 2008; Popescu et al., 2009).
Figure 6. The *IKZF1* gene, including the most common focal deletions and all sequence mutations reported in pediatric BCP ALL. Schematic illustration of the *IKZF1* gene (at 7p12.2) with the coding exons (ex2-8) in grey, the non-coding exon (ex1) in white, and the zinc finger motifs (ZF1-6) in black. The genomic extent of the recurrent focal deletions reported in article II is displayed by the horizontal bars, followed by the percentage of each deletion type (based on the 50 non-focal and focal deletions identified in total). All sequence mutations detected in BCP ALL (Mullighan et al., 2009b; Kuiper et al., 2010; Waanders et al., 2011; Roberts et al., 2012, 2014a; Article III) are given above the schematic figure of full-length IKAROS; recurrent mutations are indicated by the superscript “1”. Some of the alternative splicing isoforms, including the three dominant negative IK6, IK7, and IK8 that are detected in normal myeloid and lymphoid cells as well as in leukemia are listed below the primary structure of full-length IKAROS. This figure is based on Figure 1 in Francis et al. (2011) and Iacobucci et al. (2012b).
IKZF1 in Human Disease

Constitutional IKZF1 Abnormalities

Constitutional SNPs in IKZF1 (rs10911362 and 2362293) or upstream of the gene (rs4917014) have been shown to confer susceptibility to systemic lupus erythematosus (He et al., 2010; Cunningham Graham et al., 2011; Hu et al., 2011; Wang et al., 2013). In contrast, one SNP (rs10272724) has been associated with decreased risk of type 1 diabetes (Swafford et al., 2011). However, it is in this context perhaps more interesting that certain SNPs in IKZF1 seem to predispose to different types of leukemia: T-cell ALL (rs11978267; Treviño et al., 2009), infant and childhood AML (homozygous variants of rs11978267; Ross et al., 2013; Rudant et al., 2013), adult BCP ALL (rs4132601; Burmeister et al., 2014), and, most convincingly, childhood BCP ALL (rs4132601; Prasad et al., 2010; Papaemmanuil et al., 2009; Pastorczak et al., 2011; Xu et al., 2013). Although only the rs4132601 has been shown to decrease the mRNA expression of IKZF1 and that it only confers a moderately increase risk (odds ratios between 1.5 and 2.8) in adult and pediatric BCP ALL, it nevertheless may be an important factor in familial ALL (Hemminki & Jiang, 2002). Furthermore, it was recently reported that the rs4132601 was associated with age at diagnosis, with carriers of the GG genotype being younger (median age 3.8) compared with those with the TG or TT variants (median age 5.6); the SNP did not correlate with the presence of somatic ΔIKZF1 in pediatric BCP ALL (Górniaik et al., 2014). Interestingly, a constitutional SNP in the GATA3 gene, coding for a transcription factor involved in the regulation of T-cell development (Yagi et al., 2011), has been associated with IKZF1 aberrations. This inherited variant, which results in altered expression of GATA3, was significantly more common in ΔIKZF1-positive than in ΔIKZF1-negative BCP ALLs (Perez-Andreu et al., 2013).

As discussed below, ΔIKZF1 in BCP ALL is in all practice an acquired change. However, and interestingly, a patient with Greig cephalopolysyndactyly syndrome (GCPS), characterized by various skeletal malformations and caused by mutations of the GLI3 gene at 7p14.1, with a constitutional deletion of 7p, including the IKZF1 locus, has been reported (Mendoza-Londono et al., 2005). The patient also had ALL, unfortunately not specified with regard to immune-phenotypic subtype. Because leukemia is not a feature of GCPS, the authors proposed that the germline deletion of IKZF1 could have conferred an increased risk of ALL. To the best of my knowledge, no further GCPS cases with constitutional IKZF1 deletions have been reported. Furthermore, I did not detect any constitutional ΔIKZF1 among the 132 remission samples investigated in Articles I and III.
Acquired IKZF1 Abnormalities

Several acquired changes of IKZF1 have been reported in neoplastic disorders; those in BCP ALL will be discussed in detail in the section “IKZF1 in BCP ALL” below. As seen in Table 1, both genetic and epigenetic abnormalities of IKZF1 have been described in human neoplasia. Epigenetic changes, affecting the expression and function of IKAROS without any underlying DNA mutations, have been found in pituitary tumors, colon cancer, and AML, whereas deletions, sequence mutations, or rearrangements of IKZF1 have so far been identified in BCP ALL, CML, diffuse large B-cell lymphoma (DLBCL), myelodysplastic syndrome (MDS), T-cell ALL, and tumors of the large intestine, lung, and liver.

Considering that IKAROS is highly expressed in the pituitary gland, as mentioned above, it may perhaps come as no surprise that the DN isoform IK6 (discussed below) has been detected in half of all human pituitary tumors investigated (Table 1), suggesting that aberrant IK6 expression plays an important pathogenetic role (Ezzat et al., 2006b). There is, however, as of yet no evidence for genetic alterations of IKZF1 in pituitary tumors. Hypermethylation of the promoter region of IKZF1 has been detected in the majority of colorectal adenocarcinoma cases (Table 1), with analyses of colorectal cancer cells after reintroduction of IKAROS revealing several targets possibly related to progression of colorectal cancer (Javierre et al., 2011). In addition, overexpression of CK2, possibly leading to hyperphosphorylation of IKAROS, has been reported in AML (Kim et al., 2007; Francis et al., 2011) and T-cell lymphoma has been shown to arise in mice with aberrant expression of the catalytic subunit of CK2 in their lymphocytes (Seldin & Leder, 1995). IKZF1 deletions are very rare in T-cell ALL and AML (Crescenzi et al., 2004; Raghavan et al., 2008; Marçais et al., 2010; Koh et al., 2014; Article I; unpublished AML data); perhaps hyperphosphorylation of IKAROS may be a factor in such cases.

Moreover, genetic abnormalities other than ΔIKZF1 have been described in hematologic malignancies, such as IKZF1 translocations in DLBCL [t(3;7)(q27;p12)/IKZF1-BCL6] and MDS [t(1;7)(p36;p12)/IKZF-PRDM16] (Yoshida et al., 1999; Hosokawa et al., 2000; Duhoux et al., 2011), and rare sequence mutations in T-cell ALL (Dovat & Payne, 2010; Zhang et al., 2012b. Although IKZF1 deletions in T-cell ALL have been reported to be uncommon (Marçais et al., 2010), IKZF1 abnormalities, both ΔIKZF1 and sequence mutations, have been shown to be recurrent in the recently identified subtype early T-cell precursor ALL (Zhang et al., 2012b).
<table>
<thead>
<tr>
<th>Disorder</th>
<th>Type of aberration (percentage of cases)</th>
<th>Clinical impact</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BCP ALL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pediatric</td>
<td>ΔIKZF1 (15)</td>
<td>Poor outcome</td>
<td>Waanders <em>et al</em> (2011); van der Veer <em>et al</em> (2013); Yamashita <em>et al</em> (2013); Article III</td>
</tr>
<tr>
<td></td>
<td>Sequence mutation (5.7)</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>Adult</td>
<td>ΔIKZF1 (30-50)</td>
<td>Maybe poor outcome</td>
<td>Tokunaga <em>et al</em> (2013); Beldjord <em>et al</em> (2014)</td>
</tr>
<tr>
<td>Down syndrome</td>
<td>ΔIKZF1 (35)</td>
<td>Poor outcome</td>
<td>Buitenkamp <em>et al</em> (2012)</td>
</tr>
<tr>
<td><strong>BCR-ABL1-positive</strong></td>
<td>ΔIKZF1 (75)</td>
<td>Poor outcome</td>
<td>Mullighan <em>et al</em> (2008b); Martinelli <em>et al</em> (2009); Iacobucci <em>et al</em> (2009); van der Veer <em>et al</em> (2014)</td>
</tr>
<tr>
<td><strong>Other hematologic malignancies</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AML</td>
<td>Possible hyperphosphorylation (33)</td>
<td>Poor outcome</td>
<td>Kim <em>et al</em> (2007), Francis <em>et al</em> (2011)</td>
</tr>
<tr>
<td>CML in lymphoid blast crisis</td>
<td>ΔIKZF1 (60)</td>
<td>Poor outcome</td>
<td>Nakayama <em>et al</em> (1999)</td>
</tr>
<tr>
<td>DLBCL</td>
<td>Fusion to (BCL6) (NK)</td>
<td>NK</td>
<td>Yoshida <em>et al</em> (1999); Hosokawa <em>et al</em> (2000)</td>
</tr>
<tr>
<td>MDS</td>
<td>Fusion to (PRDM16) (NK)</td>
<td>NK</td>
<td>Duhoux <em>et al</em> (2011)</td>
</tr>
<tr>
<td>T-cell ALL</td>
<td>Sequence mutation (NK)</td>
<td>NK</td>
<td>Dovat &amp; Payne (2010); Zhang <em>et al</em> (2012b)</td>
</tr>
<tr>
<td><strong>Solid tumors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pituitary tumors</td>
<td>IK6 expression (50)</td>
<td>NK</td>
<td>Ezzat <em>et al</em> (2006a,b)</td>
</tr>
<tr>
<td>Colorectal adenocarcinoma</td>
<td>Hypermethylation (64)</td>
<td>NK</td>
<td>Javierre <em>et al</em> (2011)</td>
</tr>
<tr>
<td>Large intestine tumors</td>
<td>Sequence mutation (NK)</td>
<td>NK</td>
<td><a href="http://www.ensembl.org/index.html">http://www.ensembl.org/index.html</a></td>
</tr>
<tr>
<td>Lung tumors</td>
<td>Sequence mutation (NK)</td>
<td>NK</td>
<td><a href="http://www.ensembl.org/index.html">http://www.ensembl.org/index.html</a></td>
</tr>
<tr>
<td>Liver tumors</td>
<td>Sequence mutation (NK)</td>
<td>NK</td>
<td><a href="http://www.ensembl.org/index.html">http://www.ensembl.org/index.html</a></td>
</tr>
</tbody>
</table>

AML, acute myeloid leukemia; BCP ALL, B-cell precursor acute lymphoblastic leukemia; CML, chronic myeloid leukemia; ΔIKZF1, deletion of \(IKZF1\); DLBCL, diffuse large B-cell lymphoma; IK6, dominant negative isoform of IKAROS; MDS, myelodysplastic syndrome; NK, not known.
Most studies on the role played by IKAROS in human malignancies have addressed the impact of somatic \textit{IKZF1} changes in the leukemogenic process. Initially, those interested in the field of “IKAROS and leukemia” focused on the expression patterns of the various isoforms, of which IK6 received much attention.

The IK6 Isoform

The first studies implicating disruption of normal IKAROS expression in human leukemogenesis were published in the late 1990s. Nakayama \textit{et al} (1999) analyzed the relative expression levels of the IK1-4 and IK6 isoforms in BM samples from various types of hematologic malignancy, including CML in chronic phase as well as in lymphoid blast crisis, and observed overexpression of DN IK6 and sometimes reduced expression of IK1-4 during the blast crisis. In two of the latter patients, Southern blot analyses revealed “novel bands”, indicating the presence of a mutation in \textit{IKZF1}. No abnormalities were found in samples from CML patients in chronic phase. Thus, they concluded that an acquired altered IKAROS expression could contribute to lymphoid blast crisis. In the same year, Sun \textit{et al} (1999) reported that several DN isoforms, including IK6, IK7, and IK8 (Figure 6), were present in almost all samples from pediatric T-cell and BCP ALL patients. These findings, however, were challenged when IK7 and IK8 also were detected in normal BM. In contrast, IK6 was, and still is, considered an oncogenic isoform (Klug \textit{et al}, 1998; Nakayama \textit{et al}, 2000; Payne \textit{et al}, 2001).

Several subsequent expression studies revealed that overexpression of IK6 occur in pediatric and adult BCP ALL (mainly in the \textit{BCR-ABL1}-positive subgroup), pediatric AML, and, less frequently, in T-cell ALL, whereas IK6 was absent in adult AML (Nakase \textit{et al}, 2000; Nishii \textit{et al}, 2002a,b; Yagi \textit{et al}, 2002; Iacobucci \textit{et al}, 2008; Meleshko \textit{et al}, 2008). At that time, abnormal splicing of \textit{IKZF1} was thought to be the underlying mechanism for the altered expression of IKAROS; the Southern blot findings reported by Nakayama \textit{et al} (1999) were apparently not known (or disregarded). Though, one more group did suggest that cases expressing IK6 might have mutations in the \textit{IKZF1} locus; this was, however, not proved (Takanashi \textit{et al}, 2002).

\textbf{\textit{ΔIKZF1}}

There was a dramatic renewed interest in \textit{IKZF1} almost ten years later, when a SNP array-based study by Mullighan \textit{et al} (2008a) clearly showed that the “leukemogenic” IK6 isoform could be generated through an acquired focal deletion of \textit{IKZF1}, involving only a few exons, i.e., \textit{Δ}4-7 (\textit{Δ}3-6 according to the
previous nomenclature), in pediatric BCP ALL. Since then, numerous investigations, using genome-wide analyses on large patient series, have identified focal (Figure 6) as well as non-focal (larger deletions and monosomy 7) ΔIKZF1 in approximately 15% of pediatric and 30-50% of adult BCP ALL (Kuiper et al, 2010; Waanders et al, 2011; Chen et al, 2012; Iacobucci et al, 2012b; Caye et al, 2013; Dörge et al, 2013; Öfverholm et al, 2013; Tokunaga et al, 2013; van der Veer et al, 2013; Yamashita et al, 2013; Beldjord et al, 2014; Articles I and II).

The IKZF1 deletions are probably caused by inappropriate RAG activation during heavy chain recombination in lymphoid progenitor cells, particularly the focal deletions with breakpoints in the introns juxtaposed to exons 2, 4, and/or 7; these breakpoint regions are flanked by motifs for the RAG proteins (Mullighan et al, 2008a; Iacobucci et al, 2009; Nacheva et al, 2010; Meyer et al, 2013c).

The above notwithstanding, some recent studies have identified expression of IK6, without the corresponding genomic Δ4-7 deletion, in a few cases of BCP ALL (Tokunaga et al, 2013; Volejnikova et al, 2013), showing that splicing abnormalities also can generate this DN isoform, as already suggested in the “pre-ΔIKZF1 era”. However, it should be stressed that further investigations are needed to address this issue and that genomic IKZF1 deletions still are the dominant IKZF1 aberration in BCP ALL (Dupuis et al, 2013; Dörge et al, 2013; van der Veer et al, 2013; Article III).

**IKZF1 Sequence Mutations**

Considering that ΔIKZF1 is common in BCP ALL, one might have expected that also sequence mutations of IKZF1, perhaps in particular indels and nonsense mutations, resulting in loss of function, or splice site mutations, leading to DN isoforms, to be frequent. However, that does not seem to be the case. In fact, several studies, using Sanger sequencing, have either not identified any mutations in pediatric and adult BCP ALL (Mullighan et al, 2008a, 2011; Yang et al, 2011; Dupuis et al, 2013; Tokunaga et al, 2013; Van der Veer et al, 2013) or have only detected such changes in ~2% of the total of 1270 pediatric BCP ALL cases investigated in five studies (Mullighan et al, 2009b; Kuiper et al, 2010; Waanders et al, 2011; Roberts et al, 2012, 2014a). However, we observed a higher frequency of sequence mutations in pediatric BCP ALL by the use of targeted deep sequencing – 6% of the cases in our uniformly treated, population-based pediatric BCP ALL patient cohort (Table 1; Article III) harbored point mutations or indels (Figure 6). It is noteworthy that all the sequence mutations were predicted to have a damaging effect and that all were located in exons coding for zinc fingers, namely exons 4, 5, and 8.

It should, however, be emphasized that IKZF1 mutations are not unique to BCP ALL; they have also been reported in T-cell ALL, tumors of the large
intestine, liver, and lung (Table 1), as well as in single tumors of the central nervous system, kidney, and prostate (http://www.ensembl.org/index.html).

Clinical Features and Significance of $\Delta$IKZF1 changes

The presence of $\Delta$IKZF1 as a recurrent anomaly in BCP ALL was first reported in 2007 (Kuiper et al., 2007; Mullighan et al., 2007). Since then, several studies have provided strong evidence that $\Delta$IKZF1 plays a key role in leukemogenesis. In fact, such deletions have been detected in 75% of all $BCR$-$ABL1$-positive BCP ALL cases, more than 60% of CML in lymphoid blast crisis, 35% of Down syndrome-associated BCP ALL, ~30% of pediatric HR BCP ALL, and, as mentioned, in 15% of pediatric and 30-50% of adult BCP ALL in general (Table 1).

$\Delta$IKZF1, including both focal and non-focal deletions, has been shown to confer poor prognosis in terms of decreased OS and increased risk of relapse in these types of leukemia. In fact, $\Delta$IKZF1 is the only submicroscopic change consistently shown to have a significant impact on outcome, especially in pediatric BCP ALL cases cytogenetically grouped as “other” (Mullighan et al., 2008a, 2009b; Yang et al., 2008; Iacobucci et al., 2009, 2012b; Martinelli et al., 2009; Kuiper et al., 2010; Waanders et al., 2011; Buitenkamp et al., 2012; Chen et al., 2012; Asai et al., 2013; Caye et al., 2013; Dörge et al., 2013; Öfverholm et al., 2013; Tokunaga et al., 2013; van der Veer et al., 2013, 2014; Yamashita et al., 2013; Beldjord et al., 2014; Articles I and II).

Although $\Delta$IKZF1 has been detected in all major cytogenetic subgroups of BCP ALL, $IKZF1$ deletions are significantly less frequent in $ETV6$-$RUNX1$-positive cases and more common in the subtypes “other” and $BCR$-$ABL1$, with the latter two subgroups also being associated with an inferior prognosis when harboring an $IKZF1$ deletion (Martinelli, et al, 2009; Dörge et al., 2013; Enshaei et al., 2013; van der Veer et al, 2014; Articles I and II). One study suggested that $\Delta$IKZF1 confers a negative prognostic impact in HeH-positive cases (van der Veer et al, 2013) but this could not be confirmed by us (Article II). The presence of $\Delta$IKZF1 appears to be associated with certain risk features, such as high WBC counts and older age; consequently $\Delta$IKZF1-positive cases are more commonly stratified as HR. This does, however, not explain the increased risk of relapse in the $\Delta$IKZF1-positive group because $\Delta$IKZF1 is an independent risk factor also when risk stratification and MRD status are taken into account (Den Boer et al., 2009; Martinelli et al., 2009; Mullighan et al, 2009b; Dörge et al., 2013; Schwab et al., 2013; van der Veer et al, 2013, 2014; Yamashita et al, 2013; Articles I and II). In fact, it has been reported that a combination of $IKZF1$ and MRD status is the best predictor of treatment failure in pediatric BCP ALL (Waanders et al, 2011). It has also been suggested that $\Delta$IKZF1 should be used as an MRD marker – it is
always preserved at the time of relapse and is, in the vast majority of cases, acquired (Mendoza-Londono et al., 2005; Venn et al., 2012; Caye et al., 2013).

The clinical significance of sequence mutations in IKZF1 remains to be elucidated. In our study, the presence of mutations as such was not associated with outcome (Article III). However, the few cases with both ΔIKZF1 and a mutation all relapsed, suggesting a cooperative effect. Such an impact was also demonstrated in a previous study where cases harboring biallelic IKZF1 deletions, resulting in null expression of wt IKAROS in the leukemic cells, had a selective edge over the cells with heterozygous ΔIKZF1, becoming the clone emerging at relapse (Dupuis et al., 2013).

**How to detect IKZF1 Abnormalities**

The presence of ΔIKZF1 has already been, or is on the verge of being, implemented as a risk-stratifying marker in pediatric BCP ALL treatment protocols (Mullighan, 2011; Dörge et al., 2013; Schwab et al., 2013; van der Veer et al., 2013, Moorman et al., 2014). It is hence important to ascertain which method(s) that most accurately detects the above-mentioned diverse changes in IKZF1 and hence should be implemented in clinical routine.

Today, most laboratories apply SNP array or MLPA analyses to identify ΔIKZF1. More recently, multiplex PCR-based methods have been developed for ΔIKZF1 detection and targeted deep sequencing has been shown to be the method of choice for mutation analysis. However, all these methods have their limitations. For example, deletions may well escape detection by SNP arrays and MLPA if they are present in smaller (<10% and <30%, respectively) subclones, large, i.e., non-focal, deletions are missed by multiplex PCR, and targeted deep sequencing does not provide information on deletions at all, except small indels (Kuiper et al., 2007; Mullighan et al., 2007, 2008a; Schwab et al., 2010; Li et al., 2011; Caye et al., 2013; Dupuis et al., 2013; Meyer et al., 2013c; Articles II and III). Thus, a combination of methods, such as SNP arrays and targeted deep sequencing, seems to be required. However, none of these methods detects post-translational abnormalities.

**Pathogenetic Impact**

Deletions including exons 4, 5, and/or 6, with Δ4-7 being most common, correspond to expression of DN IKAROS isoforms that could interfere with wt proteins of the IKAROS zinc finger family, whereas focal deletions affecting exons 1-3 and 8 and non-focal deletions result in reduced levels of IKAROS, i.e., haploinsufficiency (Mullighan et al., 2008a; Iacobucci et al., 2009; Caye et al., 2013; Dupuis et al., 2013). Mouse models with DN Ikaros isoforms have shown a
more severe leukemic phenotype than mice with only lower expression of Ikaros. This would suggest that haploinsufficiency may be less “leukemogenic”. However, haploinsufficiency has been shown to be sufficient to accelerate the development of leukemia (Georgopoulos et al, 1994; Paphathanasiou et al, 2003; Ruiz et al, 2008; Virely et al, 2010; Dupuis et al, 2013) and, furthermore, there are no clinical differences between BCP ALL cases with focal and non-focal ΔIKZF1 (Dupuis et al, 2013; Dörge et al, 2013; Article II).

It has been reported that reduced Ikaros expression in mice results in a differentiation block at an early pro B-cell stage and that introduction of the DN IK6 into murine stem cells impair B-lymphocyte lineage commitment (Tonnelle et al, 2001; Kirstetter et al, 2002; Klein et al, 2006; Reynaud et al, 2008). In addition, it has been shown that the chimeric protein PICALM-MLLT10, which is present in a few cases of ALL, AML, and lymphoma, binds to and alter the subcellular localization of Ikaros in a murine BM transplantation model. This results in the development of an aggressive leukemia, suggesting that PICALM-MLLT10 interferes with the normal IKAROS function and thereby blocks lymphoid differentiation (Greif et al, 2008). Finally, in BCP ALL cell lines, IKAROS seems to be essential for response to glucocorticoid treatment, whereas IK6 expression is associated not only with treatment resistance but also with decreased apoptosis and increased proliferation compared with cell lines with wt IKAROS (Wargnier et al, 1998; Sezaki et al, 2003; Ruiz et al, 2004).

**Gene Expression Signature**

To explore the underlying pathways and downstream targets of aberrant IKAROS expression, Iacobucci et al (2012b) performed a gene expression study of adult BCP ALL. ΔIKZF1-positive cases were shown to have a unique expression signature that, interestingly, involved down-regulation of DNA repair genes and up-regulation of stem-cell self-renewal and JAK-STAT signaling. The latter has also been shown to be associated with ΔIKZF1 in expression analyses of pediatric BCP ALL (Den Boer et al, 2009; Mullighan et al, 2009b; Harvey et al, 2010b). Furthermore, in an experimental model of pre-B cells with IK6 expression, a similar up-regulation of pro-survival kinase pathways was identified (Kano et al, 2008), indicating that wt IKAROS regulates these pathways. Moreover, there is an association between aberrant IKAROS activity and stem cell features (Tonnelle et al, 2001; Iacobucci et al, 2012b). In fact, very recently, ΔIKZF1-positive BCP ALL cells were shown to be dependent on the stem cell niche, stromal cells, and integrin-signaling to survive and progress to ALL (Joshi et al, 2014); this finding may have therapeutic implications. However, in a recent study there was no increased treatment resistance in IKZF1 knockdown cells compared with wt IKZF1 controls after challenging the cells with chemotherapy.
(Vitanza et al, 2014), suggesting that the poor prognosis associated with IKZF1 deficiency in human ALL may be influenced by coexisting genetic changes.

Cooperative Changes in BCP ALL

As mentioned, IKZF1 aberrations do not seem to be sufficient for overt leukemia to arise (Virely et al, 2010) and, furthermore, not all ΔIKZF1-positive cases relapse. Thus, it is likely that certain cooperative genetic alterations may modify the pathogenetic and/or prognostic impact, for better or for worse (Article III). Also, ΔIKZF1 seems to be a secondary hit during leukemogenesis, considering that IKZF1 aberrations occasionally only are present in subclones at diagnosis and that different subclones, harboring diverse ΔIKZF1, have been found in some instances (Dupuis et al, 2013; Article I). This notwithstanding, ΔIKZF1 is always preserved at the time of relapse or only emerges at relapse, strongly indicating that IKZF1 abnormalities are associated with treatment failure (Mullighan et al, 2007, 2008a; Yang et al, 2008; Dupuis et al, 2013; Articles I and III). Cazzaniga et al (2011) reported interesting data pertaining to the latter. They investigated two pairs of twins, all four of whom were positive for the BCR-ABL1 fusion gene; three were diagnosed with ALL whereas one was healthy, carrying the fusion in a low frequency. The BCP ALL in two twins, one in each pair, also harbored a ΔIKZF1. Both these died as a result of aggressive disease, the other two, without ΔIKZF1, survived. This provides further support for the notion that ΔIKZF1 is a secondary abnormality with a dismal prognostic impact. In addition, it also begs the question: what are the cooperative changes that may modify this impact?

Abnormal Kinase Signaling Pathways and Deletions of ERG

It has been suggested that the BCR-ABL1 chimera may induce IKZF1 changes, especially the Δ4-7 deletions that are strongly associated with this cytogenetic subgroup (Klein et al, 2006; Dupuis et al, 2013). Furthermore, the dismal outcome of BCR-ABL1-positive cases that also harbor ΔIKZF1 may be due to the inappropriate activation of the JAK-STAT signaling pathway, which decreases the response to tyrosine kinase inhibitors (Iacobucci et al, 2008; Trageser et al, 2009). Also the BCR-ABL1-like group is characterized by coexisting ΔIKZF1 and such cases frequently carry other rearrangements involving genes coding for tyrosine kinases or cytokine receptors, such as ABL1 (other than fusion to BCR), CRLF2, JAK2, and PDGFRB. These may contribute to the up-regulation of JAK-STAT signaling and to the expression signature in the BCR-ABL1-like cases (Den Boer et al, 2009; Mullighan et al, 2009b; Harvey et al,
2010b; Roberts et al, 2012, 2014a; van der Veer et al, 2013). In contrast to BCR-ABL1-positive cases, the prognostic impact of ΔIKZF1 in BCR-ABL1-like cases is less clear – van der Veer et al (2013) did not observe any adverse effect on the relapse risk whereas Roberts et al (2014a) reported that IKZF1 aberrations contributed to an inferior survival. Furthermore, deregulation of CRLF2 and JAK2 mutations are significantly more common, and may also increase the risk of relapse, in ΔIKZF1-positive BCP ALL cases (Mullighan et al, 2009b,c; Chen et al, 2012; Schwab et al, 2013; Van der Veer et al, 2013; Yamashita et al, 2013; Article III). Thus, abnormal kinase signaling and ΔIKZF1 definitely goes together.

There are emerging data indicating that deletions of the transcription factor-encoding gene ERG at 21q22.2 actually may improve the outcome of ΔIKZF1-positive cases (Clappier et al, 2014; Zaliova et al, 2014); our findings did, however, not support this (Article III).

In conclusion, the genomic context in which ΔIKZF1 occurs is important, indicating that the mere presence of this deletion may not be sufficient for proper risk stratification.

Controversies

Recently, the pathogenetic impact and the clinical relevance of IKZF1 abnormalities have been questioned, either as a genomic change as such or as an independent prognostic marker.

In an analysis of more than 250 BCR-ABL1-positive and -negative pediatric BCP ALL cases, neither intragenic IKZF1 deletions nor any expression patterns indicative of such deletions were observed (Qazi et al, 2013; Uckun et al, 2013a). Furthermore, unaffected expression levels in cases with hemizygous IKZF1 deletions as well as a weak expression of IK6 in healthy controls have also been reported (Tokunaga et al, 2013; Volejnikova et al, 2013). Thus, the correlation between genomic ΔIKZF1 and IKZF1 expression patterns does not seem to be clear-cut. However, even a normal expression may be pathogenetically important, at least if IKAROS is located in the cytoplasm instead of in the nucleus (Marçais et al, 2010).

There are also conflicting data with regard to the correlation between ΔIKZF1 and outcome, particularly in adult BCR-ABL1-negative BCP ALL in which the prognostic impact of ΔIKZF1 has been reported to be low or only relevant for focal deletions or DN isoforms (Liu et al, 2012; Moorman, 2012; Moorman et al, 2012; Tokunaga et al, 2013; Beldjord et al, 2014).

There are contradictory reports also in pediatric BCR-ABL1-negative BCP ALL as to whether ΔIKZF1 is an independent risk factor when, for example MRD data and co-occurring aberrations, are taken into account (Chen et al, 2012; Palmi et al, 2013; Volejnikova et al, 2013; Clappier et al, 2014; Zaliova et al, 2014).
Furthermore, some studies have suggested that the poor outcome associated with ΔIKZF1 is due to the presence of a higher frequency of CNAs in such cases compared with non-deleted cases, i.e., the negative impact is caused by an underlying genomic instability rather than by the ΔIKZF1 as such (Palmi et al., 2013; Qazi & Uckun, 2013). However, my findings do not agree with this – ΔIKZF1 was an independent risk factor also in the context of MRD (Article II) and there was no significant difference in the frequency of CNAs between ΔIKZF1-positive and ΔIKZF1-negative cases (Article III).

The above notwithstanding, the fact that all ΔIKZF1 at diagnosis, occasionally only being present in subclones, are preserved at relapse and that ΔIKZF1 may emerge first at relapse (Mullighan et al., 2008b; Kuiper et al., 2010; Articles I and III) strongly indicates that IKZF1 aberrations as such are important for leukemic evolution.

Conclusions and Future Directions

To conclude, it has been clearly demonstrated that changes affecting IKZF1 are important genetic events in leukemogenesis and progression, with ΔIKZF1 being a strong independent risk factor for relapse and dismal outcome. Thus, it is vital to identify ΔIKZF1-positive cases already at the time of diagnosis, in particular those that today are cytogenetically grouped as “other”, in order to risk stratify and treat the patients correctly. Moreover, also among the BCP ALL cases already risk-stratified into HR groups those with ΔIKZF1 may be important to detect because they could benefit from alternative treatment, such as tyrosine kinase inhibitors.

However, several issues should be addressed in the future, such as the true frequency and the clinical impact of IKZF1 aberrations other than ΔIKZF1, e.g., sequence mutations and epigenetic or post-transcriptional silencing. Furthermore, further analyses of cooperative genetic events that may modify the effect of IKZF1 abnormalities are required and, finally, robust and standardized method(s) to rapidly detect these changes at the time of diagnosis of BCP ALL need to be developed.
Svensk Sammanfattning

Den vanligaste barncancerformen är akut leukemi, en form av blodcancer som uppstår i immunförsvarsets celler och som drabbar cirka 75 barn i Sverige varje år. Även om de flesta av barnen idag botas får en av fem återfall och då minskar överlevnaden drastiskt trots intensiv behandling. Vad det är som gör leukemin mer motståndskraftig när den återkommer vet man ännu inte i detalj. Det är sålunda viktigt att man redan vid diagnostillfället så säkert som möjligt kan förutså vilka barn som har stor risk för återfall och som av den anledningen ska få mer behandling. Det är även betydelsefullt att identifiera de leukemipatienter som kan botas med mindre aggressiv terapi, för att i görligaste mån minska bi-verkningar associerade med cytostatika.

Immunförsvarset består av många olika celltyper som cirkulerar runt i vårt blod och som skyddar oss mot infektioner. Dessa härstammar från stamceller i benmärgen vilka stegvis delar sig och mognar ut till de olika celltyperna. Vid akut leukemi tar istället ett stort antal omogna blodceller över benmärgen och blodet, med infektioner, blödningar och trötthet som följd. Beroende på vilken celltyp leukemin liknar/uppstår i får den ett specifikt namn. Akut leukemi kan således delas upp i flera undergrupper, t.ex. akut lymfatisk leukemi (ALL) som uppstår i B- eller T-celler och som står i fokus i denna avhandling. Den vanligaste formen hos barn är B-cells prekursor ALL (BCP ALL), vilken är associerad med en ansamling av omogna B-celler (celltypen som i vanliga fall producerar våra antikroppar). Korrekt klassifikation är viktigt eftersom det spelar en stor roll vid val av behandlingsstrategi.

I likhet med andra typer av cancer orsakas ALL av förändringar (mutationer) i arvsmassan (DNA) i de sjuka cellerna. I princip har alla celler i kroppen samma genetiska uppsättning men i leukemicellerna har DNA-material tillkommit, förändrats och/eller förlosts vilket leder till att gener finns i fler kopior än normalt, förvandlas eller försvinner, något som resulterar i okontrollerad tillväxt och förlorad cellfunktion. Dessa avvikelser är oftast inte medfödda utan uppstår i celler under vår livstid; de är s.k. förvärvade mutationer. ALL är en heterogen sjukdom där varje enskilt fall uppvisar ett mer eller mindre unikt mutationsmönster. Dock är vissa mutationer återkommande och har visats vara viktiga att identifiera då de har stor betydelse för prognos och behandling. Vidare har nya läkemedel som specifikt angriper genförändringar tillkommit.

Numera används genetiska analyser för att påvisa förändringar som innebär att patienten ska ha mer eller mindre intensiv behandling, något som medför en

Syftet med mitt avhandlingsarbete har varit att angripa ovanstående frågeställningar genom att kartlägga genetiska förändringar som förekommer i patienter som fått återfall (re cidiverat) eller inte svarat på den initiala behandlingen och att studera förekomsten av dessa mutationer i leukemiprov från barn som blivit botade, med målet att finna förändringar som är klart överrepresenterade i de fall som sedermera recidiverar. I avhandlingens första artikel (I) undersöktes benmärgs- eller blodprover från 228 barn med ALL, såväl T-cells ALL som BCP ALL. Jag använde mig av en metod som heter ”SNP array” och som möjliggör upptäckt av tillskott eller förlust av DNA-material i cancercellerna som finns i provet. I T-cell ALL-fallen hittade jag inga genetiska förändringar som kunde förklara återfallen. I BCP ALL, däremot, fann jag återkommande förluster av viktiga gener som kunde förutsäga återfall (redan vid diagnostikkällfallet) och som kunde förklara varför vissa återfall var motståndskraftiga mot behandlingen. En specifik gen, med gennamnet IKZF1 och som leder till ett protein som heter IKAROS, var borta i 15 % av alla BCP ALL-fall. Denna förlust var associerad med återfall och sämre överlevnad. Dessutom var tre gener (med gennamnnen BTG1, MSH6 och NR3C1) oftare förlorade i recidivceller än i celler från diagnosställfallet. Detta visar att leukemicellerna utvecklats vidare genom tillkomst av nya mutationer och på så sätt blivit resistenta mot recidivbehandling.

I de två övriga artiklarna (II och III) inkluderade i min avhandling utvärderades IKZF1-förlust som en ny riskmarkör i BCP ALL. I delarbete II insamlades flera fall runt om i Sverige med känt IKZF1 status, d.v.s. oförändrad eller förlorad, för att få ett bättre statistiskt underlag. Denna studie klargjorde att IKZF1-förlust är en oberoende och viktig riskmarkör som identifierar många fall som sedermera recidiverar och som inte upptäcks med dagens riskparametrar som t.ex. hög ålder, stor andel omogna vita blodceller i blodet vid diagnostikkällfallet och förekomst av leukemiceller som överlevt den initiala behandlingen. Dessutom visades IKZF1-förlust vara särskilt viktig att upptäcka i de fall som idag inte uppvisar några typiska ALL-associerade genetiska avvikelser. I delarbete III studerades leukemiceller från 140 BCP ALL-patienter med en annan teknik, nämligen s.k. storskalig sekvensering, där specifika arvsanlag (gener) kan läsas av och små mutationer som förändrar gener påvisas. Eftersom vi sett att IKZF1-förlust är en riskmarkör så undersökte jag om även andra mutationer i denna gen kan bidra till återfall och därmed också vara väsentliga att upptäcka redan vid
diagnostiffället för korrekt prognosbedömning och rätt behandling. Jag kunde då visa att andra förändringar i \textit{IKZF1} bara ger återfall om de samexisterar med \textit{IKZF1}-förlust. Samtidigt undersöktes alla fall med \textit{IKZF1}-förändring närmare med tonvikt på förlust av andra gener. Strukturella förändringar av \textit{CRLF2}-genen, förlust av \textit{SH2B3}-genen och mutationer i \textit{JAK2}-genen var vanligare i fall med \textit{IKZF1}-förändringar och \textit{CRLF2} verkade samverka med och påverka den kliniska effekten av \textit{IKZF1}-förändringar. Detta innebär att för att förstå varje enskilt leukemifall behöver man kartlägga även andra mutationer än \textit{IKZF1}-förlust vid diagnostiffället.

Sammanfattningsvis har undersökningarna i mitt avhandlingsarbete lett till en ökad förståelse för det komplexa mönster av genetiska förändringar som bidrar till leukemiarecidiv. Vidare har jag påvisat en viktig riskmarkör i BCP ALL, nämligen förlust av \textit{IKZF1}-genen. Förhoppningsvis kan denna kunskap leda till en förbättrad behandling av BCP ALL hos barn och ungdom i framtiden.
Acknowledgments

I would like to express my sincere gratitude to everyone who has contributed to my PhD studies and in particular thank all of you mentioned below. I hope you all know how grateful I am for your help and support during this thesis work. I could not have done this and these four years would not have been as fun without you.

My supervisor, Professor Bertil Johansson, for showing me that you can be a brilliant scientist, caring, and generous, all at the same time. For giving me the confidence to be independent but always knowing that I can ask you about anything. I could not have wished for a better supervisor and I am so happy that you gave me the opportunity to be a PhD student in your group and to learn everything there is to know, about genetics, the English language, and how a perfect table should look like, from you. The only downside has been that you not yet have visit Kalmar, but I hope to see you there next summer!

My co-supervisor, Docent Kajsa Paulsson, for helping me be (almost) as good as you in interpret SNP array data and master this technique but also for making me realize that science is best to be discussed over a glass of wine.

My friends and fellow PhD students, Anders, Axel, Charles, Daniel, Elsa, Jakob, Kristina K, Mia, Niklas, Pablo, Setareh and Sofia, for all the laughers, bib nights, and sport activities, particularly the table tennis games at the end of my thesis writing so that I got some kind of workout. But “Lindakrossen/olympiaden” is not over and I hope you will continue trying to beat me… A special thanks to Kristina K, for being a great traveling partner at congress trips and for finishing your PhD study one week before me so that I could walk in your footsteps. To Setareh, for being a cozy roommate at conferences, for allowing me to pass on my wisdom 😊, and for all the fun during these years.

Everyone in the research group; Andrea for all the help in the lab, Catarina (together with Anna Andersson) for taking care of me in Amsterdam when I was sick and without money, and Kristina L and Rebeqa for all the fun in Bologna learning bioinformatics.

Thank you Henrik for taking your time and answer every bioinformatic question (and they have not been few), especially for the great talk (at least from my
perspective) about the future of genetics and sequencing, over the clouds from Helsinki to Copenhagen, when you were stuck on an airplane with me and my NGS-thoughts.

To you who laid the groundwork for my knowledge of genetic methods, Karolin, Linda M, and Margareth. You are not forgotten and you inspired me to continue working at Clinical Genetics.

The very skilled and helpful co-workers Eleanor, Helena Å, Jenny K, Jenny N, Linda HM, Marianne and Nils H.

All principal investigators at the division of Clinical Genetics, Felix Mitelman, Fredrik Mertens, Thoas Fioretos, David Gisselsson Nord, Anna Andersson, and Marcus Järås for contributing to a creative and warm research environment with a lot of enthusiasm and knowledge.

Anette Welin, for always having your door open for every kind of question I have had during my thesis work.

All past and present co-workers in Clinical Genetics at BMC and at the University Hospital. A special thanks to, Hans, Mia, Nina, Petra, Samuel, and Ulf, at the Hospital side for all the interesting discussions and encouragement during these years.

To my friends and family;

Ett stort tack till alla er som stöttat mig under mina år som doktorand och som kommer för att fira med mig på denna dag, det gör mig så glad!

Ett särskilt tack till Mamma och Pappa för all hjälp med disputationsfesten, för allt annat som ni alltid ställer upp på och för att ni alltid finns här för mig.

Till min Petter, jag kan inte tänka mig en bättre toastmaster/festman (Jag vet att det ska vara ett å istället för ett e, bröderna Olsson, men detta är en ordvits). Du gör festen/livet mycket roligare, varmare, och mer spännande. Tack för att du vill ha denna roll på min fest och i mitt liv!
References


lymphoblastic leukaemia with t(1;19)(q23;p13): clinical and cytogenetic characteristics of 47 cases from the Nordic countries treated according to NOPHO protocols. *British Journal of Haematology*, **155**, 235-243.


Bennett, J.H. (1845) Case of hypertrophy of the spleen and liver in which death took place from suppuration of the blood. *Edinburgh Medical and Surgical Journal*, 64, 413-423.


the European clinical trials directive while surviving the administrative pressure - an alternative approach to toxicity registration in a cancer trial. *European Journal of Cancer, 50*, 251-259.


Waanders, E., Scheijen, B., van der Meer, L.T., van Reijmersdal, S.V., van Emst, L.,
Kroeze, Y., Sonneveld, E., Hoogerbrugge, P.M., van Kessel, A.G., van Leeuwen,
F.N. & Kuiper, R.P. (2012) The origin and nature of tightly clustered $BTG1$
deletions in precursor B-cell acute lymphoblastic leukemia support a model of multiclonal

Walsh, K.M., de Smith, A.J., Chokkalingam, A.P., Metayer, C., Roberts, W., Barcellos,
ancestral components in Hispanic children with ALL. *Blood*, 122, 3385-3387.

*Frontiers of Medicine*, 6, 263-274.

Wang, C., Ahlford, A., Järvinen, T.M., Nordmark, G., Eloranta, M.L., Gunnarsson, I.,
Svennungsson, E., Padyukov, L., Sturfelt, G., Jönsen, A., Bengtsson, A.A., Truedsson,
L., Eriksson, C., Rantapää-Dahlqvist, S., Sjöwall, C., Julkunen, H., Criswell, L.A.,
Graham, R.R., Behrens, T.W., Kere, J., Rönnblom, L., Sjövall, C., Sandling, J.K. (2013) Genes identified in Asian SLE GWASs are also associated with SLE in

Diseases*, 14, 10-20.

elements of the granzyme B promoter. *The Journal of Biological Chemistry*, 273,
35326-35331.

Wesolowska-Andersen, A., Borst, L., Dalggaard, M.D., Yadav, R., Rasmussen, K.K.,
Wehner, P.S., Rasmussen, M., Orntoft, T.F., Nordentoft, I., Koehler, R., Bartram,
C.R., Schrappe, M., Sicheritz-Ponten, T., Gautier, L., Marquart, H., Madsen, H.O.,
Brunak, S., Stanulla, M., Gupta, R. & Schmiegelow, K. (2014) Genomic profiling of
thousands of candidate polymorphisms predicts risk of relapse in 778 Danish and

& Mullighan, C.G. (2013) Tyrosine kinase inhibitor therapy induces remission in a
patient with refractory $EBF1$-$PDGFRB$-positive acute lymphoblastic leukemia.
*Journal of Clinical Oncology*, 31, e413-416.


$PRDM9$ associated with high hyperdiploid childhood acute lymphoblastic leukaemia.
*British Journal of Haematology*, 166, 947-949.

autonomous defects in dendritic cell populations of Ikaros mutant mice point to a
developmental relationship with the lymphoid lineage. *Immunity*, 7, 483-492.


