From Department of Molecular Medicine and Surgery – MMK
Karolinska Institutet, Stockholm, Sweden

WHOLE GENOME SEQUENCING IN PEDIATRIC ALL – A TOOL FOR UNDERSTANDING CHROMOSOMAL ABERRATIONS AND IMPROVING DIAGNOSTIC PROCEDURES

Fatemah Rezayee

Karolinska Institutet

Stockholm 2025

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2025
ISBN I: https://doi.org/10.69622/28444694
DOI I: https://doi.org/10.69622/28444694
Cover illustration: generated by Office 365 Copilot

Whole genome sequencing in pediatric ALL - a tool for understanding chromosomal aberrations and improving diagnostic procedures

Thesis for Doctoral Degree (Ph.D)

By

Fatemah Rezayee

The thesis will be defended in public at Torsten Gordh, Eugeniavagen 27, Norrbacka S2:02, Karolinska Institutet, Stockholm, on Friday 16 Maj 2025 at 9 a.m.

Principal Supervisor:
Docent Gisela Barbany
Karolinska Institutet
Department of Molecular Medicine and Surgery

Co-supervisor(s):
Professor Richard Rosenquist Brandell
Karolinska Institutet
Department of Molecular Medicine and Surgery

Jesper Eisfeldt
Karolinska Institutet
Department of Molecular Medicine and Surgery

Anh Nhi Tran
Karolinska Institutet
Department of Molecular Medicine and Surgery

Opponent:
Research Director Caroline Heckman
University of Helsinki
Institute for Molecular Medicine Finland

Examination Board:
Docent Linda Holmfeldt
Uppsala University
Department of Immunology, Genetics and Pathology

Docent Roza Chaireti
Karolinska Institutet
Department of Molecular Medicine and Surgery

Andreas Lennartsson
Karolinska Institutet
Department of Medicine, Huddinge

To my sons Edvin and Oliver

Popular science summary of the thesis

Leukemia is a disease of blood cells in which the bloodstream fills with abnormal immune cells that have lost the ability to fight off infection, called leukemic cells. Over time, leukemic cells overpopulate the bone marrow and blood, displacing normal blood cells and impairing their functions. Consequently, the body becomes more susceptible to infections and loses its ability to combat even simple illnesses. Leukemia affects individuals of all ages but depending on the type of leukemia the incidence differs between children and adults.

Acute lymphoblastic leukemia (ALL) is one of the most common types of cancer in children. With a survival rate of over 90%, the treatment of children with ALL is one of the success stories of modern medicine. However, the downside of current treatment is the long-term complications, the risk of developing other cancers and, in some cases, death due to high intensity treatment. Today, one of the challenges is to identify patients who can be treated with less intensive treatment and thus experience a better quality of life. One approach to achieving this goal is to study the DNA of leukemic cells. DNA is an instruction manual for how a cell must function and an error in the crucial steps can lead to abnormal behavior in the cell. These errors are known as genetic aberrations. Some genetic aberrations in the blood cells can turn them into leukemic cells. In clinical diagnostics, several methods are available to detect and identify genetic aberrations in the DNA. Recently, a new method with high resolution has been introduced that shows promising ability to read the whole length of human DNA (whole genome sequencing (WGS)) and detect aberrations in various forms.

In this work, the potential of WGS in detecting aberrations in the DNA of leukemic cells was investigated. The main objective was to find the aberrations that are known to transform blood cells into leukemic cells. The result showed that WGS is a promising tool that can single-handedly detect all the genetic aberrations required for in the treatment protocol of ALL. In addition, these genetic aberrations can be used as markers to monitor a patient's response to the treatment and clinicians can use this information to adjust the therapy for best possible result. We selected two to three patient-specific markers and tracked them with extremely sensitive methods in patient's samples. The results show that by using our strategy, patient-specific mutations can be detected and used to monitor patients during treatment, making fine adjustments possible.

In summary, WGS with its high resolution was able to detect all mutations required for the diagnosis of childhood ALL. Even in patients with inconclusive results in routine clinical practice, WGS could detect and identify mutations involved in the transformation of blood cells into leukemic cells. However, to fully realize the potential of WGS in clinical diagnostics, additional studies are necessary.

Abstract

Whole genome sequencing is a revolutionary technology that has changed the field of genomics. By providing unprecedented base pair resolution, WGS allows for precise identification of disease causing genetic abnormalities. While cytogenetic and molecular methods are well established in clinical practice, the application of WGS in cancer diagnostics remains at the developmental stage. Additionally, detecting complex genomic events currently requires a multi-modal approach, as each method is targeted and has limited resolution. This reliance on multiple techniques often leads to cumbersome workflows and requires input from various experts with specialized knowledge.

One of the diseases that requires genetic characterization is acute lymphoblastic leukemia (ALL). This disease is known to be genetically heterogeneous, with structural genomic changes playing a central role in leukemogenesis. Currently, ALL is divided into genetic subgroups, each associated with different outcomes. Detecting and identifying the genetic aberrations driving disease progression often requires the use of multiple methods. Therefore, having a single method capable of detecting all genetic aberrations would be highly beneficial.

In this thesis, the efforts to assess the value of WGS as a stand-alone method in clinical diagnostics are described. The aim is not only to detect the genetic aberrations described in the treatment protocols, but also to identify emerging genetic subgroups and novel gene fusions that have diagnostic significance or can be used for targeted therapy.

In **Paper I** WGS was applied to a pediatric B-cell ALL (B-ALL) cohort to investigate its feasibility to detect clinically relevant genetic aberrations. Two sequencing depths (30x versus 90x coverages) and two analysis approaches (leukemia-only vs leukemia/normal) were tested, and the results were comprehensively compared to findings from standard of care (SoC) methods. Based on our analysis, sequencing leukemia samples with 30x coverage was sufficient to detect all obligatory aberrations. Additionally, almost all patients without stratifying genetic markers could be allocated to one of the emerging subgroups. The paper demonstrates the potential of WGS in the diagnostic setting of ALL .

In **Paper II** we applied WGS to genetically characterize a case of pediatric T-cell ALL (T-ALL) that did not respond to standard therapy. WGS detected a novel *JAK2::CCDC88C* gene fusion, resulting in a chimeric transcript. This novel transcript, containing the tyrosine kinase domain of *JAK2*, produces a chimeric protein expected to respond to a JAK2 inhibitor. This case report shows how WGS can be used to find personalized treatment options for leukemia patients.

In **Paper III** we present a proof-of-concept study describing the use of patient-specific genetic markers detected by WGS to monitor disease. This is performed by designing sensitive quantitative assays targeting these genetic markers. WGS was applied to detect structural variants (SVs) and other relevant leukemic genetic alterations resulting in unique sequences in six pediatric ALL cases. We then used this data to identify patient-specific targets and designed droplet digital PCR (ddPCR) assays to enable the absolute quantification of these targets. Subsequently, the technical feasibility of these patient-specific targets was assessed for monitoring measurable residual disease (MRD) in genomic DNA and cell-free tumor DNA. The results showed that the sequences provided by WGS enabled us to identify patient-specific targets that could be used for sensitive and specific detection of leukemic cells.

Manuscript IV explores the complexity of a specific ALL subtype. We applied WGS and whole transcriptome sequencing (WTS) to perform a detailed characterization of the genomic and transcriptomic profiles of Philadelphia-positive ALL (Ph+ALL). We have integrated the findings from SoC diagnostics and MRD information to comprehensively investigate the features of Ph+ALL and potential for refining diagnostics for these patients.

Together, these studies explored the potential of WGS in clinical diagnostics and enhanced the understanding of genetic aberrations in pediatric ALL.

List of scientific papers

I. Feasibility to use whole-genome sequencing as a sole diagnostic method to detect genomic aberrations in pediatric B-cell acute lymphoblastic leukemia.
F. Rezayee, J. Eisfeldt, A. Skaftason, I. Öfverholm, S. Sayyab, A. C. Syvänen, K. Maqbool, H. Lilljebjörn, B. Johansson, L. Olsson-Arvidsson, C. Orsmark Pietras, A. Staffas, L. Palmqvist, T. Fioretos, L. Cavelier, L. Fogelstrand, J. Nordlund, V. Wirta, R. Rosenquist, and G. Barbany.
Front Oncol **2023**, 13:1217712.

II. Case Report: Whole genome sequencing identifies *CCDC88C* as a novel *JAK2* fusion partner in pediatric T-cell acute lymphoblastic leukemia.
A. Krstic, **F. Rezayee**, L. Saft, A. Hammarsjö, P. Svenberg, and G. Barbany.
Front. Pediatr. **2023**, 10:1082986.

III. Patient-specific assays based on whole-genome sequencing data to measure residual disease in children with acute lymphoblastic leukemia: a proof of concept study.
C. Arthur†; **F. Rezayee**†; N. Mogensen; L. Saft; R. Rosenquist; M. Nordenskjöld; A. Harila-Saari; E. Tham; G. Barbany
Front Oncol **2022**, 12, 899325.

IV. Unveiling the complexity of Philadelphia-positive acute lymphoblastic leukemia through the integration of genomic, phenotypic, and measurable residual disease data.
F. Rezayee, J. Eisfeldt, A. Krstic, P. Svenberg, J. Joelsson, K. Belander Strålin, H. Vogt, S. Deneberg, S. Volanthen, L. Saft, R. Rosenquist, and G. Barbany.
Manuscript **2025**.

Related scientific papers not included in this thesis

I. The prognostic impact of IKZF1 deletions and UKALL genetic classifiers in paediatric B-cell precursor acute lymphoblastic leukaemia treated according to NOPHO 2008 protocols.
I. Öfverholm[†], **F. Rezayee**[†], M. Heyman, A. Harila, L. Arvidsson, K. Schmiegelow, U. Norén-Nyström, and G. Barbany.
Br J Haematol. **2023**;00:1–9

II. Multi-omics analysis reveals multiple mechanisms causing Prader–Willi like syndrome in a family with a X;15 translocation.
J. Eisfeldt, **F. Rezayee**, M. Pettersson, K. Lagerstedt, H. Malmgren, A. Falk, G. Grigelioniene, and A. Lindstrand.
Human Mutation. **2022**;43:1567–1575.

Contents

List of abbreviations

ALL	Acute lymphoblastic leukemia
Alt	Alteration
AML	Acute myeloid leukemia
Array CGH	Array Comparative Genomic Hybridization
B-ALL	B-cell ALL
BM	Bone marrow
BCR::ABL1-L	BCR::ABL1-lymphoid involvement
BCR::ABL1-M	BCR::ABL1-multi lineage involvement
CML	Chronic myeloid leukemia
CBA	Chromosome banding analysis
Der 9	Derivative chromosome 9
Der 22	Derivative chromosome 22
ddPCR	Digital droplet PCR
FCM	Flow cytometry
FISH	Fluorescence in situ hybridization
EoT	End of treatment
GEP	Gene expression profiling
HeH	Hyperdiploidy
ICC	International consensus classification
iAMP21	Intrachromosomal amplification chromosome 21
LAIP	Leukemia associated immunophenotype
HoL	Low hypodiploidy
L-only	Leukemia-only
L/N	Leukemia/Normal
MoAb	Monoclonal antibody
MRD	Measurable residual disease
Ph	Philadelphia chromosome

Ph-like	Philadelphia-like
Ph+ALL	Philadelphia positive ALL
-r	Rearrangement
SoC	Standard of care
SV	Structural variant
T-ALL	T-cell ALL
TKI	Tyrosine kinase inhibitor
WBC	White blood cells
WHO	World health organization
WGS	Whole genome sequencing
WGTS	Whole genome-transcriptome sequencing

Introduction

Recognizing DNA as the fundamental building block of biological life was the first step towards groundbreaking discoveries in genomics and human health. The quest to resolve fundamental questions about the genome and cellular function has driven both the research community and the industry to develop advanced technologies capable of deciphering the complex structure of DNA. Today, WGS stands as the ultimate method to read the DNA sequence and map the entire genome.

WGS has proven useful in revealing unknown genetic aberrations, yet it remains a complementary tool in clinical diagnostics. One of the major areas where WGS can significantly enhance patient care is cancer diagnostics. Cancer is initiated by cells acquiring genetic aberrations, and detecting these aberrations is crucial for patient diagnosis and management. Cancer is divided into major types, hematological and solid tumors. While the role of genetic aberrations in hematological malignancies has been well established, the use of genetic diagnostics in solid tumors is a more recent development. Currently, a multimodal approach is employed to identify clinically relevant genetic aberrations in patients with hematological cancer. However, limitations in these methods prevent the detection of unknown genetic events, leaving some cases inconclusive. The following sections focus on the potential of WGS and what it can offer in the genetic diagnosis in clinical settings.

1 Background

1.1 Leukemia

Leukemia (Greek “leukos" white and “haima" blood) was first described by Doctor Rudolf Virchow in 1847 (1). He observed blood samples from patients with an abnormal amount of white blood cells and referred to this condition as “leukemia". Today we know that leukemia originates from the hematopoietic stem and progenitor cells in the bone marrow (BM), which are responsible for the production of blood cells. The first time a genomic abnormality was associated with a hematological malignancy was in 1960, when Nowell and Hungerford described an abnormal chromosome 22 in a patient with chronic myeloid leukemia (CML) and named it the Philadelphia (Ph) chromosome (2). After further technological developments, it was recognized that the Ph chromosome is the result of a reciprocal translocation between chromosomes 9 and 22(3).

Several years later, a study characterized an 8-kilobase RNA and discovered that it contained transcripts from both the *BCR* and *ABL1* genes. This study, led by Shtivelman, identified the fusion protein and suggested its role in the development of CML (4). This discovery paved the way for the development of the first targeted therapy, which involved directing a tyrosine kinase inhibitor (TKI) to the tyrosine kinase part of the BCR::ABL1 fusion protein (5).

Today, the classification of leukemia is mainly based on the type of clinical presentation and rate of progression as well as the type and maturation stage of the cell of origin. When malignant transformation occurs at the immature progenitor cells, it disrupts the production of mature, functional cells, leading to the rapid accumulation of immature cells and the development of acute leukemia. Depending on the affected cell lineage, this results in developing either acute myeloid leukemia (AML) or ALL . Consequently, the number of blasts in the BM increases and begins to move into the peripheral blood (6). The malignant cells are characterized by morphology, immunohistochemistry and by flow cytometry (FCM) using standard combinations of antibodies to determine the leukemia-associated immunophenotype (LAIP) (7). However, if cell differentiation is arrested at later stages of development, the cells remain relatively functional, leading to a slower progression of the disease, referred to as chronic leukemias (i.e. CML and chronic lymphocytic leukemia) (6).

1.2 Acute Lymphoblastic Leukemia

ALL can occur at all ages. It is the most common malignancy in children, accounting for 25% of pediatric malignancies, with the highest incidence being between 2-5 years of age (8). The median age in adults is 40 years (9). Outcomes have improved significantly over the past few decades, with survival rates exceeding 90% for pediatric patients. However, prognosis becomes less favorable with increasing age, with only 50% survival rate in adults in developed countries (10, 11). Despite this remarkable progress in pediatric survival rate, relapses and treatment failure, although rare still constitute challenges (11). In addition, long-term side effects and secondary malignancies affect patient's quality of life. It is worth noting that the survival rate in developing countries is still low due to limited access to advanced care.

The causes why children develop ALL are still unknown, however, various factors such as environmental, congenital and immunological dysregulation alone or likely in combination have been proposed to play a role (12). UV radiation or

exposure to carcinogenic chemicals are among the common environmental factors that increase the risk of developing cancer and leukemia in the adults (13). Recently, a Swedish nationwide study found that 11% of children diagnosed with solid tumors had a germline variant in a gene causing cancer predisposition (14). A follow up unpublished observation estimated <5% of leukemia cases harbored a germline variant. Studies of identical twins with concordant ALL have shown that both twins have an identical preleukemic clone and the second hit occurred in one of them (15, 16) providing evidence that the initiating lesion had occurred in utero.

ALL results from the accumulation of immature lymphoid cells in the BM and peripheral blood. It is further classified into B-ALL and T-ALL, based on whether the origin of the lymphoblast is committed to the B-cell or the T-cell lineage. Immunophenotyping, which detects cell surface markers, enables differentiation between the two lineages. LAIP is characterized at this stage to enable monitoring of response to the treatment at later stages.

Diagnosing ALL begins with the examination of BM samples, where the pathologist analyzes the morphology, immunohistochemistry and immunophenotype of the leukemic cells. This analysis helps determine the specific type of leukemia present. The initial risk assessment of ALL is based not only on genomic aberrations but also on factors such as white blood cell count (WBC) and age at diagnosis (17). To further refine the diagnosis and identify specific genetic abnormalities, additional testing is required. This typically includes cytogenetic analyses to assess chromosomal aberrations, as well as molecular testing to detect mutations or gene fusions that may be driving the disease (9). Since the initial response to treatment is closely associated with risk of relapse, most modern treatment protocols integrate residual disease measurement with genetic characterization (18). This is essential for refining risk assessment, assigning each patient to an appropriate risk category, and tailoring therapy to achieve the best possible outcome.

1.3 The Genetic Landscape of ALL

The genetic landscape of ALL has been extensively studied over the last decades and has been shown to be very heterogeneous (19). This is true for both pediatric and adult ALL however, adults harbor often high-risk genetic aberrations (20). There are three main categories of primary lesions commonly observed in ALL. First, aneuploidies, changes in chromosome numbers, define

ALL subgroups with very different outcomes. One further category is characterized by the consequences of large structural chromosomal rearrangements that occur during leukemogenesis, including translocations, deletions, and inversions. In B-ALL, these rearrangements frequently result in the formation of fusion genes, a genetic hallmark of the disease. The fusions often involve transcription factors and produce chimeric proteins that serve as primary events, driving leukemogenesis. Several recurrent fusions represent class-defining lesions in B-ALL are illustrated in Figure 1. The third category arises from structural genomic rearrangements that disrupt gene regulation patterns, leading to the deregulation of genes essential for differentiation along the lymphoid lineage. This category is particularly common in T- ALL(21).

In addition to these primary events, a few secondary lesions have been reported to impact the outcome. In the following sections well-established subgroups, emerging genetic subtypes and the secondary genetic profile of B-ALL are discussed, and the current understanding of the genetic profile for T-ALL is summarized.

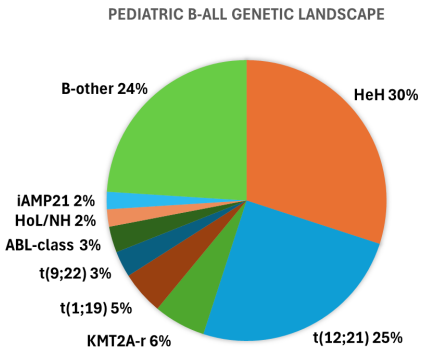

Figure 1. A pie chart that illustrates the distribution of genetic lesions mandatory to investigate according to current treatment protocol in pediatric B-ALL (unpublished data). HeH, high hyperdiploidy; *KMT2A*-r, *KMT2A* rearrangements; HoL, low hypodiploidy; NH, near haploidy; iAMP21, intrachromosomal amplification chromosome 21.

1.3.1 Genetics of B-ALL

1.3.1.1 Chromosome dosage imbalances

Aneuploidy, or numerical chromosome changes, are common recurrent abnormalities in B-ALL. High hyperdiploidy (HeH) with 51-65 chromosomes is the most common cytogenetic abnormality found in pediatric ALL (22-24). Chromosomal gains in HeH are not random, with chromosomes +4 (78%), +6 (85%), +10 (63%), +14 (84%), +17 (68%), +18(76%), +21 (99%) and +X (89%) being the most frequently affected, manifesting as trisomies and tetrasomies (25). There is evidence of a simultaneous increase of chromosomes in a single cell division causing HeH (26, 27). ALL with HeH shows excelent outcome in children when treated with standard chemotherapy, with survival rate of ≥96% at 5 years (11).

Hypodiploidies, defined as having fewer than 46 chromosomes, are also observed in ALL, but at a much lower frequency of around 5% (28). Within hypodiploidies, near-haploidy (NH, 25-29 chromosomes), mainly observed in children and low-hypodiploidy (HoL, 30-39 chromosomes), observed in young adults/adults are rare but recurrent occurring in less than 1% of cases (9). Both NH and HoL ALL have been reported to have poor outcomes and patients are therefore allocated to the high-risk genetic group. It is worth noting that often, the number of chromosomes in NH or HoL may double and appear as hyperdiploidy or NH (29). Those cases may be easily misclassified, if only the doubled clone is captured by chromosome banding analysis (CBA). It is clinically important to ensure that a duplicated NH is distinguished from hyperdiploidy since the outcome differs greatly (28, 30). The preferred method for detecting a duplicated NH is SNP arrays, which enable the identification of loss of heterozygosity (LoH) in monosomal chromosomes, even after the number of chromosomes has been duplicated (31).

Intrachromosomal amplification of chromosome 21 (iAMP21) is an aberration detected in 2% of pediatric ALL cases. Cytogenetic examination revealed that iAMP21 shows a variety of morphological changes which can be the result of complex intrachromosomal rearrangements created by breakage-fusion-bridge (32). Patients with iAMP21 have a higher risk of relapse and are regarded as high-risk (25).

1.3.1.2 Recurrent gene fusions in ALL

ETV6::RUNX1 is among the most common alterations in pediatric ALL, encompassing approximately 25% of cases (33-36). Both *ETV6* and *RUNX1* are transcription factors crucial for hematopoiesis. The translocation between chromosomes 12 and 21 brings these two genes together and results in a fusion gene *ETV6::RUNX1* on the derivative chromosome 21 (37, 38). Studies showed that monozygotic twins that developed ALL at different time points, harbored the same *ETV6::RUNX1* (39). In addition, the presence of this fusion in neonatal blood spots from children diagnosed with t(12;21) ALL provided evidence that this fusion occurred years before ALL developed and likely represents the initiating event in leukemogenesis. (16, 39, 40). The *ETV6::RUNX1* fusion can result from simple rearrangements involving only chromosomes 12 and 21, as well as from more complex genetic rearrangements, which may also involve near triploidy or tetraploidy (41, 42). Regardless of the underlying genetic complexity, this subtype is associated with a favorable response to treatment in children (34).

The recurrent translocation between chromosome arms 1q and 19p fuses two genes encoding transcription factors, and results in the *TCF3::PBX1* fusion which is the class-defining lesion for this subgroup of ALL(43-45). Both *TCF3* and *PBX1* are essential for the commitment and self-renewal of hematopoietic stem cells and for the maturation of lymphoid cells. The aberrant fusion protein causes differentiation arrest and constitutes the class-defining lesion (46). *TCF3::PBX1* is found in 5% of ALL cases overall. Originally classified as an intermediate risk category, the result for this subgroup is comparable to the standard genetic risk in modern treatment protocols (10, 47).

A rare translocation, t(17;19), involving *TCF3* and hepatic leukemia factor gene *HLF*, produces a chimeric transcript which defines another subtype of ALL with extremely poor outcome (48). Thus, patients harboring *TCF3::HLF* fusions are classified as high-risk, regardless of their initial response to treatment.

Rearrangements of *KMT2A* are found in several leukemia types including ALL(49). The *KMT2A* gene is located on the long arm of chromosome 11 and encodes a histone methyltransferase that plays an important role during embryonic development and hematopoiesis. Rearrangements of *KMT2A* define a subgroup of ALL and lead to the formation of fusion genes with more than 90 partners described so far (50) The most common partners found in infant and pediatric

ALL include *AFF1*, *MLLT1* and *MLLT3*, followed by *MLLT10*, which together are responsible for over 90% of *KMT2A*-r infant ALL and about 75% of *KMT2A*-r pediatric ALL (50). *KMT2A*-r have been associated with a poor outcome in most ALL trials in children (10, 51).

A recent study described an ALL subgroup with tyrosine kinase rearrangements involving *ABL1* (other than *BCR::ABL*), *ABL2*, or *CSFR/PDGFRB* referred to as *ABL*-class, that make blasts responsive to tyrosine kinase inhibitors (TKIs) (52). Therefore, under the current protocol, it is essential to investigate theses rearrangements since patients with these alterations will receive TKIs as part of their treatment. The *ABL*-class accounts for 3-5% of pediatric ALL cases (53).

As previously mentioned, *ABL1* is one of the genes involved in the translocation of t(9;22)(Figure 2). The *BCR::ABL1* fusion gene encodes a constitutively active tyrosine kinase which is the defining lesion in Ph+ALL as well as in CML (54, 55). CML occurs mainly in adults and only rarely in children. In normal cells, *ABL1* is involved in cell survival, differentiation and growth by phosphorylating target proteins in the RAS/MAPK, PI3K/AKT and JAK/STAT pathways (54). Two main isoforms of fusion proteins have been found. A 190 kDa protein (p190) translated from the minor transcript (m-*BCR::ABL1*), in which exon 1 of *BCR* is juxtaposed to exon 2 of *ABL1*, is found in 2/3 of ALL patients (56). The 210 kDa isoform (p210) is translated from the major transcript (M-*BCR::ABL1*), where exon 13 or 14 of *BCR* is juxtaposed to *ABL1* exon 2, and is characteristic of CML (54). While the *BCR::ABL1* fusion occurs in 2-3% of pediatric cases, its prevalence increases with age and is found in 25% of adult ALL cases (57, 58). Since the discovery of TKIs that directly target the kinase domain of the *BCR::ABL1* protein outcomes for Ph+ALL have improved. (5). Currently, only patients who do not achieve a satisfactory response to combination therapy require stem cell transplantation (59). One of the key factors in managing Ph+ALL patients, as with ALL in general, is the careful assessment of the initial response to therapy, specifically MRD, which serves as an important prognostic indicator. A more detailed discussion of MRD quantification in ALL can be found in the MRD section.

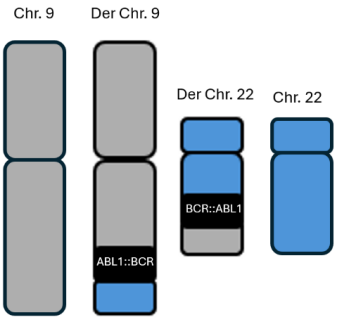

Figure 2. Illustration of reciprocal translocation between chromosomal arms 9q and 22q. The *BCR::ABL1* fusion gene is on the der 22 also known as Philadelphia chromosome. Der 9, derivative chromosome 9; der 22, derivative chromosome 22.

1.3.1.3 Emerging subtypes in B-cell ALL

B-ALL patients without established class-defining aberrations are referred to as B-other ALL. In the last decade, new sequencing technologies, including whole genome and transcriptome (WGTS), have made it possible to discover new genetic subtypes in B-ALL. Some of these new findings have been suggested to have potential prognostic significance and/or can be targeted for treatment (60). The latest WHO 2022 classification and ICC have included some of these new findings as new subtypes of B-ALL (61-63). Although these new subtypes have not been included in the treatment protocol, preliminary data from small studies suggest that the information on these new subtypes may be relevant for patient management (64-67). Some of the new subtypes of B-ALL are presented below (Figure 3).

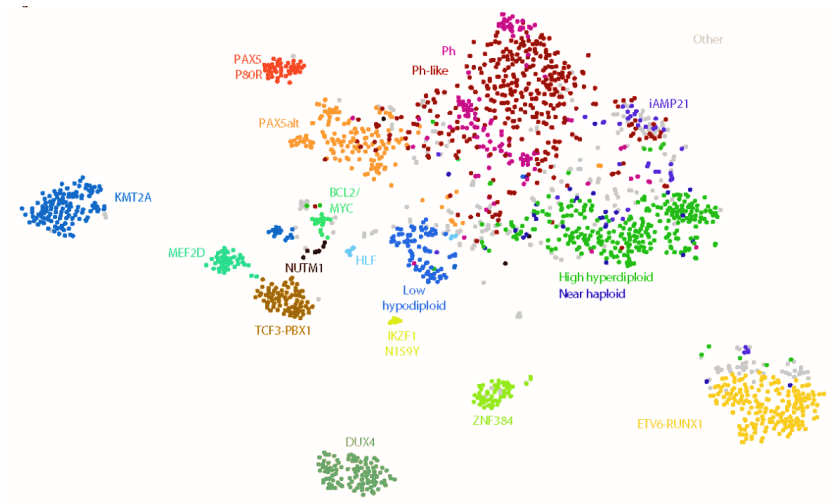

Figure 3. Gene expression profiling (GEP) of B-ALL cases. GEP is performed on 1,988 B-cell ALL showing the clustering of genetic subgroups (64).

In a comprehensive transcriptome profiling study, Gu and colleagues discovered a subgroup of ALL patients that harbored rearrangements of the *MEF2D* gene with any of six potential partner genes (*BCL9*, *CSF1R*, *DAZAP1*, *HNRNPUL1*, *SS18*, and *FOXJ2*) (65). They showed that *MEF2D* rearranged (*MEF2D*-r) ALL clusters in transcriptome profiling and in their study *MEF2D*-r ALL showed inferior overall survival rate compared to other ALL subtypes. The rearrangement of *MEF2D* enhances its activation and subsequently of, a histone deacetylase gene, *HDAC9*. Further, they suggest that a histone deacetylase inhibitor can be used as a potential targeted therapy on *HDAC9* for patients with *MEF2D*-r.

In the same study, they also identified another cluster of patients harboring *ZNF384* rearrangements (*ZNF384*-r) with six potential partner genes (*ARID1B*, *CREBBP*, *EP300*, *SMARCA2*, *TAF15*, and *TCF3*) (65). Further characterization revealed that patients with the *TCF3::ZNF384* fusion had an inferior outcome compared to those with *EP300::ZNF384* or wild-type *ZNF384* (66). Furthermore, *ZNF384*-r has a unique immunophenotype and it is proposed as a new subgroup of ALL (61).

Another interesting subgroup identified by transcriptome profiling and subsequent clustering is ALL carrying *DUX4* rearrangements (*DUX4*-r). *DUX4* is a transcription factor that is involved in stem cell programming and is silenced in somatic cells. However, the gene reactivates in a subgroup of ALL through

rearrangement with the *IGH* locus. This rearrangement locates the *DUX4* under *IGH* enhancer while altering the DUX4 protein by disrupting its C terminus all of which results in oncogenic activity. In an *in vivo* study in mice, the expression of *DUX4*::*IGH* transformed pro B-cells into leukemic cells and the authors concluded that this rearrangement is capable of inducing leukemogenesis (67). In addition, *DUX4* overexpression due to the rearrangement, binds within intron 6 of *ERG* gene and induces expression of a novel *ERG* isoform (*ERG*alt). ERGalt disturbs *ERG* activity and thus promotes leukemogenesis. Also, DUX4 protein binding to *ERG*, potentially affects the chromatin accessibility to recombination-activating gene (*RAG*) endonucleases that can induce the deletion of *ERG* (68, 69). Hence the deletion of *ERG* has been proposed as an indicator to *DUX4*-r.

Aberrations of *CRLF2*, encoding a kinase involved in the JAK-STAT pathway, have also been recurrently detected in ALL. In those cases, the expression of *CRLF2* is elevated either by abnormal rearrangements with the *IGH* locus (*IGH*::*CRLF2*) or by a deletion in the pseudo-autosomal region 1 (PAR1), resulting in *CRLF2*::*P2RY8* fusion. In these scenarios, *CRLF2* is controlled by either the *IGH* or the *P2RY8* promoter region respectively, which causes the overexpression of the gene (70). Deregulation of *CRLF2* has been observed to occur in higher frequency in Down syndrome-ALL patients and in Ph-like B-other ALL (71).

PAX5 encodes a transcription factor essential for B lymphoid development and lineage commitment (72). An integrative genomic analysis of pediatric and adult ALL described two new subtypes involving *PAX5*. This study revealed two clusters defined by distinct genetic lesions: i) intragenic deletions and amplifications or interchromosomal rearrangements affecting *PAX5* (*PAX5*-alt), and ii) the single nucleotide variant p.Pro80Arg or biallelic PAX5 alterations. They conclude that disruptions in *PAX5* deregulate B-cell differentiation and represent an early event in B-cell ALL (64). It is, however, not completely established to what extent all *PAX5* aberrations represent primary lesions.

Gene expression profiling (GEP) studies revealed that a significant number of B-other ALL samples did not exhibit the canonical fusion characteristic of their respective cluster. Thus, a fraction of samples clustered together with Ph+ALL but lacked the *BCR*::*ABL1* fusion. This phenomenon led to the identification of *BCR*::*ABL1*-like or Philadelphia-like (Ph-like) ALL, a subgroup whose expression profile closely resembles that of Ph+ALL yet lacks the *BCR*::*ABL1* fusion. Later studies showed that aberrations affecting other tyrosine kinase genes such as

ABL2, *CRLF2*, *PDGFRB*, *EPOR* or *JAK2* alternatively *ABL1,* with partners other than *BCR*, were involved in this subtype (73-75). Similarly to Ph+ALL, patients with a Ph-like profile responded poorly to conventional therapy regimens (75). Hence, TKI have been added to the treatment of the Ph-like patients.

Another subgroup of ALL with an expression profile like *ETV6::RUNX1* ALL but lacking the fusion was also identified with transcriptome profiling and referred as *ETV6::RUNX1*-like (76). However, additional studies are needed to investigate the clinical relevance and prognostic implications of *ETV6::RUNX1*-like ALL.

1.3.2 Secondary genetic events

In recent years, the role of secondary genetic events in the development of ALL has gathered significant attention. Intragenic deletions of *IKZF1*, initially identified as a hallmark of Ph+ B-ALL (77), were the first to be associated with an unfavorable prognosis (73, 78, 79). These findings were later validated by several research groups, confirming the prognostic significance of *IKZF1* deletions in B-ALL (80, 81). Furthermore, a copy number classifier (CNAs) based on *IKZF1,* and seven further genes recurrently deleted in ALL was developed, known as the UKALL-CNA classifier (82, 83). The genes/regions included in this classifier are *IKZF1*, *EBF1*, *RB1*, *ETV6*, *PAX5*, *BTG1*, *CDKN2A/B*, and a deletion in the pseudoautosomal region 1 (PAR1) that results in a *P2RY8::CRLF2* fusion. By integrating UKALL-CNA status with cytogenetic risk groups, Moorman et al. proposed a modified genetic risk classification (82). In addition, Stanulla and colleagues described a different CNA profile, which includes *IKZF1* deletions along with deletions in *CDKN2A*, *CDKN2B*, *PAX5*, or PAR1, but without concomitant *ERG* deletion. This profile, referred to as *IKZF1*plus, was then combined with MRD status to refine the risk groups (84).

1.3.3 Genetics of T-ALL

T-lymphoblasts are another type of cells that can develop into ALL (T-ALL) and represent 15 and 25 % of pediatric and adult ALL respectively (85, 86). This form of leukemia is aggressive and shows considerable genetic heterogeneity (21). Within T-ALL, the WHO classification identifies one distinct subgroup based on immunophenotypic markers: early precursor T-ALL (ETP-ALL). The only genetic entity described by ICC is *BCL11B*-r (61, 62). Unlike B-ALL, where genetic information plays an important role in risk stratification, T-ALL lacks universally recognized genetic markers that provide prognostic insights. However, a few recurrent aberrations, such as *KMT2A* rearrangements, *ABL*-class

rearrangements, and *BCR::ABL1* fusions, are also observed in T-ALL and are routinely investigated as they impact therapy. The latter two aberrations can be targeted with TKIs (21, 87).

Recently, large-scale GEP studies have identified subgroups within T-ALL, with a recurring theme of deregulating in transcription factors, critical for lymphocyte differentiation and proliferation (88-90). Based on recent GEP and WGS studies the ICC describes a number of provisional entities for T-cell ALL (91). These include *TAL1/2*-r (92, 93), *TXL1/3*-r (94, 95), *HOXA* dysregulations, *LMO1/2*-r (96), *BCL11B*-a (97), *SPI1*-r, *NKX2*-r (98) and *BHLH* (21). ETP-ALL is characterized by abnormal activation of the *BCL11B* gene (*BCL11B*- alteration) (99, 100).

Detecting many of these subgroups is challenging, even with RNA sequencing, as the rearrangements occur in non-coding regions and lead to gene deregulation rather than expressed gene fusions (90). Detection by WGS can be difficult as well, as the lesions underlying the GEP subgroups remain incompletely defined. Moreover, high-throughput methods like WGS and GEP clustering are not widely accessible and not validated for diagnostic purposes. A recent review provides comprehensive guidelines for detecting these potentially relevant lesions in T-ALL using more widely available techniques, such as CBA and FISH analysis (101). Clinical studies are essential for determining the clinical impact associated with genetic subgroups in T-ALL, while further investigations are needed to better understand the genomic lesions underlying these subgroups.

1.4 Genetic characterization of ALL in the clinical setting

The genetic techniques used to characterize leukemia samples have advanced alongside technological developments in genome research. Initially, the detection of leukemic cells in ALL relied on morphological examination under the microscope. The discovery of the Ph chromosome in CML in the 1960s marked the beginning of cancer genetics (2), and of leukemia genetics in particular.

CBA with a resolution of 5-10 megabases can reveal large scale chromosomal aberrations such as aneuploidies, translocations, inversions and deletions. However, the estimation of breakpoints is inaccurate, and is unable to pinpoint the affected genes. FISH analysis is a targeted approach with a resolution of 50-100 kilobases and represents a targeted analysis used to investigate specific lesions. During 2000s, array CGH was developed which provided higher

resolution for more precise mapping of breakpoints and the identification of aneuploidies, chromosomal imbalances and minor gains and losses (102).

With the advent of the era of molecular genetics, Sanger sequencing and RT-PCR were applied to detect lesions at base pair resolution and allowed the identification of genes involved in leukemogenesis. In the diagnostic setting of ALL molecular methods are also used to examine specific genes, particularly fusions. At the end of the 20th century, gene expression and methylation arrays began to unravel further layers of lesions involved in the development and progression of ALL (103).

Today, in clinical diagnostics, due to the diverse nature of ALL genetic lesions, a combination of cytogenetic and molecular analysis is used. While the cytogenetics methods, CBA, FISH and array CGH detect large chromosomal abnormalities, molecular methods detect point mutations, gene fusions and smaller insertion/deletion/inversion alterations. Although these methods are well established, the use of different approaches not only requires experts with different skills but also increases the turnaround time. With the advent of massive parallel sequencing (MPS), the simultaneous detection of many genetic lesions becomes possible.

Currently, WGS is only used in a centralized manner in a few countries in the diagnostic setting and also to a limited extend as a complementary method to investigate inconclusive cases (104). Therefore, the full potential of WGS and transcriptome sequencing in the clinical diagnosis of ALL remains to be tested. In this work, the feasibility of WGS in the diagnosis and monitoring of pediatric ALL is investigated.

1.5 Measurable residual disease

Around the same time that treatment for ALL became more intensive and survival rates began to improve, the concept of monitoring MRD, was introduced. As technology advanced, FCM and PCR-based methods were developed, allowing for highly sensitive quantification of MRD. These advanced tools enabled researchers to establish a link between the persistence of leukemic cells after induction therapy and future outcomes (105).

Today, FCM is a well-established method for detecting residual leukemic cells by labelling cell surface markers with monoclonal antibodies (MoAb). Due to its widespread use in diagnostic laboratories, the EuroFlow Consortium was

established to create guidelines and standardize quantification methods (7). Currently, the approach employs eight monoclonal antibody (MoAb) combinations to identify leukemic cells based on their LAIP. Each MoAb targets a specific cell surface marker, aiding in the identification of cell types involved in ALL and allowing for the quantification of leukemic cells in the sample. With a sensitivity of finding 1 leukemic cell per 10,000 cells (1^{-4}) and a short analysis time, FCM was the first method used to measure MRD (106). However, some patients may lack informative LAIPs, or the LAIP may change during treatment, making it impossible to quantify residual disease using FCM (107).

Another approach to quantify MRD is to identify the clonal IG/TR gene rearrangement present at presentation in the leukemic blasts which can later be measured by quantitative PCR (qPCR) during follow up (108). During the maturation of B and T cells, the immunoglobulin (IG) and T cell receptors (TR) genes undergo physiological rearrangements to generate the entire antigen receptor (IG/TR) repertoire. Since ALL develops from a single blast cell with a unique IG/TR sequence, the resulting ALL population has the same IG/TR sequence that can be used as an MRD marker (109). To reduce inter-laboratory variation and achieve comparable MRD across different treatment protocols, the European Study Group on MRD detection in ALL (ESG-MRD-ALL) has initiated and developed guidelines and standards for PCR-based MRD quantification (ESG-MRD-ALL). The guidelines contain suggestions for the experimental setup, how to determine the quantitative/sensitivity range, how to define MRD positivity and negativity as well as for the quantification of follow-up samples (110). The absolute sensitivity of PCR-based MRD quantification is usually high and depends primarily on the amount of DNA input. One important limitation of the method is however that sometimes the leukemia-associated IG/TR rearrangements are not very different from physiological rearrangements within the immune repertoire, and this significantly may limit the sensitivity of particular assays. Also, it has been observed that the IG rearrangement may change during the course of the disease and not be present at relapse(111).

In Ph+ALL, the methods mentioned above are often supplemented with the quantification of the *BCR::ABL1* fusion transcript using real-time RT-PCR (112, 113). While it is widely used to monitor MRD in Ph+ALL adult patients (112, 114, 115), in pediatric FCM and IG/TR rearrangement appeared to be more reliable to predict outcome (116). However, by comparing the results of FCM, IG/TR and *BCR::ABL1*

transcript, studies have shown that quantification of transcript was more sensitive to predict a relapse (116, 117).

Interestingly, the term CML-like Ph+ALL has been suggested for a subgroup of Ph+ALL who showed higher MRD measured by *BCR::ABL1* transcript level than by IG/TR rearrangement as well as showed a multi-lineage involvement (117, 118). It is worth noting that while measuring *BCR::ABL1* by RT-PCR offers high sensitivity and specificity, it cannot estimate the number of malignant cells. Instead, it provides an average measure of fusion gene expression in the sample. A new generation of quantitative PCR, droplet digital PCR (ddPCR), is being increasingly adopted in clinical settings, enabling the absolute quantification of targets without the need for cumbersome dilutions of positive samples to create a standard curve. Instead, ddPCR statistically calculates the absolute copy number of *BCR::ABL1*, and a reference gene in parallel, allowing for the accurate determination of the percentage of *BCR::ABL1*-positive cells in the sample (119, 120).

2 Research aims

The overall aim of this thesis is to assess the feasibility of WGS application and to improve understanding of chromosomal aberrations in pediatric ALL.

Specific aims:

1) To investigate the feasibility to use WGS as the sole method to detect genomic aberrations in pediatric B-ALL in the diagnostic setting.

2) To test the hypothesis whether the junctions left by the genomic rearrangements driving leukemia and detected by WGS can be exploited as patient specific targets to quantify MRD

3) To characterize Ph+ALL by integrating immunophenotypic, genomic and transcriptomic features together with MRD assessments.

3 Materials and methods

3.1 Patient samples

Paper I, a cohort consisting of 88 retrospective pediatric B-ALL samples treated according to the NOPHO trials at Uppsala University and Karolinska University Hospitals, Sweden, was investigated. We aimed to include samples representing all mandatory aberrations (n = 49). The remaining 36 samples lacked mandatory aberrations, i.e. represented B-other ALL, and three samples corresponded to Down syndrome ALL.

Paper II, BM DNA from a child with T-ALL, who failed first line therapy was analyzed by WGS.

Paper III, six samples corresponding to four B-ALL and two T-ALL cases were included in a proof-of-concept study at Karolinska University Hospital in Sweden. Three B-ALL patients belonged to B-other sub-group and one to HeH ALL DNA from diagnostic and follow-up BM samples as well as cell-free DNA (cf-DNA) isolated from plasma and cerebrospinal fluid (CSF) were analyzed in this study. Blood and CSF were collected in BCT tubes (10 ml) (STRECK, La Vista, NE, USA) and cfDNA, once isolated, was stored at -20°C for a maximum of 3 months.

Paper IV, leukemia samples where the standard of care (SoC) genetic investigations had identified the *BCR::ABL1* gene fusion were retrieved from the Karolinska University Hospital, Stockholm, Sweden. Samples from pediatric and young adult patients were prioritized. A total of 15 patients diagnosed with Ph+ALL (median age 12 years, range 3-60 years).

3.2 DNA and RNA extraction

In all papers, depending on the amount of BM or blood material available, different extraction methods were applied. Briefly, Genomic DNA was isolated using Tissue kit and EZ1™ automated instrument (Qiagen, Hilden, Germany) or extracted manually together with RNA from frozen cell pellets with the AllPrep DNA/RNA Mini Kit Qiagen (Qiagen, Hilden, Germany). For samples with small amount of frozen BM taken at remission, DNA was extracted with a Gentra Puregene Blood core kit (Qiagen, Hilden, Germany). DNA was stored at -20°C and RNA at -80°C until use.

RNA purification in **paper IV** was done with RNase-Free DNase Set, Qiagen, according to the manufacturer's guidelines.

3.3 Sequencing and analysis

Paper I, patients were divided into an exploratory and a validation set. The exploratory set was sequenced at the National Genomics Infrastructure Science for Life Laboratory, Stockholm (n = 36) and Uppsala (n = 22), Sweden. The validation set (n = 30) was sequenced at Clinical Genomics, SciLifeLab, Stockholm, Sweden. The details on input DNA and sequencing are provided in Table 1 and Figure 1. Samples were sequenced either with 90x leukemia samples or 30x coverage for normal/germline on Illumina NovaSeq 6000 platform (Figure 1).

In this study, the accuracy of detecting the obligatory aberrations required by the ALLTogether study protocol was investigated using WGS. We evaluated two analysis approaches: leukemia-only (L-only) and leukemia/normal (L/N). In both approaches, to reduce complexity, the detected variants were first filtered using a local database containing normal variants from unaffected individuals, followed by a positive filter based on a list of loci relevant to ALL. The gene list included all genes that are mandatory according to the current ALL treatment trial protocol, genes included in the UKALL CNA classification and key genes in B-other ALL or included in the WHO 2022 classification as well as the International Consensus Classification (ICC) of hematological malignancies (61, 62). For L/N, the variants of each leukemia case were first filtered against the normal counterpart, followed by the application of the aforementioned filters. To maximize the chance of detecting recurrent aberrations, the leukemia samples (n = 64) were sequenced with 90x coverage, while the paired normal samples were sequenced with 30x coverage as is standard for the detection of germline conditions (121, 122). Later, the results of L-only and L/N were compared with each other and with the results of SoC. In addition, the potential of 30x coverage was investigated by sequencing 30 leukemia cases and down-sampling an additional 10 cases that had originally been sequenced with 90x coverage. A schematic chart illustrates the distribution of WGS coverage and filtering steps in Figure 1.

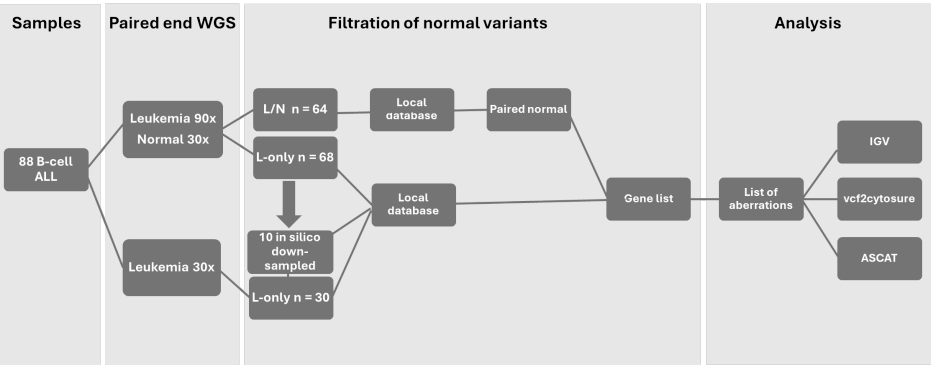

Figure 1. WGS coverage distribution and down-stream analysis in **paper I**.

Table 1. WGS sample distribution in **paper I**.

	Exploratory set (n = 58)	Validation set (n = 30)	Library	Instrument
Input DNA 1 µg	36	-	TruSeq DNA PCR-free	HiSeq X platform
Input DNA 100 ng	22	-	TruSeq DNA Nano protocol	HiSeq X platform
Input DNA 200 ng	-	30	NxSeq® AmpFREE Low DNA Library Kit	NovaSeq 6000

Paper II, **III** and **IV**: WGS was performed at Clinical Genomics, SciLifeLab, Stockholm, Sweden. Briefly, 1.1 µg genomic DNA isolated from diagnostic BM was sequenced by PCR-free, paired-end WGS protocol. Samples in papers **II and III** sequenced on an Illumina HiSeq X platform, and **paper IV** on Illumina NovaSeq 6000 platform in, with 30x coverage.

The analysis of the sequencing data in **paper I** was performed with Sarek (123). **Papers II** and **III** was analyzed using the MIP (124) from Clinical Genomics. In **paper II**, we aimed to detect the existence of potential SNVs or SVs that might be involved in the development of T-cell ALL. Therefore, both SNV callers, GATK HaplotypeCaller, and SV callers, Delly (125)and TIDDIT v3.0.0 (126), were used. In **paper III**, the focus was on SVs, as the study aimed to identify driver lesions and develop patient-specific assays. CNVnator V0.3.2 (127), TIDDIT V2.0.0 (126)and Manta (128) were used for this purpose. In **Paper IV**, the data was processed by the BALSAMIC pipeline (129) (a pipeline developed for detection of somatic variants) which contained and TIDDIT V2.0.0, Manta, CNVnator and vcf2Cyrosure (130).(Table 2).

For manual inspection of the variants, CytoSure™ Interpret Software (Oxford Gene Technologies, Oxford, UK), Integrative Genomics Viewer (IGV) and ASCAT were used.

Table 2. Bioinformatics tools used on each paper.

	I	II	III	IV
SNV calling	Manually	GATK HaplotypeCaller	NA	Manually
SV calling	Manta v1.6.0, Delly TIDDIT v3.0.0	Delly Manta TIDDIT v3.0.0	Delly TIDDIT V2.0.0 Manta	TIDDIT V2.0.0 Manta
Aneuploidies and CNAs	CNVnator vcf2cytosure v0.7.1 ASCAT version 4.5.0	CNVnator	CNVnator	CNVnator vcf2cytosure v0.7.1
DUX4 detection	Samtools command	NA	NA	NA

3.4 RNA-sequencing and analysis

Paper IV, purified RNAs originating from the diagnostic BM (n = 14) or blood (n =1) samples of patients diagnosed with Ph+ALL as well as the relapse sample (P02) and the molecular relapse sample (P04) were sequenced. 300 ng input RNA was prepared with Illumina Stranded mRNA Library Preparation. and sequenced Paired end 2x150bp on the NovaSeqX sequencing platform (Illumina) at Clinical Genomics, SciLifeLab, Stockholm, Sweden.

In this study we aimed to utilize the bulk RNA-seq data on 15 Ph+ALL cases and attempt to allocate them into one of the sub-groups, lymphoid or multi-lineage. Initially, RNA-seq data was analyzed through nf-core/rnafusion (131). A bioinformatic tool that is specifically directed towards Ph+ALL was developed by Beder et al. ALLCatchR-bcrabl1 (132). This tool is based on the original ALLCatchR tool, which is a machine learning algorithm to classify B-ALL into subtypes, trained and validated on several independent transcriptome datasets.

3.5 Droplet digital PCR

Paper III, the limit of detection (LoD) and limit of quantifiability (LoQ) were assessed for each assay by a serial dilution of gDNA diluted in normal control ranging from 10^{-1} to 10^{-6}. After determining the specificity and sensitivity of the assays, targets were measured in BM gDNA, cfDNA from blood and CSF.

Paper IV, the junctions created by t(9;22)(q34;q11) on der 9 with the *ABL1::BCR* fusion and der 22 with the active *BCR::ABL1* fusion were identified using WGS data. Separate ddPCR assays were designed towards each junction. These assays allowed detection of each junction separately.

3.6 Ethical considerations

This thesis project is based on genetic data obtained from WGS. This raises several ethical questions that are important to reflect upon.

Firstly, the generation of whole genome data raises significant concerns regarding traceability and data security. Since the data encompasses an individual's entire genome, it is crucial to ensure that it is pseudonymized before being stored in databases. Access to this data, along with the individual's disease history, should be restricted to healthcare providers or researchers with ethical permits. Both the sequencing facility and the hospital have the responsibility of implementing secure data management systems and maintaining highly secure databases for storing patient data. However, it is important to acknowledge that no security system is entirely foolproof. Hackers with specific interests may attempt to breach these databases. Therefore, routine maintenance and updates are imperative to keep the servers secure.

The second concern is the identification of unexpected germline genetic abnormalities. This raises the recurrent question of which unsolicited genetic abnormalities should be reported to patients and responsible clinicians. The answer is not a simple "Yes, we do report!" or "No, we do not." Each case is unique and requires careful and thoughtful ethical discussion. During these discussions, all aspects of the outcome should be considered, including the patient's disease history, the outcome of the primary disease, age, and more, to make the most suitable decision. Even after thorough ethical deliberation, it can sometimes be challenging to arrive at the right decision.

A third ethical aspect is the immediate benefit of this study for participants. Since our studies included all patients retrospectively, participants did not benefit directly from the results of this project. However, the data generated from our studies provides a valuable foundation for prospective research and the implementation of WGS in routine clinical laboratories. Our studies indicate that WGS, as a standalone method, performs better than the current multi-method procedure, refining genetic diagnostics and allowing for adjustments to

therapy regimens. This raises the ethical question of whether it is appropriate to continue using our current methodology instead of adopting WGS. One important consideration is the cost-benefit balance, and preliminary health economic studies suggest that the increased costs associated with WGS are justified by its superior performance. Given that WGS is an unbiased method for studying the human genome, it has significant diagnostic potential. My research aims to realize part of this potential.

4 Results and discussions

4.1 Paper I

In this paper we assessed the diagnostic yield and accuracy of WGS as sole method to detect genomic lesions of clinical relevance in pediatric ALL

4.1.1 Detection of mandatory fusions and copy-number alterations

In the first step, we evaluated the diagnostic accuracy of whole-genome sequencing (WGS) as a stand-alone method by analyzing 38 B-ALL cases with various mandatory genetic events, as outlined in the ALLTogether trial. To reduce data complexity and efficiently identify relevant lesions, we applied a final filtering step to the variant files, utilizing a predefined list of pertinent genes and genomic regions. This process narrowed the number of variants to approximately 20-25 per sample, which were then manually inspected using the visualization tools IGV (133), CytoSure™ Interpret Software, and ASCAT(134) (Figure 1). The latter two tools were particularly effective in visualizing aneuploidies, as well as large gains and losses. We were able to successfully detect cases with HeH, HoL, and iAMP21, and also distinguish trisomies and tetrasomies, with results consistent across analyses of L-only and L/N samples, as well as SoC results. However, a challenge with CytoSure arises when interpreting cases with near-haploid ALL (n±, <30 chromosomes) or near-tri/tetraploidy (3n±/4n±, 58-80/81-103 chromosomes), as CytoSure averages the signals from all chromosomes to set the diploid threshold and does not assess loss of heterozygosity. This can lead to misclassification of these samples as HeH. Being aware that losses and gains in specific chromosomes are not random can help mitigate this issue (Safavi et al., 2018). Using bioinformatic tools that can detect heterozygosity from WGS data, (135), in a way similar to SNP arrays, can fully resolve this challenge.

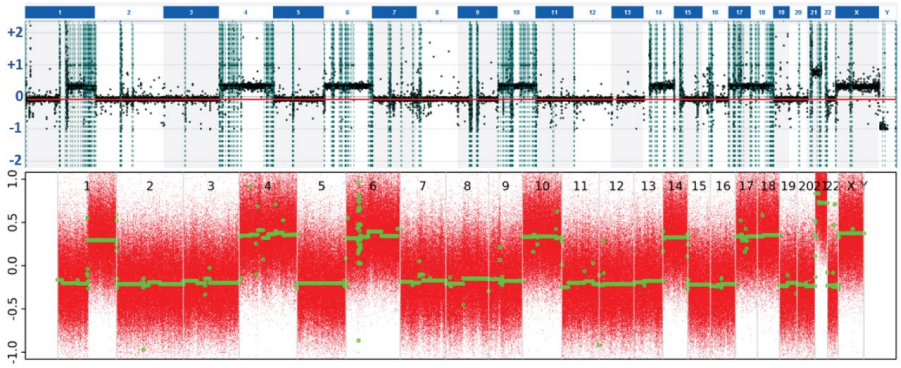

Figure 1. Graphs illustrating ploidy changes in HeH ALL. WGS data is visualized by Cytosure (upper panel), and allele-Specific Copy number is shown by ASCAT (lower panel). Gain of chromosomes 4, 6, 10, 14, 17, 18, 21 and X are visible in both panels (136).

Furthermore, relevant gene fusions were sought among the lists of variants identified by WGS in all B-ALL samples. The *ETV6::RUNX1* and *TCF3::PBX1* fusions were identified successfully by both L-only and L/N analysis. However, in the L/N analysis, only four out of six *KMT2A*-r were detected. Further investigation showed that the paired normal samples used for the L/N analysis had been collected shortly before both patients relapsed and the events had therefore been filtered out as the normal sample contained cells with the *KMT2A*-r. On the other hand, all six *KMT2A*-r were found through L-only analysis. Two Ph+ALL cases with the t(9;22)(q34;q11) and two *ABL*-class rearrangements were also detected in four further samples. However, one *ABL*-class ALL harboring a *RANBP2::ABL1* fusion was not picked up by the filter. Manual inspection in IGV revealed that the breakpoint on *ABL1* was located upstream of the gene and the discordant reads supporting the rearrangement had not been retained by the filter. By replacing the gene ID with the genomic coordinates adding 5 kb upstream of the *ABL1* gene, the rearrangement was retrieved by the filter. Moreover, 10 additional samples were processed through the clinical platform pipeline (SCOUT, a VCF visualization interface, Clinical genomics, Scilifelab, Stockholm, Sweden) (137) and analyzed in a blinded manner to test that diagnostic accuracy. The results completely aligned with L-only, L/N and SoC findings. In summary, our approach efficiently and correctly detected all mandatory aberrations (Figure 2).

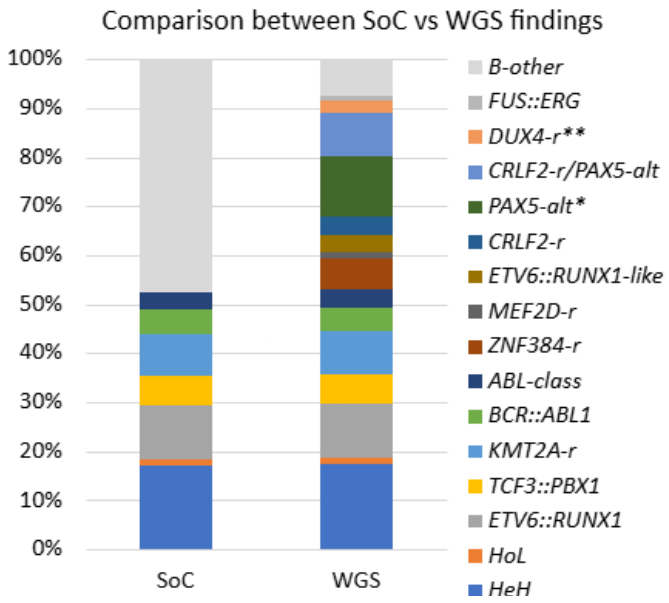

Figure 2. Bar diagrams showing the comparison of class-defining genetic lesions detected by SoC and WGS in B-ALL (modified from Rezayee et al Frontiers Oncology 2023)(136).

4.1.2 Diagnostic yield in B-other ALL

We also aimed to identify potentially clinically relevant subsets within the samples that lacked the established mandatory lesions in B-ALL. To achieve this, we analyzed 36 B-other ALL samples and three DS-ALL samples, focusing on identifying recurrent aberrations that may define subgroups within the B-other group or influence patient outcomes.

The most frequent aberration amongst the B-other ALL was rearrangements that affected *PAX5* (n = 17), such as *PAX5* fusions and intragenic deletions or amplifications followed by *CRLF2* rearrangements (n = 10). The observation that the haploinsufficiency of *PAX5* was the sole potentially class-defining lesion identified in several samples, suggests that it may be a primary driver of leukemogenesis in some B-ALL cases. On the other hand, its co-occurrence with other established lesions suggests that it contributes as a secondary alteration and influences disease progression, as has been proposed by Gu and coworkers (64). We found that *PAX5* rearrangements often co-occurred with *CRLF2*

rearrangements (*CRLF2*-r). *CRLF2* can rearrange with the *IGH* locus due to a t(X;14)(p22.33;q32) translocation or with the *P2RY8* locus (*P2RY8::CRLF2*) as a result of a deletion in the PAR1 region. Additionally, *CRLF2*-r was identified as an isolated lesion in the three DS-ALL cases in our study, a lesion that occurs in approximately half of ALL cases within this patient group (138). Furthermore, WGS identified various other rare recurrent fusions, including *ZNF384*-r, *ETV6::IKZF1*, *MEF2D::BCL9* fusions, *IGH::ID4* and a very rare *FUS::ERG* (Figure 2).

In this cohort, we included one sample that according to RNA-seq (139) harbored a *DUX4*-r. Our initial analysis, only called a structural variant in *IGH* but no discordant reads mapped to *DUX4*. A In addition, a *ERG* deletion, strongly associated with *DUX4*-r, was detected in another sample. However, no *IGH::DUX4* rearrangement was captured in the WGS data by the filter in neither sample. The likely reason for this is the existence of multiple copies for *DUX4* embedded in a highly repetitive D4Z4 region in addition to pseudogenes scattered throughout the genome and making the discordant reads impossible to accurately align to the reference genome (140). Consequently, the discordant read pairs will be discarded by the callers. To address this challenge, we employed a targeted bioinformatics approach that focused exclusively on discordant reads linking the *IGH* locus to any of the *DUX4* genes or pseudogenes and returned the number of such reads for each sample. Using this method, we successfully identified a significant number of discordant reads between these genomic regions in the two samples suspected to contain the *IGH::DUX4* rearrangement, as well as in two additional samples, representing a total of 4% of B-other ALL cases. In contrast, no or only isolated discordant reads were found in the other 85 samples, confirming the specificity of our approach. Interestingly, other studies have reported a higher frequency (16% to 41%) of *DUX4*-r (76, 141) The discrepancy may be attributed to differences in bioinformatic approaches or sample selection bias.

4.1.3 Analysis of WGS L-only 30× coverage

We tested the diagnostic sensitivity of low coverage WGS by down-sampled computationally ten samples to a mock depth of 30x and compared the findings to the original high coverage 90x. To further test whether L-only analysis on 30x coverage sequencing data had sufficient sensitivity to detect the relevant lesions, we sequenced 20 additional ALL samples to 30x depth. We also included two cases with low leukemic burden, 37% and 14% blasts as determined by FISH analysis. The results of *in silico* down-sampling were fully consistent with

previous analyses. Moreover, the results were also fully consistent with the SoC results and detected a clinically relevant lesion in 19/20 additional samples, including HeH, HoL, and several gene fusions. Even in samples with a low percentage of blasts, 30x coverage proved sufficient to detect both *ETV6::RUNX1* and *KMT2A::AFF1* lesions. Interestingly, in one sample with an *ERG* deletion detected by WGS, *IGH::DUX4*-r was also identified by the targeted approach described above. These findings suggest that 30x coverage may be adequate for diagnostic purposes in ALL. Notably, a study by Duncavage and colleagues, which achieved a mean coverage of 50x, reported 100% sensitivity for detecting CNVs and SVs, though sensitivity for SNVs was lower at 84.6% (142).

In the current ALL trial, high-risk genetic information and targetable aberrations must be available by day 14, which eliminates the possibility of using the remission sample as a normal/germline sample or culturing fibroblasts for L/N analysis. Since our results with L-only analysis were equally effective in identifying all mandatory aberrations, we propose that this approach could help meet the required turnaround time. Another critical consideration for the clinical implementation of WGS is the cost-benefit aspect. The use of 30x L-only samples improves cost-effectiveness, and the decreasing cost of sequencing in recent years makes WGS a viable alternative to multimodal SoC testing.

4.2 Paper II

The case report represents a real-world example of precision medicine illustrating how WGS may be used in ALL to identify potential therapeutic targets. The case describes a 7 year old child diagnosed with T-ALL, that showed no response to initial treatment. WGS was performed on DNA from the diagnostic BM, with the specific aim to identify clinically actionable therapeutic targets as described in the individualized therapy for relapsed malignancies in childhood (INFORM) (143).

The WGS data once filtered with a gene list derived from INFORM revealed 260 SNVs and 18 SVs. Further filtering using gnomAD or local observations together with annotations in ClinVar narrowed the number of SNVs to 11, which were all discarded after manual inspection as no targetable lesion was found. Among the SVs, two lesions, an inversion and a translocation shared a common breakpoint on 9p24. This shared breakpoint mapped to intron 15 (NM_001322194.2) of the *JAK2* locus making the lesion highly interesting. The reciprocal reads mapped, to 14q32.11, intron 25 (NM_001080414) of the *CDCC88C* locus. We interpreted the

rearrangement as a translocation with a concomitant inversion and as a result a fusion between *CDCC88C* and *JAK2* was formed. The WGS data supported a rearrangement resulting in a gene fusion that linked intron 25 of *CDCC88C* to intron 15 of JAK2. However, the genes were oriented in the opposite direction and it was the concomitant inversion of the 9p fragment that led to the fusion, predicted to be in-frame (Figure 4). Thus, according to the WGS, the fusion gene on derivative chromosome 9 consists of exons 1 to 15 of *CDCC88C* and exons 16 to 25 of *JAK2*.

The rearrangement of *JAK2* was verified by metaphase FISH, a break apart probe, which confirmed that the rearrangement involved the *JAK2* locus. Because this translocation involved small, distal regions of chromosomes 9 and 14, CBA could not detect the rearrangement. The fusion transcript was confirmed by RT-PCR and the exons involved by Sanger sequencing. Thus, this chimeric gene leads to an in-frame hybrid protein with an intact kinase domain of *JAK2*.

JAK2 protein, a tyrosine kinase, is involved in cell growth and differentiation via the JAK/STAT signaling pathway, and alterations in *JAK2* gene have been found in cancers including myeloid and lymphoid malignancies (144). In B-ALL, *JAK2* fusions are frequently observed in Ph-like ALL, exhibiting gene expression patterns similar to those seen in Ph+ALL (145). Furthermore, previous *JAK2* fusions identified in Ph-like ALL preserve exons 19-25, which encode the kinase domain, making these lesions amenable to targeting with JAK2 inhibitors (146). In T-ALL, *JAK2* fusions have been reported in isolated cases (146), although point mutations in members of the JAK/STAT pathway are more frequent, occurring in 25% of T-ALL cases (21, 146). While Jak2 inhibitors have demonstrated antileukemic activity in *ex vivo* experiments, *in vivo* experience is limited (147).

Although over 30 partner genes have been identified for *JAK2*, we have discovered *CDCC88C* as a novel partner gene. While the exact function of the *CDCC88C* protein remains unclear, there is evidence of a potential role in leukemia. *CDCC88C* is known to be involved in Wnt signaling and is highly expressed in pediatric B-cell and T-cell ALL. Previous studies have reported that *CDCC88C* may be implicated in fusions with *FLT3* and *PDGFRB* in hematological malignancies (Gosenca, Kellert et al. 2014; Ear, Dunkel et al. 2019; Kurihara, Mizuno et al. 2022).

The patient responded to second line therapy and later received allogeneic stem cell transplantation (SCT) and was thus not treated with Jak2 inhibitors.

Nonetheless the case report illustrates the utility of WGS to identify potential therapeutic targets refractory or relapsed pediatric ALL.

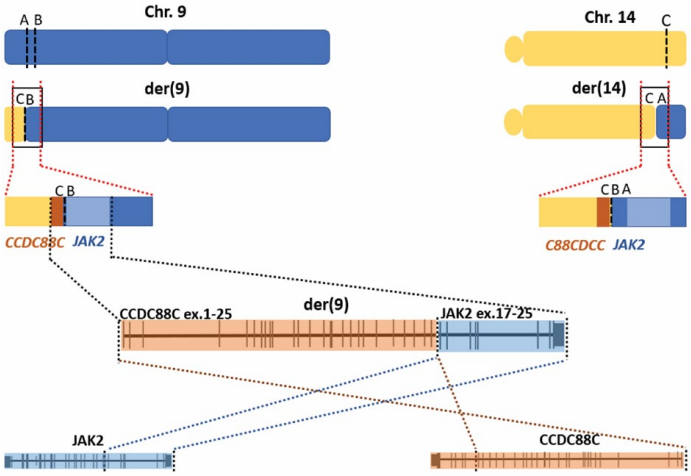

Figure 3. Cartoon showing a schematic representation of chromosomes 9 (blue) and 14 (yellow) at the derivative chromosomes 9 and 14 (upper panel), the black squares mark the junctions. Middle cartoon represents a blow up of the junctions illustrating the creation of the fusion gene on derivative chromosome 9 (der9). The lower panel represents a magnification of the fusion gene showing exons 1–25 in *CCDC88C* joined to exons 16–25 in *JAK2* (148).

4.3 Paper III

This paper serves as a proof of principle, demonstrating how WGS data can be leveraged to identify molecular markers for designing patient-specific quantitative assays to measure and monitor MRD in pediatric ALL patients. We quantified the number of targets using ddPCR in BM gDNA, as well as in cell-free tumor DNA (ctDNA) from blood and cerebrospinal fluid (CSF).

4.3.1 Target Identification by WGS

In order to develop patient-specific ddPCR assays from WGS data to detect the major leukemic clone during patient follow-up, three criteria were established: i) a unique sequence at a junction created by the genomic rearrangements in the leukemic blasts ii) event supported by reads that indicate their presence in the majority of leukemic cells, iii) the event should be eihter recurrent in ALL or a putative driver event. Our goal was to develop two to three assays per patient. Once the junctions were identified, the probe generating the fluorescent signal

was placed over the junction to achieve a high specificity of the assay. We successfully identified suitable markers and developed ddPCR assays for all patients selected in this study.

The WGS data detected a reciprocal translocation in patient 1, t(12;19)(p13;p13) resulting in a *TCF3::ZNF384* fusion gene and a deletion in the KRAS gene. Since *ZNF384* rearrangements in B-ALL have been reported as a distinct entity, the *TCF3::ZNF384* fusion was selected for ddPCR design. Furthermore, *KRAS* plays a crucial role in the RAS/MAPK signaling pathway, and aberrations in this gene are known in many cancers, making it a good candidate for MRD detection. For patient 2 we had to discard most SV as the junctions were located on stretches of repetitive sequence and were not suited to design assays. However, small intragenic deletions in *IKZF1* and *PAX5*, were identified and their junctions selected as ddPCR targets. Patient 3 had been diagnosed with HeH ALL with 54 chromosomes. Upon inspection of the WGS data, it was evident that most of the structural genomic rearrangements had occurred in highly repetitive regions, making them unsuitable as targets. Only one SV junction, resulting from a rearrangement between chromosomes 2 and 8, was deemed suitable for design. To maximize the signal and increase the probability of detecting this marker, two non-overlapping probes were designed to target the junction, one on the plus strand and one on the minus strand.

In patient 4, four unrelated SVs located in chromosome arms 4p, 9p, 20p and 21q were reported by the callers which were also present at relapse. After inspection of all junctions, three targets were selected, a deletion in *CDKN2A/B* genes, a deletion resulted from a pericentric inversion of chromosome 4, and a junction left by a deletion in chromosome 20. In patient 5, cytogenetic analysis had identified a t(5;7)(q35;q21) translocation and WGS evidenced a *TLX3::CDK6* fusion, previously described in T-ALL (149). However, the junctions were located in a highly repetitive region, making it unsuitable for design purposes. Finaly, the junctions resulting from deletions on 9p, 8p and 16q were selected for ddPCR assays. In patient 6, while CBA and array CGH detected several SVs on chromosome 6, as well as a recurrent deletion on 9p21.3. WGS confirmed these events and additionally detected a *STIL::TAL1* fusion, recurrent in T-ALL (150). Both *STIL::TAL1* and a deletion on 9p were selected as targets.

During the selection process, we ensured that all chosen targets were present in the majority of cells in the diagnostic sample. This was achieved by reviewing

clinical data from FCM, CBA, FISH, and array CGH, in conjunction with WGS findings, as this was a critical criterion to ensure that our assays accurately measured the entire leukemic burden. Although WGS detected structural SVs in all patients, not all of them were suitable for use as targets, which represents a limitation of this approach. The presence of SVs in repetitive regions is a known feature of the human genome (151), and this poses a significant challenge when attempting to use the junctions created by SVs as targets for MRD monitoring. Additionally, in patients with aneuploidies as class-defining lesions, identifying suitable targets can be particularly difficult.

4.3.2 Optimization and Performance of Droplet Digital PCR Assays

The optimization of the assays started with the determination of the appropriate annealing temperature in single assays and multiplexed with the reference gene assay. DNA from normal samples was also included in the optimization as negative controls (NC) to ensure the specificity of the assays. The annealing temperature was optimized to clearly distinguish the reference from the target assay, ensuring no false-positive signals were detected in the negative controls NC. Of 18 ddPCR assays, 15 were successful in detecting the targets and showed no background signal. This underscores the advantage of selecting multiple targets per patient to ensure the detection of leukemic cells, to minimize the risk of losing targets during clonal evolution, and to guarantee that at least one assay will be effective.

Successful assays were evaluated for linearity and sensitivity using serial 10-fold dilutions to determine the range of quantifiability (LOQ) and detectability (LOD)(Figure 4A). The assays for four patients reached a LOQ down to 10^{-5} and the remaining two reached 10^{-4}. Importantly, for the two T-ALL patients in the study, our assays showed LoQ/LoD at least one order of magnitude beyond the RQ-PCR markers. The LoQ/LoD can be pushed even further by increasing the input DNA running more replicates for each sample. The strategy to include two probes in a single assay, was successful to achieve a sensitivity comparable to the other assays. The results showed that the patient-specific SV assays were highly specific and sensitive (Figure 4B) and that the limiting factor was the input DNA.

Since we expected the amount of ctDNA to be low in follow up and CSV samples, multiplexing the assays for individual patients was also tested. We further investigated the lowest amount of DNA input that achieved a reliable signal by testing 1 ng of gDNA input per well (from 10^{-1} and 10^{-2} dilutions of PC in NC). All targets were detectable in 100 pg of patient DNA in 900 pg of normal control DNA while only traces of targets (<3 target-positive droplets in three replicate wells) were detected in when the ratio was 10/990 pg. Notably, the measured number of targets was consistently lower than the theoretical number calculated assuming no pipetting losses and a PCR efficiency of 100%. We concluded that the quantification of the targets, even in low DNA input is possible with the patient-specific assays.

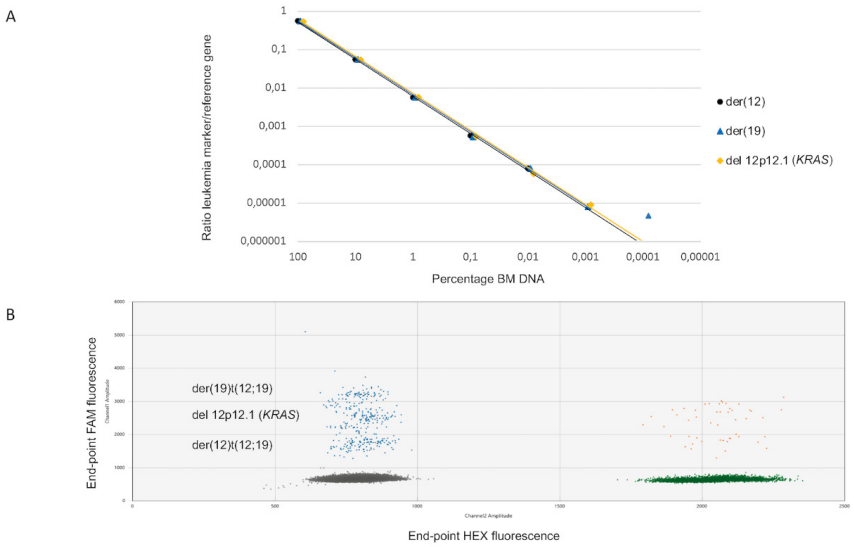

Figure 4. (A, B) Dilution series and ddPCR fluorescence plot for leukemic targets in patient 1. (A) Sensitivity testing for the three targets selected for patient 1, der (12), der(19) and deletion 12p. The ratios of CN target to CN reference obtained from the ddPCR assays are plotted against the percentage of gDNA from the diagnostic BM for the different dilutions using 500 ng of input gDNA. (B) 2D fluorescence amplitude plot generated by the QuantaSoft™ software showing the droplet clusters corresponding to the three different targets for patient 1 (der(12), der(19), and deletion 12p using 1 ng of PC in a background of 9 ng of NC. y-Axis, end-point FAM-fluorescence from targets. x-Axis, end-point HEX-fluorescence from reference gene. Negative droplets (gray). Target positive–reference negative droplets

(green). Double-positive droplets (orange). ddPCR, droplet digital PCR; CN, copy number; gDNA, genomic DNA; BM, bone marrow; PC, positive control; NC, negative control (152); CN, copy number; der 12, derivative chromosome 12; der 19, derivative chromosome 19.

The molecular monitoring of residual disease in ALL quantifies the clonal IG/TR rearrangement with qPCR and the methodology is standardized by the Euro-MRD consortium (110). Our assays complied with the Euro-MRD criteria for detection of MRD, the performance of all our assays was comparable to, or even exceeded, that of IG/TR-PCR in terms of specificity, sensitivity and range of quantifiability while not showing any background amplification from the germline normal DNA. The only factor limiting the sensitivity and range of quantifiability of the targets is the amount of input DNA. Although, the amount of input template in the ddPCR reaction was limited by availability and technical constraints in our study, this could easily be overcome in the clinical setting by using more replicates. Interestingly, even though patient 4 assays had the lowest LoD/LoQ among all six patients, ddPCR detected higher MRD levels than the routine methods at both end of induction (EoI) and end of first block consolidation therapy (CB1).

It was still early days regarding the use of WGS in clinical routine and challenging to generate and analyze the data meeting the turnaround time required in the diagnostic setting. However, we were able to show that the approach has big potential to yield highly specific and sensitive assays for the majority of ALL patients.

4.3.3 Detection of targets in cfDNA

Next, we estimated the total amount of cfDNA recovered from plasma by analyzing the copy numbers of reference genes measured by ddPCR, as total cfDNA in itself has been proposed as a marker for MRD (153). A few samples resulted in no signal in the ddPCR reaction. Careful inspection revealed that the samples had very high cell counts with extremely high concentration of cfDNA, which caused oversaturated ddPCR reactions and no signals generated. To address this, we diluted the sample to approximately 10^{-5} which solved the issue.

The data showed the amount of total cfDNA was directly proportional to the white blood cell count in the sample. Overall, the amount of cfDNA recovered decreased during induction, and in two patients, it dropped by approximately 0.5 to 3 log10 at the end of the first block of consolidation therapy. In a study by

Schwartz and coworkers, they suggest that total cfDNA could potentially be used as an MRD marker (153). Discrepancies between plasma cfDNA and BM-MRD were observed in three patients due to severe treatment-related pancreatitis, post-sepsis, renal insufficiency and coagulation. Thus, according to our observations, total cfDNA reflects pathophysiological processes in the body and is not specific for leukemic burden.

In general, the kinetics of ctDNA targets followed the same pattern as the results from BM. Interestingly, in patient 4, who suffered from relapse, ddPCR detected the targets in cfDNA two months before the first relapse was diagnosed. These results open the possibility of using cfDNA for monitoring patient specific targets.

We also investigated ctDNA isolated from the CSF and observed that the results did not follow the same pattern as BM-MRD nor plasma ctDNA. In one of our patients, for example, despite a high amount of leukemic target in the plasma ctDNA, only a few copies of the reference and no leukemic target at all were detected. In another patient, who was positive for leukemic targets in the CSF, only trace amounts were found in later samples with no correlation to the level in the paired plasma samples. These results suggest that leukemic targets in CSF do not originate from leaked cells from plasma but rather may represent the existence of the malignant cells in the CSF.

4.4 Paper IV (Manuscript)

The study was initiated following the observation that two pediatric Ph+ALL patients exhibited persistent *BCR::ABL1* transcripts at the end of treatment (EoT), despite being in complete remission according to FCM and IG/TR qPCR. Our primary objective was to investigate whether we could identify differentiating features between the various types of *BCR::ABL1*-driven acute leukemias. To achieve this, we studied a retrospective cohort of well-characterized Ph+ALL samples (n = 15) and conducted a detailed analysis of their genomic and transcriptomic features. Additionally, we integrated these data with findings from SoC diagnostics and MRD information, aiming to uncover distinguishing features among Ph+ALL subgroups.

As a first step, we examined the available data on MRD quantification by the different methodologies. We found that in four patients the MRD measured by FCM and the qRT-PCR measuring the *BCR::ABL1* transcript differed at least at one timepoint. Two of these patients (P02 and P04) were the pediatric patients who triggered our study and for those we subsequently investigated all available follow-up samples.

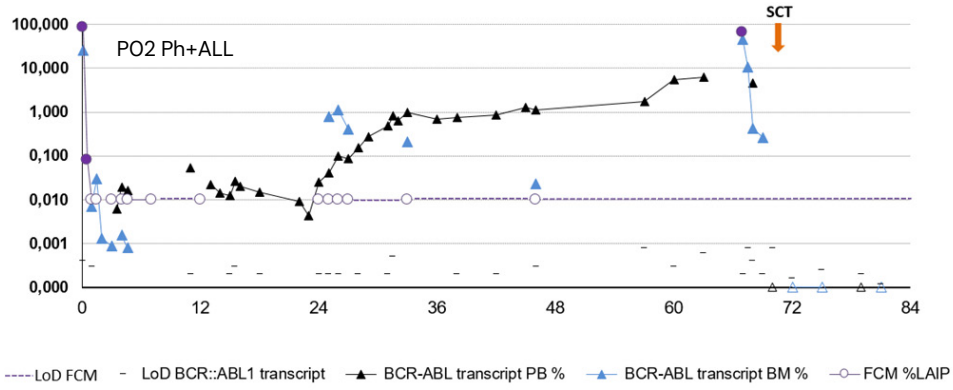

Figure 5. Comparison of MRD measurements over time by FCM and qPCR, belong to P02 case. The figure illustrates the percentage of leukemic blasts as quantified by FCM, (purple dots), and by qRT-PCR *BCR::ABL1* transcript in BM (blue symbols) and blood (black symbols). Empty symbols refer to not detected. Orange arrows indicate the timepoint for stem cell transplantation (SCT).

For Ph02 patient, the LAIP was undetectable by FCM from the end of induction until relapse. However, the *BCR::ABL1* transcript level was detected in both blood and BM (Figure 5). The level of transcript increased from EoT onwards and remained high despite TKI treatment. Five years after initial diagnosis and one year after discontinuation of Dasatinib, the *BCR::ABL1* transcript level in blood increased again and 6 months later the patient presented with overt relapse. Interestingly, at this stage, the blasts exhibited a subtle immunophenotypic shift, along with the presence of aberrant myeloid markers by FCM. Additionally, morphologic examination of the BM revealed a left shift in granulopoiesis, accompanied by an elevated granulocyte count in the peripheral blood - all indicative of CML as the underlying disease for this patient (Figure 6).

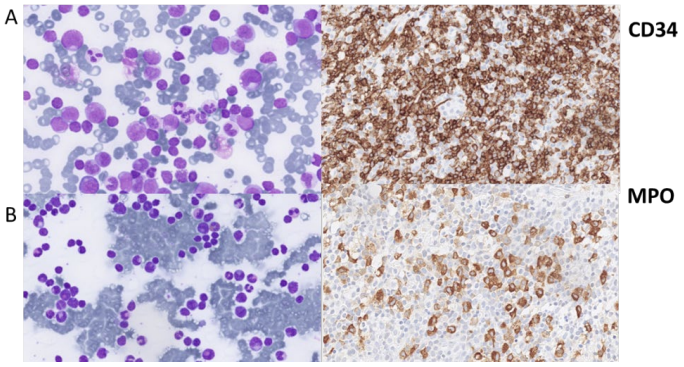

Figure 6. BM (A) and blood smears from P02 relapse stained with Wright-Giemsa B) Immunohistochemistry showing CD34 and MPO staining.

Ph04 showed a similar discrepancy, while FCM failed to detect leukemic blasts, *BCR::ABL1* expression was consistently detectable from d29 to EoT both on PB and BM. Notably, both patients expressed the *BCR::ABL1* transcript in sorted CD34+, CD33+, CD29+, and CD3+ cell fractions.

Next, we sought to investigate whether the follow up samples harbored the additional copy of the Ph chromosome detected by CBA and FISH at presentation. For the purpose we developed ddPCR assays targeting the genomic *BCR::ABL1* on the der 22 and the reciprocal *ABL1::BCR* on der 9 and, quantified both derivative chromosomes at presentation and at follow-up. The results showed that in both patients at diagnosis, the number of der 22 was twice that of der 9, a ratio of 2:1, but in all follow-up samples including the relapse sample, the ratio was 1:1 (Figure 7A). Our interpretation is that the clone with additional der 22, which was responsible for driving the leukemia to acute phase, was eradicated by the therapy, and that the original underlying pathological cells were not affected by treatment. These cells accounted for *BCR::ABL1* transcript expression and harbored only one copy of the Ph chromosome. This raised another question, namely whether these leukemic cells were of lymphoid or myeloid origin. To answer this question, both the *BCR::ABL1* transcript and the genomic *BCR::ABL1* junctions were measured in sorted cells from one follow-up sample for each patient. The results showed that the cell fractions CD3 (B cells), CD4 (T cells), CD19 (granulocytes), and CD34 (progenitor cells) all expressed the *BCR::ABL1* transcript, as well as the der 22 and der 9 junction sequences in varying proportions, with a 1:1 ratio (Figure 7B). Taken together, these results

suggest that at presentation, the disease manifested as acute leukemia, while the genetically simpler clone without lineage commitment persisted after treatment and later caused relapse in P02 and expressed the *BCR::ABL1* transcript in P04. These data support the hypothesis that CML transformed into blast crisis as the underlying disease, which could be confirmed for P02 at relapse.

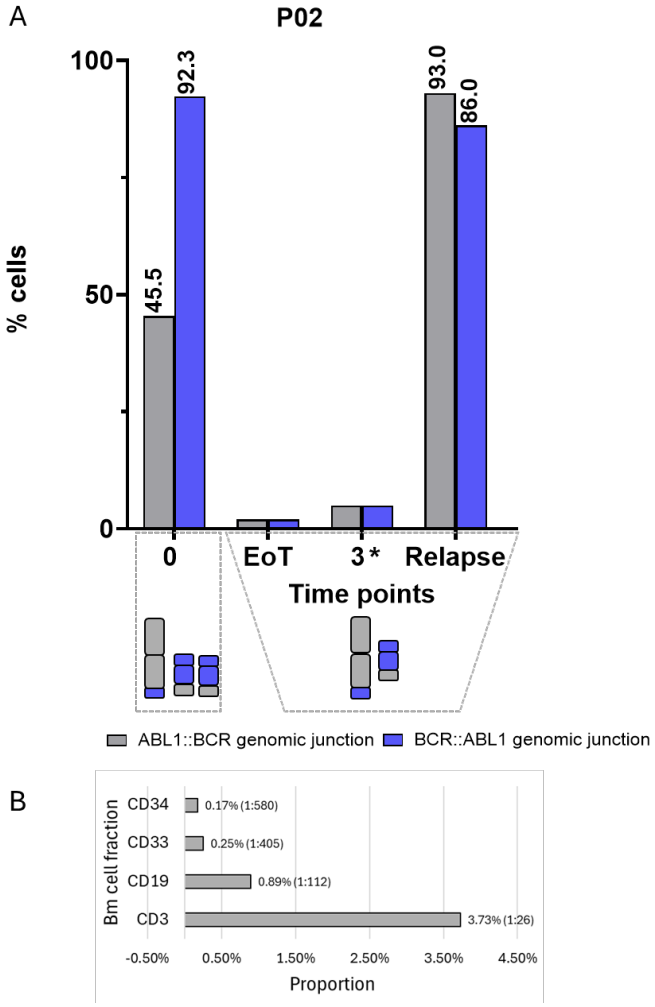

Figure 7. Quantification of the genomic junctions on der 9 and der 22 by ddPCR. A) The staple diagram shows the percentage of *ABL1::BCR* vs *BCR::ABL1* junctions in P02 patient. B) proportion of CD3+, CD19+, CD33+ and CD34+ cells that harbor *BCR::ABL1* junction in the time point marked with asterisk.

Subsequently, we performed a supervised analysis from all Ph+ALL samples using the ALLCatchR-bcrabl1 bioinformatics tool developed by Beder et al (132, 154). This far we have processed data on 13 samples from 11 Ph+ALL patients.

The tool assigned Ph02 sample taken at presentation as lymphoid subtype and the relapse sample as multi lineage. This discrepancy is likely explained by the high number of lymphoid blasts at presentation which made the underlying disease impossible to detect neither by GEP, nor immunophenotypically. However, FISH analysis had shown that one small clone (4%) with only one Philadelphia chromosome coexisted at presentation with the dominant clone harboring two Ph chromosomes. Of the 11 samples taken at presentation, the tool assigned three samples (P08, P11, and P16) to the multilineage subgroup (*BCR::ABL1*-M), while the remaining samples were classified as lymphoid (*BCR::ABL1*-L). Neither of the presentation samples from P02 nor P04 were predicted to belong to the *BCR::ABL1*-M subgroup. In fact, the classifier did not classify P04 as Ph+ALL, instead assigned a high score (0.8) for the Ph-like-ALL category. This unexpected result may be due to the RNA for P04 being extracted from peripheral blood, whereas the other samples were derived from BM. Interestingly, the relapse sample from P02 was assigned to the *BCR::ABL1*-M subgroup by the classifier. All the samples classified as *BCR::ABL1*-M (n = 4) were found to express aberrant myeloid markers, though not all samples with aberrant myeloid markers were classified as *BCR::ABL1*-M. Notably, only P16 showed discordant MRD levels between flow cytometry (FCM) and qPCR results. Furthermore, the classifier grouped three *BCR::ABL1*-M samples into the subgroup defined by the deletion of *HBS1L*, while P11 was assigned to the *BCR::ABL1*-M subgroup characterized by monosomy 7, which had also been detected by cytogenetic analysis (CBA) as a subclonal lesion.

5 Conclusions

Taken together, the work presented in this thesis highlights the potential of WGS as a valuable tool in the diagnostic workup of ALL. By utilizing WGS as the sole method for detecting aneuploidies, iAMP21, SVs and CNA profiles, we successfully identified all genetic aberrations specified in the treatment protocol. The WGS results were in complete concordance with SoC findings, enabling accurate allocation to the correct genetic subgroups in all cases. Additionally, we were able to classify most of the B-other cases into emerging subgroups, including *DUX4*-r. Our L-only approach demonstrated equal accuracy compared to paired analysis. We also showed how the WGS data can be used to identify personalized treatment options. In a T-ALL patient, who failed first line therapy, WGS identified a novel *JAK2* fusion. Further investigations revealed that the fusion gene product is a chimeric protein with tyrosine kinase activity, for which targeted experimental therapy is available.

In addition, the sequence data provided by WGS enabled the design of patient-specific ddPCR quantitative assays. These assays demonstrated superior specificity and sensitivity compared to conventional quantitative methods. Furthermore, results from ctDNA indicate that monitoring leukemia markers in plasma can provide valuable information for the individualized follow-up of ALL patients. We conclude that designing patient-specific targets identified by WGS has significant potential for MRD monitoring.

Furthermore, the application of WGS and WTS enabled a detailed investigation of the genomic and transcriptomic characteristics of Ph+ALL. A comprehensive analysis of all available data at the diagnostic, follow-up, genomic and transcriptomic levels was performed to identify features that could differentiate Ph+ALL-L from Ph+ALL-M. Ongoing analysis suggests that the insights gained from WGS and WTS could further elucidate the complexities of Ph+ALL and contribute to a better understanding of its subtypes.

In summary, the work presented underscores the transforming potential of WGS in the diagnostic and follow-up management of ALL. The ability of WGS to accurately detect genetic aberrations, classify subgroups, and guide personalized treatment decisions illustrates its value in clinical practice. Furthermore, the application of WGS and WTS provides a deeper understanding of Ph+ALL, offering insights that could inform future diagnostic strategies and therapeutic approaches. Ultimately, the findings highlight how integrating WGS

into routine clinical workflows could not only enhance diagnostic precision but also pave the way for more tailored and effective treatment options for ALL patients.

6 Points of perspective

The field of genetic diagnostics has made significant strides with the introduction of high-throughput technologies. To fully harness the potential of these advanced methods in clinical diagnostics, it is essential to standardize result generation and interpretation. This standardization ensures that clinicians can diagnose patients using consistent criteria and provide appropriate treatments. In Sweden, the Genomic Medicine Sweden (GMS) initiative is leading the effort to coordinate and harmonize high-throughput technologies to support precision medicine in diagnostics. Since 2020, GMS has specifically aimed to implement WGS and Whole Transcriptome Sequencing (WTS) as routine clinical diagnostics for all pediatric cancer cases (155). Additionally, GMS collaborates with other genomic medicine initiatives across Europe, exchanging knowledge and learning from each other's experiences.

In our studies, WGS results were found to be comparable to conventional methods, yet they provided superior resolution. The adoption of WGS has the potential to transform clinical diagnostics, shifting from a multi-modal approach to a unified method. However, for WGS to become the primary methodology in clinical genetic diagnostics, it requires skilled professionals in both genetics and bioinformatics. In addition to assessing the WGS method itself, it is crucial to consider and validate the entire workflow, including sample handling, professional training, data generation, security, storage, pre-analysis, downstream analysis, result interpretation, and reporting to clinicians. Moreover, the development of a user-friendly interface will help professionals quickly adapt to the method, reducing the risk of misinterpretation. In the near future, the widespread use of WGS across all diagnostics could lead to a more data-driven clinical workflow, which will likely require advanced big data handling and machine learning bioinformatics.

Although the cost of WGS remains relatively high, micro-costing calculations comparing the SoC and WGS indicate that, if samples can be batched, the additional cost of WGS is reasonable. Importantly, the cost of WGS has decreased over time due to its increased usage. However, unlike traditional methods, implementing WGS necessitates the creation of new infrastructure within healthcare systems, which should be factored into the overall cost. Ultimately, the benefits of WGS in patient management should be the main focus when considering its value.

WGS enables precise identification of genetic abnormalities in individual patients, allowing clinicians to create tailored treatment plans and targeted therapies core components of precision medicine. However, translating this complex genetic data into actionable clinical targets remains a challenge. It is also important to avoid overemphasizing genetic factors, which could lead to neglecting environmental and lifestyle influences on health.

In summary, human genetics is undergoing a revolutionary transformation, with data-driven oncology diagnostics becoming increasingly inevitable. While our study focused on the feasibility of using WGS in the clinical setting for ALL diagnostics, similar studies are being conducted in other areas of human genetics. By sharing insights on WGS data analysis and interpretation, we are progressing toward the ultimate goal: precision medicine that saves lives and improves the quality of life for survivors.

7 Acknowledgements

I would like to thank my main supervisor Gisela Barbany. Thank you for giving me the opportunity to be your doctoral student. Your enthusiasm in finding the answer and your commitment to your patients are inspiring. You have not only been a supervisor to me, but also my mentor. Thank you for the many interesting conversations and laughter.

To my co-supervisor Richard Rosenquist, thank you for welcoming me to the clinical genetics research group and giving me valuable feedback whenever I asked for it.

To my co-supervisor Jesper Eisfeldt, thank you for helping me with the bioinformatic analysis in all my projects. You have always been very flexible and open to all our ideas and have always found a way to make it work. I would also like to thank you for the many interesting conversations.

I would like to thank my co-supervisor Anh Nhi Tran for her willingness to be my co-supervisor and for her valuable feedback.

Ingegerd and Cecilia, thank you for your patience during the time we worked on our projects. I really enjoyed the time during our collaboration. I look forward to continuing to work with you.

I would like to thank you Magnus Nordenskjöld for your kind encouragement and sharing fun and informative stories.

I would also like to thank all my colleagues in clinical genetics and genomics. Anna Lindstrand, thank you for welcoming me to the geneticist group. Thank you Lucia for your kind encouragement and for sharing your work experiences with me. Thank you Isabel for your sincere friendship and motherly advice. Thank you, Nilofar, for sharing your life experiences with me and making me feel like I always have someone to talk to. Thank you Malin, my wonderful friend, for your kind words of encouragement. Thank you Fabjola, Alexandra, Mikaela, Inger, Eva, Monika, Helen and Anita for making me feel welcome in the lab. Thank you Karin Kindberg for being so helpful and kind to me. Anna Hammarsjö, thank you for being an inspiration.

Thank you to my colleagues and friends in clinical genetics, Larry, Cecilia Ö, Cecilia J, Karin S, Gilbert, Magdalena, Marlene, Fulya, Fan, Sakshi, Litika, Blaz, Birna, Daniel, Jose, Sailan, Zahra, Jakob, Sofia, Emma Ehn, Emma Tham, Tatjana and

Mikaela for the many nice conversations over lunch. Thank you Teresa and Nina for your support and kind company. Thank you, Leily and Mastoureh, for being my Farsi speaking friends and sharing your life experiences with me.

To my parents, thank you for your support and encouragement. To my brothers and sister, thank you for your unconditional support and love. To my husband, you are my “Swiss army knife”, there is nothing you can't do to solve my problems. Thank you for having patience with me during all the deadlines and the most stressful moments during my PhD, this is an achievement for both of us. Till mina underbara söner Edvin och Oliver, ni påminner mig om vad som är viktig i livet, tack för att ni kom i vårt liv.

8 Declaration about the use of generative AI

The following AI assisted tools, ChatGPT, Office365 Copilot, Gemini were used in writing the “kappa”/comprehensive summary of the thesis , for language refining and grammar editing to enhance clarity.

I take full responsibility for the content of the “kappa”/comprehensive summary of the thesis

9 References

1. Wirchow R. Zur pathologischen Physiologie des Blutes. II. Weisses Blut. Archives of Pathology, Anatomy and Physiology. 1847:I: 563-72.

2. NOWELL PCH, D A. Chromosome studies in human leukemia. II. Chronic granulocytic leukemia. Journal of the National Cancer Institute. 1961;27:1013-35.

3. Rowley JD. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. Nature. 1973;243(5405):290-3.

4. Shtivelman E, Lifshitz B, Gale RP, Canaani E. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. Nature. 1985;315(6020):550-4.

5. Druker BJ, Guilhot F, O'Brien SG, Gathmann I, Kantarjian H, Gattermann N, et al. Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. N Engl J Med. 2006;355(23):2408-17.

6. Irons RD, Stillman WS. The process of leukemogenesis. Environ Health Perspect. 1996;104 Suppl 6(Suppl 6):1239-46.

7. van Dongen JJ, Lhermitte L, Böttcher S, Almeida J, van der Velden VH, Flores-Montero J, et al. EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. Leukemia. 2012;26(9):1908-75.

8. Hjalgrim LL, Rostgaard K, Schmiegelow K, Söderhäll S, Kolmannskog S, Vettenranta K, et al. Age- and sex-specific incidence of childhood leukemia by immunophenotype in the Nordic countries. J Natl Cancer Inst. 2003;95(20):1539-44.

9. Johansson B. CYTOGENETISKA OCH MOLEKYLÄRGENETISKA

FÖRÄNDRINGAR I MALIGNA HEMATOLOGISKA SJUKDOMAR. In: Avdelningen för klinisk genetik Lu, editor. 2014.

10. Toft N, Birgens H, Abrahamsson J, Griškevičius L, Hallböök H, Heyman M, et al. Results of NOPHO ALL2008 treatment for patients aged 1-45 years with acute lymphoblastic leukemia. Leukemia. 2018;32(3):606-15.

11. Norén-Nyström U, Andersen MK, Barbany G, Dirse V, Eilert-Olsen M, Engvall M, et al. Genetic Subtypes and Outcome of Patients Aged 1 to 45 Years Old With Acute Lymphoblastic Leukemia in the NOPHO ALL2008 Trial. Hemasphere. 2023;7(5):e883.

12. Greaves M. A causal mechanism for childhood acute lymphoblastic leukaemia. Nat Rev Cancer. 2018;18(8):471-84.

13. Tebbi CK. Etiology of Acute Leukemia: A Review. Cancers (Basel). 2021;13(9).

14. Tesi B, Robinson KL, Abel F, Díaz de Ståhl T, Orrsjö S, Poluha A, et al. Diagnostic yield and clinical impact of germline sequencing in children with CNS and extracranial solid tumors-a nationwide, prospective Swedish study. Lancet Reg Health Eur. 2024;39:100881.

15. Cazzaniga G, van Delft FW, Lo Nigro L, Ford AM, Score J, Iacobucci I, et al. Developmental origins and impact of BCR-ABL1 fusion and IKZF1 deletions in monozygotic twins with Ph+ acute lymphoblastic leukemia. Blood. 2011;118(20):5559-64.

16. Bang B, Eisfeldt J, Barbany G, Harila-Saari A, Heyman M, Zachariadis V, et al. A somatic UBA2 variant preceded ETV6-RUNX1 in the concordant BCP-ALL of monozygotic twins. Blood Adv. 2022;6(7):2275-89.

17. Schmiegelow K, Forestier E, Hellebostad M, Heyman M, Kristinsson J, Söderhäll S, et al. Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. Leukemia. 2010;24(2):345-54.

18. A Treatment Protocol for Participants 0-45 Years With Acute Lymphoblastic Leukaemia [Available from: https://clinicaltrials.gov/ct2/show/NCT03911128.

19. Zaliova M, Stuchly J, Winkowska L, Musilova A, Fiser K, Slamova M, et al. Genomic landscape of pediatric B-other acute lymphoblastic leukemia in a consecutive European cohort. Haematologica. 2019;104(7):1396-406.

20. Liu YF, Wang BY, Zhang WN, Huang JY, Li BS, Zhang M, et al. Genomic Profiling of Adult and Pediatric B-cell Acute Lymphoblastic Leukemia. EBioMedicine. 2016;8:173-83.

21. Liu Y, Easton J, Shao Y, Maciaszek J, Wang Z, Wilkinson MR, et al. The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. Nat Genet. 2017;49(8):1211-8.

22. Pui CH, Raimondi SC, Dodge RK, Rivera GK, Fuchs LA, Abromowitch M, et al. Prognostic importance of structural chromosomal abnormalities in children with hyperdiploid (greater than 50 chromosomes) acute lymphoblastic leukemia. Blood. 1989;73(7):1963-7.

23. Moorman AV, Richards SM, Martineau M, Cheung KL, Robinson HM, Jalali GR, et al. Outcome heterogeneity in childhood high-hyperdiploid acute lymphoblastic leukemia. Blood. 2003;102(8):2756-62.

24. Heerema NA, Nachman JB, Sather HN, Sensel MG, Lee MK, Hutchinson R, et al. Hypodiploidy with less than 45 chromosomes confers adverse risk in childhood acute lymphoblastic leukemia: a report from the children's cancer group. Blood. 1999;94(12):4036-45.

25. Johansson CJHaB. Acute lymphoblastic leukemia. In: Mitelman SHaF, editor. Third edition Cancer Cytogenetics. 3 ed. Ubited State of America: John Wiley and Sons; 2009. p. 233-96.

26. Onodera N, McCabe NR, Nachman JB, Johnson FL, Le Beau MM, Rowley JD, et al. Hyperdiploidy arising from near-haploidy in childhood acute lymphoblastic leukemia. Genes Chromosomes Cancer. 1992;4(4):331-6.

27. Paulsson K, Mörse H, Fioretos T, Behrendtz M, Strömbeck B, Johansson B. Evidence for a single-step mechanism in the origin of hyperdiploid childhood acute lymphoblastic leukemia. Genes Chromosomes Cancer. 2005;44(2):113-22.

28. Harrison CJ, Moorman AV, Broadfield ZJ, Cheung KL, Harris RL, Reza Jalali G, et al. Three distinct subgroups of hypodiploidy in acute lymphoblastic leukaemia. Br J Haematol. 2004;125(5):552-9.

29. Safavi S, Forestier E, Golovleva I, Barbany G, Nord KH, Moorman AV, et al. Loss of chromosomes is the primary event in near-haploid and low-hypodiploid acute lymphoblastic leukemia. Leukemia. 2013;27(1):248-50.

30. Gibbons B, MacCallum P, Watts E, Rohatiner AZ, Webb D, Katz FE, et al. Near haploid acute lymphoblastic leukemia: seven new cases and a review of the literature. Leukemia. 1991;5(9):738-43.

31. Simons A, Stevens-Kroef M, El Idrissi-Zaynoun N, van Gessel S, Weghuis DO, van den Berg E, et al. Microarray-based genomic profiling as a diagnostic tool in acute lymphoblastic leukemia. Genes Chromosomes Cancer. 2011;50(12):969-81.

32. Robinson HM, Harrison CJ, Moorman AV, Chudoba I, Strefford JC. Intrachromosomal amplification of chromosome 21 (iAMP21) may arise from a breakage-fusion-bridge cycle. Genes Chromosomes Cancer. 2007;46(4):318-26.

33. Romana SP, Poirel H, Leconiat M, Flexor MA, Mauchauffé M, Jonveaux P, et al. High frequency of t(12;21) in childhood B-lineage acute lymphoblastic leukemia. Blood. 1995;86(11):4263-9.

34. Shurtleff SA, Buijs A, Behm FG, Rubnitz JE, Raimondi SC, Hancock ML, et al. TEL/AML1 fusion resulting from a cryptic t(12;21) is the most common genetic lesion in pediatric ALL and defines a subgroup of patients with an excellent prognosis. Leukemia. 1995;9(12):1985-9.

35. McLean TW, Ringold S, Neuberg D, Stegmaier K, Tantravahi R, Ritz J, et al. TEL/AML-1 dimerizes and is associated with a favorable outcome in childhood acute lymphoblastic leukemia. Blood. 1996;88(11):4252-8.

36. Harrison CJ, Moorman AV, Barber KE, Broadfield ZJ, Cheung KL, Harris RL, et al. Interphase molecular cytogenetic screening for chromosomal abnormalities of prognostic significance in childhood acute lymphoblastic leukaemia: a UK Cancer Cytogenetics Group Study. Br J Haematol. 2005;129(4):520-30.

37. Golub TR, Barker GF, Bohlander SK, Hiebert SW, Ward DC, Bray-Ward P, et al. Fusion of the TEL gene on 12p13 to the AML1 gene on 21q22 in acute lymphoblastic leukemia. Proc Natl Acad Sci U S A. 1995;92(11):4917-21.

38. Romana SP, Mauchauffé M, Le Coniat M, Chumakov I, Le Paslier D, Berger R, et al. The t(12;21) of acute lymphoblastic leukemia results in a tel-AML1 gene fusion. Blood. 1995;85(12):3662-70.

39. Ford AM, Bennett CA, Price CM, Bruin MC, Van Wering ER, Greaves M. Fetal origins of the TEL-AML1 fusion gene in identical twins with leukemia. Proc Natl Acad Sci U S A. 1998;95(8):4584-8.

40. Greaves MF, Maia AT, Wiemels JL, Ford AM. Leukemia in twins: lessons in natural history. Blood. 2003;102(7):2321-33.

41. Attarbaschi A, Mann G, König M, Steiner M, Dworzak MN, Gadner H, et al. Near-tetraploidy in childhood B-cell precursor acute lymphoblastic leukemia is a highly specific feature of ETV6/RUNX1-positive leukemic cases. Genes Chromosomes Cancer. 2006;45(6):608-11.

42. Raimondi SC, Zhou Y, Shurtleff SA, Rubnitz JE, Pui CH, Behm FG. Near-triploidy and near-tetraploidy in childhood acute lymphoblastic leukemia: association with B-lineage blast cells carrying the ETV6-RUNX1 fusion, T-lineage immunophenotype, and favorable outcome. Cancer Genet Cytogenet. 2006;169(1):50-7.

43. Carroll AJ, Crist WM, Parmley RT, Roper M, Cooper MD, Finley WH. Pre-B cell leukemia associated with chromosome translocation 1;19. Blood. 1984;63(3):721-4.

44. Michael PM, Levin MD, Garson OM. Translocation 1;19--a new cytogenetic abnormality in acute lymphocytic leukemia. Cancer Genet Cytogenet. 1984;12(4):333-41.

45. Williams DL, Look AT, Melvin SL, Roberson PK, Dahl G, Flake T, et al. New chromosomal translocations correlate with specific immunophenotypes of childhood acute lymphoblastic leukemia. Cell. 1984;36(1):101-9.

46. Hunger SP. Chromosomal translocations involving the E2A gene in acute lymphoblastic leukemia: clinical features and molecular pathogenesis. Blood. 1996;87(4):1211-24.

47. Bergfelt Lennmyr E, Engvall M, Barbany G, Fogelstrand L, Rhodin H, Hallböök H. Cytogenetic aberrations in adult acute lymphoblastic leukemia-A population-based study. EJHaem. 2021;2(4):813-7.

48. Inaba T, Roberts WM, Shapiro LH, Jolly KW, Raimondi SC, Smith SD, et al. Fusion of the leucine zipper gene HLF to the E2A gene in human acute B-lineage leukemia. Science. 1992;257(5069):531-4.

49. Meyer C, Hofmann J, Burmeister T, Gröger D, Park TS, Emerenciano M, et al. The MLL recombinome of acute leukemias in 2013. Leukemia. 2013;27(11):2165-76.

50. Meyers S, Alberti-Servera L, Gielen O, Erard M, Swings T, De Bie J, et al. Monitoring of Leukemia Clones in B-cell Acute Lymphoblastic Leukemia at

Diagnosis and During Treatment by Single-cell DNA Amplicon Sequencing. Hemasphere. 2022;6(4):e700.

51. Richard-Carpentier G, Kantarjian HM, Tang G, Yin CC, Khoury JD, Issa GC, et al. Outcomes of acute lymphoblastic leukemia with KMT2A (MLL) rearrangement: the MD Anderson experience. Blood Adv. 2021;5(23):5415-9.

52. Roberts KG, Morin RD, Zhang J, Hirst M, Zhao Y, Su X, et al. Genetic alterations activating kinase and cytokine receptor signaling in high-risk acute lymphoblastic leukemia. Cancer Cell. 2012;22(2):153-66.

53. Tasian SK, Loh ML, Hunger SP. Philadelphia chromosome-like acute lymphoblastic leukemia. Blood. 2017;130(19):2064-72.

54. Deininger MW, Goldman JM, Melo JV. The molecular biology of chronic myeloid leukemia. Blood. 2000;96(10):3343-56.

55. Ponzetto C, Guerrasio A, Rosso C, Avanzi G, Tassinari A, Zaccaria A, et al. ABL proteins in Philadelphia-positive acute leukaemias and chronic myelogenous leukaemia blast crises. Br J Haematol. 1990;76(1):39-44.

56. Gabert J, Beillard E, van der Velden VH, Bi W, Grimwade D, Pallisgaard N, et al. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia - a Europe Against Cancer program. Leukemia. 2003;17(12):2318-57.

57. Pui CH. Genomic and pharmacogenetic studies of childhood acute lymphoblastic leukemia. Frontiers of Medicine. 2015;9(1):1-9.

58. Burmeister T, Schwartz S, Bartram CR, Gökbuget N, Hoelzer D, Thiel E. Patients' age and BCR-ABL frequency in adult B-precursor ALL: a retrospective analysis from the GMALL study group. Blood. 2008;112(3):918-9.

59. Fielding AK, Rowe JM, Richards SM, Buck G, Moorman AV, Durrant IJ, et al. Prospective outcome data on 267 unselected adult patients with Philadelphia chromosome-positive acute lymphoblastic leukemia confirms superiority of allogeneic transplantation over chemotherapy in the pre-imatinib era: results from the International ALL Trial MRC UKALLXII/ECOG2993. Blood. 2009;113(19):4489-96.

60. Mullighan CG. How advanced are we in targeting novel subtypes of ALL? Best Pract Res Clin Haematol. 2019;32(4):101095.

61. Arber DA, Orazi A, Hasserjian RP, Borowitz MJ, Calvo KR, Kvasnicka HM, et al. International Consensus Classification of Myeloid Neoplasms and Acute Leukemias: integrating morphologic, clinical, and genomic data. Blood. 2022;140(11):1200-28.

62. Board WCoTE. WHO Classification of Tumours of Haematolymphoid Tumours. 5th ed. Cancer LIAfRo, editor2022.

63. Alaggio R, Amador C, Anagnostopoulos I, Attygalle AD, Araujo IBO, Berti E, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. Leukemia. 2022;36(7):1720-48.

64. Gu Z, Churchman ML, Roberts KG, Moore I, Zhou X, Nakitandwe J, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. Nat Genet. 2019;51(2):296-307.

65. Gu Z, Churchman M, Roberts K, Li Y, Liu Y, Harvey RC, et al. Genomic analyses identify recurrent MEF2D fusions in acute lymphoblastic leukaemia. Nat Commun. 2016;7:13331.

66. Hirabayashi S, Ohki K, Nakabayashi K, Ichikawa H, Momozawa Y, Okamura K, et al. ZNF384-related fusion genes define a subgroup of childhood B-cell precursor acute lymphoblastic leukemia with a characteristic immunotype. Haematologica. 2017;102(1):118-29.

67. Yasuda T, Tsuzuki S, Kawazu M, Hayakawa F, Kojima S, Ueno T, et al. Recurrent DUX4 fusions in B cell acute lymphoblastic leukemia of adolescents and young adults. Nat Genet. 2016;48(5):569-74.

68. Zhang J, McCastlain K, Yoshihara H, Xu B, Chang Y, Churchman ML, et al. Deregulation of DUX4 and ERG in acute lymphoblastic leukemia. Nat Genet. 2016;48(12):1481-9.

69. Zaliova M, Potuckova E, Hovorkova L, Musilova A, Winkowska L, Fiser K, et al. ERG deletions in childhood acute lymphoblastic leukemia with DUX4 rearrangements are mostly polyclonal, prognostically relevant and their detection rate strongly depends on screening method sensitivity. Haematologica. 2019;104(7):1407-16.

70. Russell LJ, Jones L, Enshaei A, Tonin S, Ryan SL, Eswaran J, et al. Characterisation of the genomic landscape of CRLF2-rearranged acute lymphoblastic leukemia. Genes Chromosomes Cancer. 2017;56(5):363-72.

71. Hanada I, Terui K, Ikeda F, Toki T, Kanezaki R, Sato T, et al. Gene alterations involving the CRLF2-JAK pathway and recurrent gene deletions in Down syndrome-associated acute lymphoblastic leukemia in Japan. Genes Chromosomes Cancer. 2014;53(11):902-10.

72. Nutt SL, Heavey B, Rolink AG, Busslinger M. Pillars Article: Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. Nature. 1999. 401: 556-562. J Immunol. 1999;195(3):766-72.

73. Mullighan CG, Su X, Zhang J, Radtke I, Phillips LA, Miller CB, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. N Engl J Med. 2009;360(5):470-80.

74. Den Boer ML, van Slegtenhorst M, De Menezes RX, Cheok MH, Buijs-Gladdines JG, Peters ST, et al. A subtype of childhood acute lymphoblastic

leukaemia with poor treatment outcome: a genome-wide classification study. Lancet Oncol. 2009;10(2):125-34.

75. Roberts KG, Li Y, Payne-Turner D, Harvey RC, Yang YL, Pei D, et al. Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. N Engl J Med. 2014;371(11):1005-15.

76. Lilljebjorn H, Henningsson R, Hyrenius-Wittsten A, Olsson L, Orsmark-Pietras C, von Palffy S, et al. Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia. Nat Commun. 2016;7:11790.

77. Mullighan CG, Miller CB, Radtke I, Phillips LA, Dalton J, Ma J, et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. Nature. 2008;453(7191):110-4.

78. Kuiper RP, Waanders E, van der Velden VH, van Reijmersdal SV, Venkatachalam R, Scheijen B, et al. IKZF1 deletions predict relapse in uniformly treated pediatric precursor B-ALL. Leukemia. 2010;24(7):1258-64.

79. Waanders E, van der Velden VH, van der Schoot CE, van Leeuwen FN, van Reijmersdal SV, de Haas V, et al. Integrated use of minimal residual disease classification and IKZF1 alteration status accurately predicts 79% of relapses in pediatric acute lymphoblastic leukemia. Leukemia. 2011;25(2):254-8.

80. Ofverholm I, Tran AN, Heyman M, Zachariadis V, Nordenskjöld M, Nordgren A, et al. Impact of IKZF1 deletions and PAX5 amplifications in pediatric B-cell precursor ALL treated according to NOPHO protocols. Leukemia. 2013;27(9):1936-9.

81. Olsson L, Ivanov Ofverholm I, Noren-Nystrom U, Zachariadis V, Nordlund J, Sjogren H, et al. The clinical impact of IKZF1 deletions in paediatric B-cell precursor acute lymphoblastic leukaemia is independent of minimal residual disease stratification in Nordic Society for Paediatric Haematology and Oncology treatment protocols used between 1992 and 2013. Br J Haematol. 2015;170(6):847-58.

82. Moorman AV, Enshaei A, Schwab C, Wade R, Chilton L, Elliott A, et al. A novel integrated cytogenetic and genomic classification refines risk stratification in pediatric acute lymphoblastic leukemia. Blood. 2014;124(9):1434-44.

83. Hamadeh L, Enshaei A, Schwab C, Alonso CN, Attarbaschi A, Barbany G, et al. Validation of the United Kingdom copy-number alteration classifier in 3239 children with B-cell precursor ALL. Blood Adv. 2019;3(2):148-57.

84. Stanulla M, Dagdan E, Zaliova M, Moricke A, Palmi C, Cazzaniga G, et al. IKZF1(plus) Defines a New Minimal Residual Disease-Dependent Very-Poor Prognostic Profile in Pediatric B-Cell Precursor Acute Lymphoblastic Leukemia. J Clin Oncol. 2018;36(12):1240-9.

85. Hunger SP, Mullighan CG. Acute Lymphoblastic Leukemia in Children. N Engl J Med. 2015;373(16):1541-52.

86. Guru Murthy GS, Pondaiah SK, Abedin S, Atallah E. Incidence and survival of T-cell acute lymphoblastic leukemia in the United States. Leuk Lymphoma. 2019;60(5):1171-8.

87. Steimlé T, Dourthe ME, Alcantara M, Touzart A, Simonin M, Mondesir J, et al. Clinico-biological features of T-cell acute lymphoblastic leukemia with fusion proteins. Blood Cancer J. 2022;12(1):14.

88. Dai YT, Zhang F, Fang H, Li JF, Lu G, Jiang L, et al. Transcriptome-wide subtyping of pediatric and adult T cell acute lymphoblastic leukemia in an international study of 707 cases. Proc Natl Acad Sci U S A. 2022;119(15):e2120787119.

89. Ferrando AA, Neuberg DS, Staunton J, Loh ML, Huard C, Raimondi SC, et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell. 2002;1(1):75-87.

90. Pölönen P, Di Giacomo D, Seffernick AE, Elsayed A, Kimura S, Benini F, et al. The genomic basis of childhood T-lineage acute lymphoblastic leukaemia. Nature. 2024;632(8027):1082-91.

91. Duffield AS, Mullighan CG, Borowitz MJ. International Consensus Classification of acute lymphoblastic leukemia/lymphoma. Virchows Arch. 2023;482(1):11-26.

92. Begley CG, Aplan PD, Davey MP, Nakahara K, Tchorz K, Kurtzberg J, et al. Chromosomal translocation in a human leukemic stem-cell line disrupts the T-cell antigen receptor delta-chain diversity region and results in a previously unreported fusion transcript. Proc Natl Acad Sci U S A. 1989;86(6):2031-5.

93. Xia Y, Brown L, Yang CY, Tsan JT, Siciliano MJ, Espinosa R, 3rd, et al. TAL2, a helix-loop-helix gene activated by the (7;9)(q34;q32) translocation in human T-cell leukemia. Proc Natl Acad Sci U S A. 1991;88(24):11416-20.

94. Hatano M, Roberts CW, Minden M, Crist WM, Korsmeyer SJ. Deregulation of a homeobox gene, HOX11, by the t(10;14) in T cell leukemia. Science. 1991;253(5015):79-82.

95. Bernard OA, Busson-LeConiat M, Ballerini P, Mauchauffé M, Della Valle V, Monni R, et al. A new recurrent and specific cryptic translocation, t(5;14)(q35;q32), is associated with expression of the Hox11L2 gene in T acute lymphoblastic leukemia. Leukemia. 2001;15(10):1495-504.

96. Natkunam Y, Zhao S, Mason DY, Chen J, Taidi B, Jones M, et al. The oncoprotein LMO2 is expressed in normal germinal-center B cells and in human B-cell lymphomas. Blood. 2007;109(4):1636-42.

97. Nagel S, Kaufmann M, Drexler HG, MacLeod RA. The cardiac homeobox gene NKX2-5 is deregulated by juxtaposition with BCL11B in pediatric T-ALL cell lines via a novel t(5;14)(q35.1;q32.2). Cancer Res. 2003;63(17):5329-34.

98. Homminga I, Pieters R, Langerak AW, de Rooi JJ, Stubbs A, Verstegen M, et al. Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia. Cancer Cell. 2011;19(4):484-97.

99. Montefiori LE, Bendig S, Gu Z, Chen X, Pölönen P, Ma X, et al. Enhancer Hijacking Drives Oncogenic BCL11B Expression in Lineage-Ambiguous Stem Cell Leukemia. Cancer Discov. 2021;11(11):2846-67.

100. Fang H, Wang W, El Hussein S, Morita K, Beird HC, Mitra A, et al. B-cell lymphoma/leukaemia 11B (BCL11B) expression status helps distinguish early T-cell precursor acute lymphoblastic leukaemia/lymphoma (ETP-ALL/LBL) from other subtypes of T-cell ALL/LBL. Br J Haematol. 2021;194(6):1034-8.

101. De Bie J, Quessada J, Tueur G, Lefebvre C, Luquet I, Toujani S, et al. Cytogenetics in the management of T-cell acute lymphoblastic leukemia (T-ALL): Guidelines from the Groupe Francophone de Cytogénétique Hématologique (GFCH). Curr Res Transl Med. 2023;71(4):103431.

102. Ferguson-Smith MA. History and evolution of cytogenetics. Mol Cytogenet. 2015;8:19.

103. Durmaz AA, Karaca E, Demkow U, Toruner G, Schoumans J, Cogulu O. Evolution of genetic techniques: past, present, and beyond. Biomed Res Int. 2015;2015:461524.

104. Noll AC, Miller NA, Smith LD, Yoo B, Fiedler S, Cooley LD, et al. Clinical detection of deletion structural variants in whole-genome sequences. NPJ Genom Med. 2016;1:16026.

105. Borowitz MJ, Devidas M, Hunger SP, Bowman WP, Carroll AJ, Carroll WL, et al. Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia and its relationship to other prognostic factors: a Children's Oncology Group study. Blood. 2008;111(12):5477-85.

106. Szczepański T. Why and how to quantify minimal residual disease in acute lymphoblastic leukemia? Leukemia. 2007;21(4):622-6.

107. Zhang J, Oak J. Challenges of detecting measurable/minimal disease in acute leukemia. Semin Diagn Pathol. 2023;40(3):216-20.

108. Thörn I, Forestier E, Thuresson B, Wasslavik C, Malec M, Li A, et al. Applicability of IG/TCR gene rearrangements as targets for minimal residual disease assessment in a population-based cohort of Swedish childhood acute lymphoblastic leukaemia diagnosed 2002-2006. Eur J Haematol. 2010;84(2):117-27.

109. Sigaux F. The V(D)J recombination in acute lymphoid leukemias: a short review. Leuk Lymphoma. 1994;13 Suppl 1:53-7.

110. van der Velden VH, Cazzaniga G, Schrauder A, Hancock J, Bader P, Panzer-Grumayer ER, et al. Analysis of minimal residual disease by Ig/TCR gene rearrangements: guidelines for interpretation of real-time quantitative PCR data. Leukemia. 2007;21(4):604-11.

111. Szczepański T, Willemse MJ, Brinkhof B, van Wering ER, van der Burg M, van Dongen JJ. Comparative analysis of Ig and TCR gene rearrangements at diagnosis and at relapse of childhood precursor-B-ALL provides improved strategies for selection of stable PCR targets for monitoring of minimal residual disease. Blood. 2002;99(7):2315-23.

112. Arunachalam AK, Janet NB, Korula A, Lakshmi KM, Kulkarni UP, Aboobacker FN, et al. Prognostic value of MRD monitoring based onBCR-ABL1copy numbers in Philadelphia chromosome positive acute lymphoblastic leukemia. Leukemia & Lymphoma. 2020;61(14):3468-75.

113. Scheuring UJ, Pfeifer H, Wassmann B, Brück P, Atta J, Petershofen EK, et al. Early minimal residual disease (MRD) analysis during treatment of Philadelphia chromosome/Bcr-Abl-positive acute lymphoblastic leukemia with the Abl-tyrosine kinase inhibitor imatinib (ST1571). Blood. 2003;101(1):85-90.

114. Pane F, Cimino G, Izzo B, Camera A, Vitale A, Quintarelli C, et al. Significant reduction of the hybrid BCR/ABL transcripts after induction and consolidation therapy is a powerful predictor of treatment response in adult Philadelphia-positive acute lymphoblastic leukemia. Leukemia. 2005;19(4):628-35.

115. Yanada M, Sugiura I, Takeuchi J, Akiyama H, Maruta A, Ueda Y, et al. Prospective monitoring of BCR-ABL1 transcript levels in patients with Philadelphia chromosome-positive acute lymphoblastic leukaemia undergoing imatinib-combined chemotherapy. British Journal of Haematology. 2008;143(4):503-10.

116. Cazzaniga G, De Lorenzo P, Alten J, Röttgers S, Hancock J, Saha V, et al. Predictive value of minimal residual disease in Philadelphia-chromosome-positive acute lymphoblastic leukemia treated with imatinib in the European intergroup study of post-induction treatment of Philadelphia-chromosome-positive acute lymphoblastic leukemia, based on immunoglobulin/T-cell receptor and BCR/ABL1 methodologies. Haematologica. 2018;103(1):107-15.

117. Zaliova M, Fronkova E, Krejcikova K, Muzikova K, Mejstrikova E, Stary J, et al. Quantification of fusion transcript reveals a subgroup with distinct biological properties and predicts relapse in BCR/ABL-positive ALL: implications for residual disease monitoring. Leukemia. 2009;23(5):944-51.

118. Hovorkova L, Zaliova M, Venn NC, Bleckmann K, Trkova M, Potuckova E, et al. Monitoring of childhood ALL using BCR-ABL1 genomic breakpoints identifies a subgroup with CML-like biology. Blood. 2017;129(20):2771-81.

119. Martinez RJ, Kang Q, Nennig D, Bailey NG, Brown NA, Betz BL, et al. One-Step Multiplexed Droplet Digital Polymerase Chain Reaction for Quantification of p190 BCR-ABL1 Fusion Transcript in B-Lymphoblastic Leukemia. Arch Pathol Lab Med. 2022;146(1):92-100.

120. Kim Y, Kim S, Lee JM, Ahn A, Yoo JW, Lee JW, et al. Discontinuation of tyrosine kinase inhibitors based on BCR-ABL1 monitoring by digital droplet PCR in pediatric chronic myeloid leukemia. Frontiers in Pediatrics. 2022;10.

121. Stranneheim H, Lagerstedt-Robinson K, Magnusson M, Kvarnung M, Nilsson D, Lesko N, et al. Integration of whole genome sequencing into a healthcare setting: high diagnostic rates across multiple clinical entities in 3219 rare disease patients. Genome Med. 2021;13(1):40.

122. Cuppen E, Elemento O, Rosenquist R, Nikic S, M IJ, Zaleski ID, et al. Implementation of Whole-Genome and Transcriptome Sequencing Into Clinical Cancer Care. JCO Precis Oncol. 2022;6:e2200245.

123. Garcia M, Juhos S, Larsson M, Olason PI, Martin M, Eisfeldt J, et al. Sarek: A portable workflow for whole-genome sequencing analysis of germline and somatic variants. F1000Res. 2020;9:63.

124. Clinical genomics S. Mutation Identification Pipeline [Available from: https://github.com/Clinical-Genomics/MIP.

125. Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics. 2012;28(18):i333-i9.

126. Eisfeldt J, Vezzi F, Olason P, Nilsson D, Lindstrand A. TIDDIT, an efficient and comprehensive structural variant caller for massive parallel sequencing data. F1000Res. 2017;6:664.

127. Abyzov A, Urban AE, Snyder M, Gerstein M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome Res. 2011;21(6):974-84.

128. Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics. 2016;32(8):1220-2.

129. Asl HF. BALSAMIC: Github; [Available from: https://github.com/Clinical-Genomics/BALSAMIC.

130. Lindstrand A, Eisfeldt J, Pettersson M, Carvalho CMB, Kvarnung M, Grigelioniene G, et al. From cytogenetics to cytogenomics: whole-genome sequencing as a first-line test comprehensively captures the diverse spectrum of disease-causing genetic variation underlying intellectual disability. Genome Med. 2019;11(1):68.

131. nf-core/ranfusion. RNA-seq analysis pipeline for detection of gene-fusions github2023 [Available from: https://github.com/nf-core/rnafusion?tab=readme-ov-file.

132. Beder T, Hansen BT, Hartmann AM, Zimmermann J, Amelunxen E, Wolgast N, et al. The Gene Expression Classifier ALLCatchR Identifies B-cell Precursor ALL Subtypes and Underlying Developmental Trajectories Across Age. Hemasphere. 2023;7(9):e939.

133. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24-6.

134. Van Loo P, Nordgard SH, Lingjærde OC, Russnes HG, Rye IH, Sun W, et al. Allele-specific copy number analysis of tumors. Proc Natl Acad Sci U S A. 2010;107(39):16910-5.

135. SMD-Bioinformatics-Lund. Gens [Available from: https://github.com/SMD-Bioinformatics-Lund/gens.

136. Rezayee F, Eisfeldt J, Skaftason A, Öfverholm I, Sayyab S, Syvänen AC, et al. Feasibility to use whole-genome sequencing as a sole diagnostic method to detect genomic aberrations in pediatric B-cell acute lymphoblastic leukemia. Front Oncol. 2023;13:1217712.

137. Clinical genomics S. SCOUT: Github; [Available from: https://github.com/Clinical-Genomics/scout

138. Russell LJ, Capasso M, Vater I, Akasaka T, Bernard OA, Calasanz MJ, et al. Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. Blood. 2009;114(13):2688-98.

139. Nordlund J, Marincevic-Zuniga Y, Cavelier L, Raine A, Martin T, Lundmark A, et al. Refined detection and phasing of structural aberrations in pediatric acute lymphoblastic leukemia by linked-read whole-genome sequencing. Sci Rep. 2020;10(1):2512.

140. Lemmers RJ, van der Vliet PJ, Klooster R, Sacconi S, Camaño P, Dauwerse JG, et al. A unifying genetic model for facioscapulohumeral muscular dystrophy. Science. 2010;329(5999):1650-3.

141. Ryan SL, Peden JF, Kingsbury Z, Schwab CJ, James T, Polonen P, et al. Whole genome sequencing provides comprehensive genetic testing in childhood B-cell acute lymphoblastic leukaemia. Leukemia. 2023.

142. Duncavage EJ, Coleman JF, de Baca ME, Kadri S, Leon A, Routbort M, et al. Recommendations for the Use of in Silico Approaches for Next-Generation Sequencing Bioinformatic Pipeline Validation: A Joint Report of the Association for Molecular Pathology, Association for Pathology Informatics, and College of American Pathologists. J Mol Diagn. 2023;25(1):3-16.

143. Worst BC, van Tilburg CM, Balasubramanian GP, Fiesel P, Witt R, Freitag A, et al. Next-generation personalised medicine for high-risk paediatric cancer patients - The INFORM pilot study. Eur J Cancer. 2016;65:91-101.

144. Tang G, Sydney Sir Philip JK, Weinberg O, Tam W, Sadigh S, Lake JI, et al. Hematopoietic neoplasms with 9p24/JAK2 rearrangement: a multicenter study. Mod Pathol. 2019;32(4):490-8.

145. Roberts KG, Yang YL, Payne-Turner D, Lin W, Files JK, Dickerson K, et al. Oncogenic role and therapeutic targeting of ABL-class and JAK-STAT activating kinase alterations in Ph-like ALL. Blood Adv. 2017;1(20):1657-71.

146. Downes CE, McClure BJ, McDougal DP, Heatley SL, Bruning JB, Thomas D, et al. JAK2 Alterations in Acute Lymphoblastic Leukemia: Molecular Insights for Superior Precision Medicine Strategies. Front Cell Dev Biol. 2022;10:942053.

147. Ding YY, Stern JW, Jubelirer TF, Wertheim GB, Lin F, Chang F, et al. Clinical efficacy of ruxolitinib and chemotherapy in a child with Philadelphia chromosome-like acute lymphoblastic leukemia with GOLGA5-JAK2 fusion and induction failure. Haematologica. 2018;103(9):e427-e31.

148. Krstic A, Rezayee F, Saft L, Hammarsjö A, Svenberg P, Barbany G. Case Report: Whole genome sequencing identifies CCDC88C as a novel JAK2 fusion partner in pediatric T-cell acute lymphoblastic leukemia. Front Pediatr. 2022;10:1082986.

149. Huret JL, Ahmad M, Arsaban M, Bernheim A, Cigna J, Desangles F, et al. Atlas of genetics and cytogenetics in oncology and haematology in 2013. Nucleic Acids Res. 2013;41(Database issue):D920-4.

150. Janssen JW, Ludwig WD, Sterry W, Bartram CR. SIL-TAL1 deletion in T-cell acute lymphoblastic leukemia. Leukemia. 1993;7(8):1204-10.

151. Audano PA, Sulovari A, Graves-Lindsay TA, Cantsilieris S, Sorensen M, Welch AE, et al. Characterizing the Major Structural Variant Alleles of the Human Genome. Cell. 2019;176(3):663-75.e19.

152. Arthur C, Rezayee F, Mogensen N, Saft L, Rosenquist R, Nordenskjöld M, et al. Patient-Specific Assays Based on Whole-Genome Sequencing Data to Measure Residual Disease in Children With Acute Lymphoblastic Leukemia: A Proof of Concept Study. Front Oncol. 2022;12:899325.

153. Schwarz AK, Stanulla M, Cario G, Flohr T, Sutton R, Möricke A, et al. Quantification of free total plasma DNA and minimal residual disease detection in the plasma of children with acute lymphoblastic leukemia. Ann Hematol. 2009;88(9):897-905.

154. Bastian L, Beder T, Barz MJ, Bendig S, Bartsch L, Walter W, et al. Developmental trajectories and cooperating genomic events de fi ne molecular subtypes of BCR:: ABL1-positive ALL. Blood. 2024;143(14).

155. Berglund E, Barbany G, Orsmark-Pietras C, Fogelstrand L, Abrahamsson J, Golovleva I, et al. A Study Protocol for Validation and Implementation of Whole-Genome and -Transcriptome Sequencing as a Comprehensive Precision Diagnostic Test in Acute Leukemias. Front Med (Lausanne). 2022;9:842507.