Investigating Ikaros deletions in a cohort of South African Acute Lymphoblastic Leukemia patients

Research by Mishalan Moodley: 0108860T

A dissertation submitted to the Faculty of Health Sciences, University of Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Master of Science

Promotor: Dr. P Willem

Co-promotor: Desmond Schnugh

Declaration

Investigating Ikaros deletions in a cohort of South African Acute Lymphoblastic Leukemia patients

I, Mishalan Moodley declare that this dissertation is my own work, unless otherwise stated in the text. It is being submitted for the degree of Bachelor of Science with Masters (MSc Med) at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other university.

……

Mishalan Moodley

6 Day of May 2014

Dedication

Sadhu Ram Ji

Investigating Ikaros deletions in a cohort of South African Acute Lymphoblastic Leukemia patients

Abstract

INTRODUCTION: Despite best current therapy, acute lymphoblastic leukemia (ALL) still remains the most common cause of cancer-related death in children and young adults. Relapse is the main reason for treatment failure in ALL patients and occurs in 15-20% of these patients. Current risk stratification criteria have not been sufficient to predict relapse in ALL patients. The Philadelphia (Ph) chromosome is a chromosomal abnormality found in a subset of high risk ALL patients and is associated with a poor prognosis. Recent genome-wide studies have identified focal deletions of the Ikaros gene (*IKZF1*) in 70-80% of B-cell ALL patients that have the Philadelphia (Ph) chromosome. Subsequent studies have also found a strong correlation between *IKZF1* deletions and ALL patients (Ph+ and Ph-) that relapse. *IKZF1* is required for normal lymphoid development and loss of *IKZF1* results in haploidinsufficiency or the overexpression of dominant negative *IKZF1* isoforms, in particular Ik6 in high risk ALL patients. Most studies used DNA microarrays to detect *IKZF1* deletions. Multiplex ligation probe dependent amplification (MLPA) is a low cost, rapid technique that can detect small DNA copy number changes of up to 50 targets in a single reaction and is not as technically challenging to analyse as arrays. MLPA has also been suggested to be used as an alternative to array based techniques in developing countries. **METHODS:** There were 31 ALL (paediatric and adult) patients that were tested using MLPA and 24 ALL patients tested using reverse transcriptase PCR (RT-PCR) to detect *IKZF1* copy number changes and *IKZF1* isoform expression pattern respectively. RT-PCR was validated with DNA sequencing and MLPA was validated with Fluorescent *in situ* hybridization (FISH). MLPA was also compared to cytogenetics in certain cases. **RESULTS:** MLPA detected 156 copy number changes (7.1 aberrations per sample) in 22 leukemic patients. *IKZF1* deletions accounted for the majority of the aberrations (41%) and were detected in 53% of Ph+ ALL patients (n=15) by MLPA. *IKZF1* deletions were detected at presentation and relapse in Ph+ and Ph- ALL patients. *IKZF1* isoform Ik6 was detected in 70% of Ph+ and relapsed ALL patients after performing RT-PCR. *IKZF1* deletions of exons 4-7 resulted in exclusive expression of Ik6. MLPA results were also correlated with certain aneuploidies detected

with cytogenetics. **CONCLUSION:** This study showed that *IKZF1* deletions could have assisted with prognosis in certain ALL cases and thus, newly diagnosed ALL patients should be screened for *IKZF1* deletions. MLPA proved to be a reliable, rapid and cost effective technique to detect small copy number changes in multiple genes and should be implemented as a diagnostic test to detect *IKZF1* deletions.

Acknowledgments

I would like to thank Sadhu Ram for the guidance and strength that he has given me over the last two years. Without his constant presence, completing my thesis to the best of my ability would not have been possible.

I would also like to thank:

My supervisor, Dr. Pascale Willem for all the time she took out of her busy schedule to assist me (always with a smile), for her priceless advice, support and encouragement.

My co supervisor Desmond Schnugh for his guidance and support.

Dr. Wainright and her team at the paediatric oncology unit for their interest in this project.

All the individuals who participated in this study.

Ria Lassauniere for her assistance with MLPA analysis.

The cytogenetics staff for their support and assistance.

Any other friends or family who intentionally or unintentionally assisted in this thesis.

The University of the Witwatersrand and the National Health Laboratory Services.

Table of contents

List of figures

List of tables

Nomenclature

ALL: Acute lymphoblastic leukemia

AML: Acute myeloid leukemia

BAC: bacterial artificial chromosome

Bps: base pairs

BCR-ABL1: break point cluster region-abelson

BTG1: B-cell translocation gene 1

CDKN2: cyclin-dependent kinase inhibitor 2

CML: Chronic myeloid leukemia

EBF1: early B-cell factor 1

ETV6: ets variant gene 6

CRLF2: Cytokine receptor-like factor 2

CSF2RA: colony stimulating factor 2 receptor

dNTP: Deoxyribonucleotide triphosphate

ddNTP: Dideoxynucleotide triphosphate

DQ: Dosage Quotient

FISH: Fluorescent *in situ* hybridization

GAPDH: glyceraldehyde-3-phosphate dehydrogenase

HOX: Homeobox gene

IKZF1: Ikaros gene

Ik1-10: IKZF1 isoform 1-10

IL3RA: interleukin 3 receptor, alpha

Jak2: Janus kinase 2

$MgCl_2$: magnesium chloride

MLPA: multiplex ligation-dependant probe amplification

NaOH: sodium hydroxide

NCBI: National Center for Biotechnology Information

NTC: non-template control

PAX5: paired box 5

PCR: polymerase chain reaction

Ph: Philadelphia chromosome

QR-PCR quantitative real-time PCR

Rb: retinoblastoma

RT-PCR: Reverse transcriptase-PCR

SNP: single nucleotide polymorphism

UNITS AND SYMBOLS

%: percentage

ml: millilitre

kb: kilobase

l: litre

mg: milligram

ml: millilitre

ng: nanograms

nm: nanometres

PBX1: Pre-B-cell leukemia transcription factor 1

RUNX: Runt-related transcription factor

µg: microgram

µl: microlitre

®: registered trademark

Chapter 1. Introduction

1.1 Acute Lymphoblastic Leukaemia

Acute lymphoblastic leukaemia (ALL) is a malignant disorder of the blood that affects adults and children. ALL is characterised by an excess number of white blood cells which originate from immature B and T lymphocytes in the bone marrow (Onciu, 2009). These cells are often referred to as lymphoblasts or leukemic cells because they proliferate and accumulate in the bone marrow (Reichmann, 2000). Leukemic cells proliferate uncontrollably and overcrowd the bone marrow consequently hindering the production of normal blood cells. Leukemic cells leak out of the bone marrow and accumulate in extra medullary sites, especially the lymph nodes, liver, spleen, central nervous system and gonads (Daly, 2010).

1.1.1 Epidemiology

ALL is the most common form of childhood cancer and constitutes 25% of all childhood malignancies globally (Garcia et al, 2007). In Europe, United States of America and Australia the cure rate of ALL is ~80% in children (0 – 14 years old) (Pui & Jeha, 2007). The epidemiology of ALL in children in South Africa has yet to be established. However, from cases that have been reported, it has been estimated that leukaemia is the most prevalent cancer in children and accounts for 26% of all childhood cancers (Stefan et al, 2010).

1.2 Clinical symptoms and diagnosis of ALL

The main symptoms of ALL include: anaemia, increased bleeding and bruising, infections and bone pain. Initially a full blood count of red and white blood cells and platelets are done to diagnose ALL. These three types of cells can be abnormal in ALL (Daly, 2010).

Children with ALL usually have low platelet and low red blood cell (haemoglobin) levels accompanied with high white cell count with an excess amount of immature blast cells. The blast cells do not mature properly, instead they grow and proliferate uncontrollably and also hinder the development of normal blood cells. ALL is confirmed by the percentage of blast cells in the bone marrow of a patient. Under normal physiological conditions there are less than 5% blast cells present in leukemic free individuals. In ALL patients blast cells may range between 20% - 95% (Daly, 2010).

White blood cells eventually develop into two types of cells: B-cells or T-cells. ALL affects both B-cells and T-cells. B-cell ALL is found in ~85% and T-cell ALL is found in ~15% of ALL cases. Both subtypes are characterised by assessing the morphology of cells in a blood or bone marrow specimen (Daly, 2010). ALL has been classified into different systems based on morphology, cytochemistry and immunophenotyping.

The traditional criteria used to classify ALL are based on the French-American-British (FAB) group (Bennet et al, 1976). The FAB group defines the 3 subtypes of ALL (L1, L2 and L3) based on their morphology when viewed under a microscope. These features include the cell size, prominence of nucleoli, appearance of the cytoplasm and the presence of cytoplasmic basophilia and cytoplasmic vacuoles (Bennett et al, 1985). The ALL-L1 subtype involves mainly small size blast cells and is found in the majority (70-80%) of childhood ALL cases. ALL-L2 subtype involves a mixed group of small and large blast cells with the majority being large sized. The ALL-L3 subtype involves medium to large sized blast cells. The European Group for the Immunological Classification of Leukemias (EGIL) classifies acute leukemia solely on the basis of immunophenotyping (Abdul-Hamid, 2011).

1.2.1 Immunophenotyping

Immunophenotyping is a crucial component that is used to classify cases according to the different stages of B-cell and T-cell maturation.

T-cell ALL is characterised by an abnormally high white blood cell count and it involves the central nervous system in ~50% of the cases. In most cases the thymus gland in the centre of the chest (mediastinal mass) is enlarged, this can be viewed using X-rays during diagnosis (Daly, 2010).

The different subtypes of B-cells that are produced during maturation are: common precursor B-cell ALL, Pro-B ALL and Pre-B ALL. These subtypes are characterised by the presence of certain antigens on their surface and they can be distinguished by using antibody staining or immunophenotyping. Flow cytometry is the method used for immunophenotyping (Jennings & Foon, 1997) which can distinguish between different B-cells by detecting the surface antigens also known as clusters of differentiation (CD). CD antigens are present during each stage of B-cell development. A positive reaction for CD10,

CD19, CD20, CD22, CD34 and CD38 antigens identifies the specific B ALL subtype (Nakase et al, 1996). In 80% of childhood cases, ALL arises from the early stages of B-cell development in the bone marrow. The type of ALL in these cases are referred to as precursor B-cell ALL (Pre-B-cell ALL). Pre-B-cell ALL can be further subdivided into 4 different categories:

- Common ALL accounts for around 80% of pre-B ALL. The CD10 antigen is expressed on the surface of these cells. This antigen is also referred to as the Common Acute Lymphocytic Leukemia Antigen (cALLa) and is associated with a good prognosis.

- Pro B-ALL expresses human leucocyte antigen-DR (HLA-DR), terminal deoxynucleotidyl transferase (TdT) and CD19 and is negative for CD10. This type of ALL accounts for 10% of adult ALL cases.

- Pre B-ALL expresses cytoplasmic immunoglobin and CD10 and is also found in around 10% of adult ALL cases.

- Mature B-ALL is found in around 4% of adult ALL cases. It arises in more mature developing lymphocytes. Leukemic cells spread to regions outside of the blood and bone marrow and lymphoblasts can accumulate in the neck, head and abdominal regions. Mature B-ALL is biologically similar to Burkitt's lymphoma (Abdul-Hamid, 2011).

Mature B-ALL corresponds to the L3 FAB subgroup, the other 3 ALL subtypes correspond to the L1 and L2 FAB subgroups (Abdul-Hamid, 2011).

1.2.2 Cytogenetic and molecular genetic aberrations in ALL

The karyotype and the genetic makeup of tumour cells have an important prognostic impact in ALL patients. Cytogenetic analysis and molecular genetic techniques have been used to determine the status of critical genes that provide an indication of the severity of disease. It has been shown that entire chromosomes can be gained or genes can be rearranged after a translocation event, consequently fusion genes are produced which influence the pathophysiology. This genetic information assists doctors and pathologists with prognosis and choosing appropriate therapy treatment (Armstrong et al, 2002).

Several cytogenetic and molecular genetic aberrations have been detected in paediatric (age 0-14 years) ALL patients. An illustration of these abnormalities and the proportion in which they have been detected in paediatric cases is shown in figure 1.1.

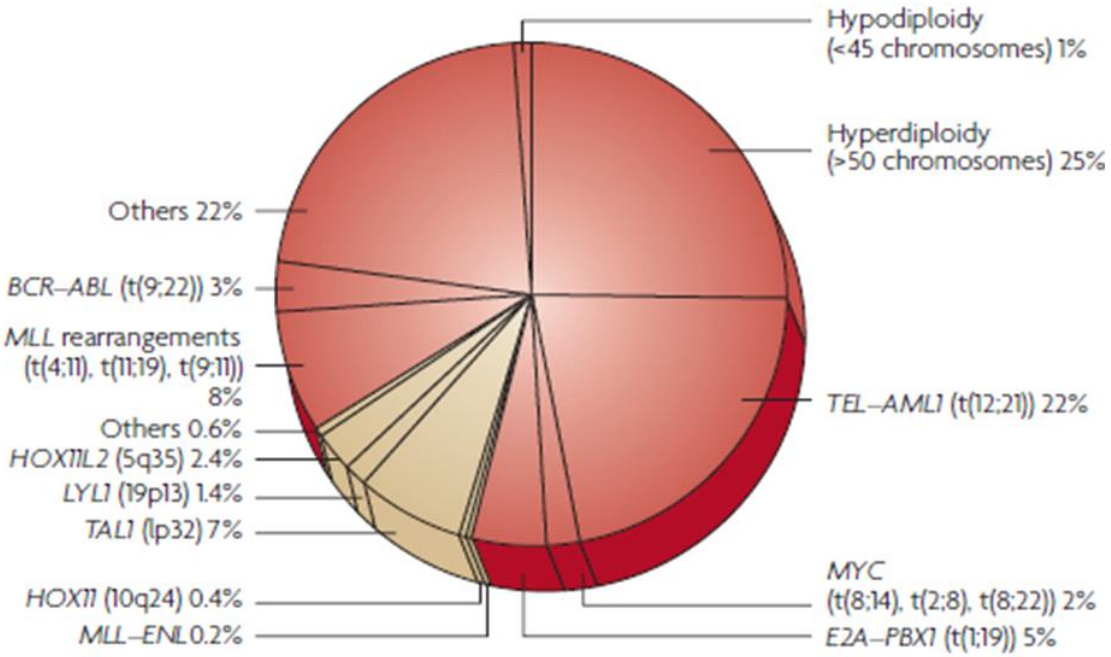

Figure 1.1 Frequency of Cytogenetic and molecular genetic aberrations detected in paediatric ALL patients
B cells (beige colour) and T cells (red colour). Illustration taken from Pui & Jeha, 2007.

The genetic aberrations illustrated in figure 1.1 have important prognostic impacts that assists in management of patients. In summary: Hyperdiploidy (>50 chromosomes present in a leukemic cell) and the *TEL-AML1* fusion gene are found in 25% and 22% of all childhood ALL cases respectively and both are associated with a good prognosis (Pui & Jeha, 2007).

In contrast hypodiploidy (less than 45 chromosomes), the break point cluster region-abelson (*BCR-ABL1*) fusion gene (also known as the Philadelphia chromosome (Ph), t(9;22)(q34;q11)) and the *MLL-AF4* t(4;11)(q21;q23) fusion gene are associated with a poor prognosis in ALL patients (Pui & Jeha, 2007).

The t(1;19)(q23;p13.3) genetic translocation is found in 5% of ALL patients. This genetic translocation results in the juxtaposition of the transcription factor E2A on chromosome 19 to the homeobox (*HOX*) gene on chromosome 1. The E2A gene affects lymphoid development. Dysregulated HOX expression has a negative impact on leukemic stem cells

transformation (Argiropoulos & Humphries, 2007). Ultimately the t(1;19)(q23;p13.3) genetic translocation leads to reduced levels of another gene, Pre-B-cell leukemia transcription factor 1 (*PBX1*), which causes deregulation of lymphoid development and is associated with a poor prognosis (Aspland et al, 2001).

The myeloid lymphoid leukaemia (*MLL)* gene on chromosome 11 can be rearranged and fused with 40 different partners. *MLL* regulates *HOX* activity during hematopoiesis. Rearrangements involving *MLL* with other genes results in reduced *MLL* function are associated with a poor prognosis (Pui & Jeha, 2007). It often fuses with the *AF4* gene on chromosome 4, this rearrangement is known as the t(4;11)(q21;q23) translocation. This abnormality is most common in children and occurs in 85% of ALL children younger than 1 year of age (Secker-Walker, 1998).

The myelocytomatosis (*c-MYC)* gene is an oncoprotein that regulates the cell cycle (Oster et al, 2002). C-*MYC* is frequently overexpressed in mature B-cell ALL and this results in cell-cycle progression (Eisenman, 2001). The t(8;14)(q24;q32) translocation involves the *c-MYC* gene on 8q24 is juxtaposed and the immunoglobin heavy chain (*IgH)* locus on 14q32 (Boerma et al, 2008). The t(8;14)(q24;q32) translocation results in *c-MYC* being placed under the influence of *IgH.* Consequently *MYC* expression is deregulated which results in tumorigenesis (Jankovic et al, 2009).

The t(12;21) (p13;q22) translocation has been identified in 22% of paediatric ALL patients. This genetic abnormality results from the juxtaposition of two genes, *TEL (ETV6* - ets variant gene 6*)* and the *AML1* gene which locate at 12p13 and 21q22.94, respectively. *ETV6* is a tumour suppressor gene. This abnormality is associated with a favourable prognosis when seen in isolation in children with pre-B ALL irrespective of their age or leukocyte count at presentation (Forestier et al, 2008).

1.3 Outcome and relapse in medium and high risk group ALL patients: the need for a prognostic marker

ALL patients are classified into different risk groups according to clinical and laboratory features such as immunophenotyping and genetic aberrations mentioned above. In addition to these factors, other factors that are used for risk assessment include the age of the

patient and a day-14 bone marrow response to treatment. Early response to treatment is determined by monitoring bone marrow morphology on day 7 and day 14 after induction therapy (Schultz et al, 2007). Induction therapy involves the use of combination chemotherapy with drugs such as daunorubicin, asparaginase, vincristine and dexamethasone (corticosteroid). These drugs are used as an attempt to induce a remission by destroying all leukemic cells in the blood and bone marrow. After induction therapy the patient is given consolidation therapy in an attempt to prevent the cancer from re-appearing which is often referred to as relapse. ALL patients younger than 10 years of age with a white blood cell (WBC) count of less than $50X10^9$cells/L (50 000cells/µL) are stratified as standard risk patients. The converse is true for high risk ALL patients, high risk patients are older than 10 years of age with a WBC > $50X10^9$cells/L (50 000cells/µL) (Daly, 2010).

Recently risk stratification in ALL patients has also included minimal residual disease (MRD) in response in treatment (Conter et al, 2010). Rearrangement of the T cell receptor gene occurs in more than 90%of childhood pre-B ALL (Szczepanski et al, 1999). There are two rearrangements that involve the T-cell receptor gene and the immunoglobulin (Ig) gene. MRD uses polymerase chain reaction (PCR) to assess these two markers with a sensitivity level of 10^{-4}or less at days 33 (time point 1 - TP1) and 78 (time point 2 - TP2) after treatment (Conter et al, 2010).Patients are classified as standard risk if they are negative for both markers at TP1 and TP2. Medium risk group patients are positive for one or both markers at TP1 and are less than a level of 10^{-3} at TP2 for both markers. Patients with a MRD of 10^{-3}or more at TP2 for any marker are classified as a high risk (Conter et al, 2010).

However, all of the mentioned factors have not been sufficient to identify ALL patients at risk for relapse which often leads to death. ALL is the most common cause of cancer-related death in children and young adults (Nguyen et al, 2008; Gaynon et al, 1998 & Saarinen-Pihkala et al, 2006). Although the cure rate and event free survival in newly diagnosed patients is 80% (in developed countries), there is still about 20% of ALL patients who relapse (Hunger et al, 2011). Very few ALL patients that relapse survive for more than 5 years and die of the disease. A study analysed the incidence of relapse in a cohort of 10 000 paediatric ALL patients (age 0 - 10 years) from 1988 to 2002. This study showed that only 11% of relapsed patients survive (within 18 months) (Nguyen et al, 2008). It was also shown that ALL relapse occurs in all currently defined risk groups: low, medium and high (Nguyen et al,

2008 & Pui & Jeha, 2007). Therefore, despite significant improvements in the survival rates of children with ALL, there has been no or very little improvement to predict relapse clearly indicating that the above mentioned factors were not effective in stratifying ALL patients into a risk of relapse. Thus recent research has focused on identifying prognostic markers to predict relapse in all ALL patients. Such markers would assist clinicians and pathologists to manage their patients better with a view to improve survival.

1.4 Ikaros (*IKZF1*) gene

Ikaros, often referred to as *IKZF1* is a transcription factor that binds to regulatory sequences of several genes in human lymphoid cells and acts as a central regulator of lymphoid differentiation and development (Nakase et al, 2000 & Hosokawa et al, 2000). *IKZF1* is also critical in the generation of B cell precursors (Georgopoulos et al, 1997). In addition it has been found that monoallelic and biallelic deletions of *IKZF1* in mice resulted in severe defects in B and T cell lymphoid development (Wang et al, 1996). Genome wide studies over the past 4 years have been carried out on ALL patients in various different countries and it was found that *IKZF1* was deleted in high risk ALL patients that relapse or have the Philadelphia (Ph) chromosome (Mullighan et al, 2008; Iacobucci et al, 2009 & Martinelli et al, 2009).

1.4.1 *IKZF1* structure and its' isoforms

The *IKZF1* gene is over 129 080 base pairs (bps) in length, is made up of 8 exons and is located on the human chromosome 7 band p11-p13. It has been reported that *IKZF1* can generate up to 13 different wild type isoforms each containing 2 C-terminal zinc finger domains through alternate splicing of its pre-mRNA (Sun et al 1999a; Iacobucci et al, 2008b & Mullighan et al, 2008). However, critical to *IKZF1* function are the amount of N - terminal zinc fingers present on each isoform. An illustration of the isoforms generated by *IKZF1* is shown in figure 1.2.

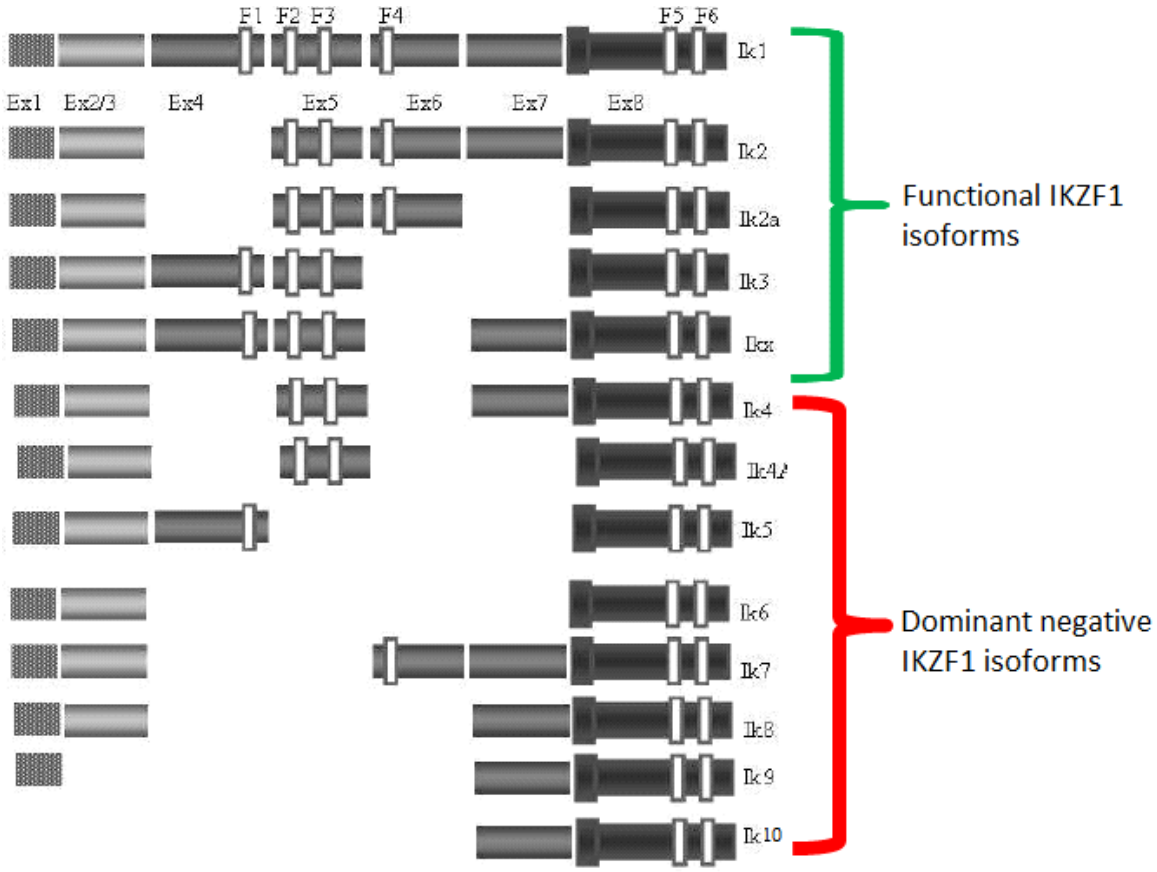

Figure 1.2. Various isoforms generated after alternate splicing of *IKZF1*
The vertical white bars (F1-F6) represent zinc fingers, N - terminal zinc fingers (F1-F4) are found in exons 4, 5 and 6. C - Terminal zinc fingers (F5 and F6) are found in exon 8. The green parenthesis represents the functional isoforms and the red parenthesis represents the dominant negative isoforms of *IKZF1*. Illustration taken and adapted from Mullighan et al, 2008.

Figure 1.2 depicts the functional isoforms (Ik1, Ik2, Ik2a, Ik3 and Ikx) and the dominant negative isoforms (Ik4, Ik4a, Ik6, Ik7, Ik8 Ik9 and Ik10) that are generated by *IKZF1*. It has also been shown that alternate splicing of exons 4, 5 and 6 generated isoforms that had zero to four N - terminal zinc fingers which defines each isoform. An *IKZF1* isoform with 3 or 4 N - terminal zinc fingers is able to bind DNA sequences which contain TGGGA or AGGAA motifs (Iacobucci et al, 2008b) and carry out *IKZF1* function during lymphoid differentiation and development (Nakase et al, 2000 & Hosokawa et al, 2000).

There are five *IKZF1* isoforms (Ik1, Ik2, Ik2a, Ik3 and Ikx) that contain at least 3 N - terminal zinc fingers which allow DNA binding and subsequently carried out normal *IKZF1* function as a transcription factor. These isoforms are referred to as the functional isoforms of *IKZF1*. All the other *IKZF1* isoforms, Ik4-Ik10 are often referred to as the "dominant negative isoforms" since they only have 2 or no N - terminal zinc fingers and thus cannot carry out a normal

IKZF1 function. Instead these isoforms form homo and hetero dimers with the functional *IKZF1* isoforms via their C –terminal zinc fingers and inhibit *IKZF1* function (Sun et al, 1999a; Nishii et al, 2002; Mullighan et al, 2008 & Martinelli et al, 2009).

1.5 *IKZF1* in ALL patients

1.5.1 *IKZF1* isoform expression in ALL patients

The longer functional *IKZF1* isoforms containing 3 or 4 N-terminal zinc fingers are expressed in healthy cells. It has also been shown that Ik1 or Ik2 are generally present in healthy individuals while the dominant negative isoforms Ik4-Ik10 are predominantly expressed in Ph+ and relapsed ALL patients, in particular *IKZF1* isoform Ik6 (Sun et al, 1999a; Nishii et al 2002; Olivero et al, 2000; Mullighan et al; 2008, Iacobucci et al; 2009 & Kuiper et al, 2010).

1.5.2 *IKZF1* deletions in ALL patients

Subsequent to genome wide studies, various deletions involving different exons of *IKZF1* were detected. The scope of *IKZF1* deletions were categorized from four landmark studies that detected deletions in a total cohort of 159 ALL (mostly Ph+) patients (Mullighan et al, 2008; Martinelli et al, 2009; Kuiper et al, 2010 & Dupuis et al, 2012). Different combinations of *IKZF1* exons were deleted at different frequencies (figure 1.3).

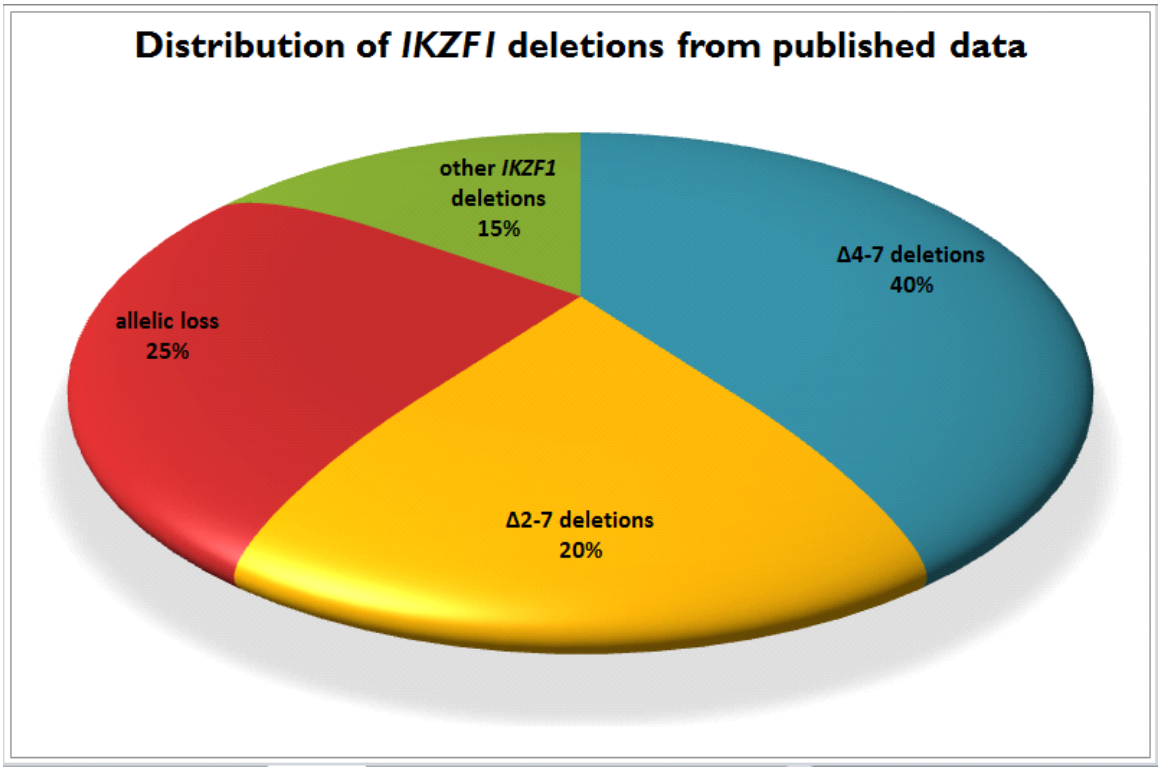

Figure 1.3. Distribution of *IKZF1* deletions in B-ALL cases (children and adult)
The percentages of each deletion was gathered from data taken from Mullighan et al, 2008; Martinelli et al, 2009; Kuiper et al, 2010 & Dupuis et al, 2012. "Δ4-7" involves a simultaneous deletion of exons 4,5,6 and 7,"Δ2-7"involves a simultaneous deletion of exons 2,3,4,5,6 and 7.

The most common *IKZF1* deletion involves exons 4-7 which results in high levels of Ik6 protein. Ik6 dimerises with the functional Ikaros proteins in a dominant negative manner inhibiting Ikaros function (Olivero et al, 2000; Mullighan et al; 2008, Iacobucci et al; 2009 & Kuiper et al; 2010).The loss of exons 4-7 of *IKZF1* results in a juxtaposition of exon 3 next to exon 8 and the exclusive expression of isoform Ik6 which has been confirmed using RNA and reverse transcriptase - polymerase chain reaction (RT-PCR) (Mullighan et al, 2008 & Iacobucci et al, 2009). This deletion is the only deletion that correlates with *IKZF1* isoform expression. Any other type of*IKZF1* deletion has shown coexistence of both functional and dominant negative *IKZF1* isoforms after RT-PCR (Olivero et al, 2000; Mullighan et al; 2008, Iacobucci et al; 2009 & Kuiper et al, 2010).

IKZF1 deletions that involve exons 2-7 or an allelic loss of *IKZF1* lead to haploinsufficiency (Kastner et al, 2013) since most of the *IKZF1* exons are required for *IKZF1* function. Exons 3-5 of *IKZF1* encodes the N-terminal zinc fingers which carry out *IKZF1* function, loss of these exons result in overexpression of dominant negative isoforms. Exon 2 contains the

translational start site; loss of this exon prevents translation of *IKZF1*. Exon 8 is required for*IKZF1* dimerization (Iacobucci et al, 2008a & Martinelli et al, 2009). Hence, deletions involving any of these exons are equivalent to an effect of a whole gene deletion of *IKZF1* (Iacobucci et al, 2009; Dupuis et al, 2012 & Caye et al, 2013).

1.5.3 *IKZF1* deletions in ALL patients that relapse

Two studies carried out by the Dutch leukaemia group (Kuiper et al, 2010) and the Children's Oncology Group in the USA (Mullighan et al, 2009) have shown that 28.6% (n=221) and 53% (n=131) of paediatric ALL patients who relapsed contained an *IKZF1* deletion respectively. Furthermore, *IKZF1* gene deletions were detected in low, medium and high risk ALL patients (Mullighan et al, 2009; Iacobucci et al, 2008a&Martinelli et al, 2009).

It has since been shown that *IKZF1* deletions together with signs of minimal residual disease (MRD) were able to predict up to 79% of ALL cases who relapsed (Waanders et al, 2010). Accumulated evidence has also shown that *IKZF1* deletions are maintained from presentation until relapse or it can be acquired at the time of relapse (Mullighan et al, 2011, Yang et al, 2008 & Kuiper et al, 2010). Subsequent studies have also shown that patients with *IKZF1* deletions respond poorly to therapy (Waanders et al, 2010).

1.5.4 *IKZF1* in ALL patients that contain *BCR/ABL1*

IKZF1 deletions have also been detected frequently in ALL patients that have the Ph chromosome and in CML patients in blast phase. The presence of the Ph chromosome, often referred to as *BCR/ABL1*, is known to be associated with high risk ALL patients (adults and in children) with a poor prognosis. Treatment failure is also common in Ph+ ALL patients (Mullighan et al, 2008).

1.5.4.1 Biology of *BCR/ABL1*

The Ph chromosome results from the fusion of the *ABL* gene on chromosome 9 and the *BCR* gene on chromosome 22 forming the *BCR/ABL1* fusion gene (Nowell& Hungerford, 1960). It is a hallmark of chronic myeloid leukaemia (CML) patients. In adult ALL, the Ph chromosome is observed in 33% of patients and in children it is observed in 3% (Pui & Jeha, 2007). There are three possible breakpoints in the *BCR* gene when it combines with the second or third

exon of the *ABL* gene to form the Ph chromosome (Kurzrock et al, 2003). The two most common forms are referred to as the major breakpoint cluster region (M-bcr) breakpoint and the minor breakpoint cluster region (m-bcr) breakpoint. The M-bcr breakpoint is the hallmark of CML whilst the m-bcr is found in the more aggressive ALL cases (Kurzrock et al, 2003). M-bcr breakpoint gives rise to the p210 protien product and the m-bcr breakpoint gives rise to the p190 protein product. CML patients who have a Ph chromosome with the p210 protein product respond well to Gleevec© (Novartis) therapy and usually achieve a normal blood count (Druker & Lydon, 2000). However, paediatric and adult ALL patients who have a Ph chromosome resulting from a p190 protein product do not respond as well to Gleevec© therapy and still experience relapse of the disease (Hunger et al, 2011).

Mullighan et al (2008) published a landmark paper that detected *IKZF1* deletions using DNA microarrays in a cohort of adult and paediatric ALL patients. In their study it was found that *IKZF1* was deleted in 83.7% of *de novo* Ph+ ALL patients and was an acquired lesion at the time of transformation of CML to ALL (lymphoid blast crisis) in adult leukemic patients. This study consequently lead to numerous additional investigations of *IKZF1* gene alterations in ALL patients and have confirmed the same findings, shown in table 1.1.

Table 1.1 B-cell ALL patients (Ph+ and Ph-) that were tested for *IKZF1* deletions

Research group	ALL patients(paediatric or adult),cohort size and Ph status	% of *IKZF1* deletions
Mullighan et al, 2008	43 Ph+ paediatric ALL	84%
Mullighan et al, 2009	221 high risk Ph- paediatric ALL	28.6%
Iacobucci et al, 2009	106 Ph+ ALL adult and CML blast crisis (69)	75%
Kuiper et al, 2010	34 paired presentation and relapse ALL	38%
	131 Ph+ and Ph- paediatric ALL	14%
Martinelli et al, 2009	83 adult Ph+ ALL	63%
Harvey et al, 2010	207 high risk Ph- paediatric ALL	30%
Mi et al, 2012	582 Ph+ and Ph- paediatric and adult ALL	21%
Caye et al, 2013	60 Ph+ paediatric ALL	75%
	512 Ph- paediatric ALL	16%
Dupuis et al, 2012	46 Ph+ and Ph- adult	52%
	93 Ph+ and Ph- paediatric	27%

From table 1.1, if we only look at the Ph+ ALL patients (adult and paediatric), collectively *IKZF1* deletions were detected in 78% (n=292) patients. Generally Ph+ ALL patients are considered as high risk patients, are poor responders to therapy, often experience relapse and die of the disease (Hunger et al, 2011). The high frequency of *IKZF1* deletions detected in Ph+ ALL patients from the studies shown in table 1 confirmed *IKZF1* deletions as the hallmark in these patients. *IKZF1* deletions have also been detected in chronic myeloid leukaemia (CML) patients who progress to lymphoid blast crisis (Mullighan et al, 2008 & Iacobuuci et al, 2009).

1.6 *IKZF1* as a prognostic marker

Over the past 3 years *IKZF1* gene deletion has therefore become a marker strongly associated with Ph+ ALL patients and high risk ALL patients that relapse (Ph+ and Ph-).Since *IKZF1* deletions are preserved from presentation until relapse, researchers and pathologists have suggested that all newly diagnosed ALL patients should be screened for *IKZF1* deletions. It has also been suggested that *IKZF1*deletions should be considered when

stratifying ALL patients into a particular risk group. Early identification of high risk ALL patients would allow doctors to give more intensive therapy to these patients and increase their survival (Mullighan et al, 2008 & Martinelli et al, 2009).

The above mentioned studies used genome wide SNP array analysis (DNA microarrays) to detect focal deletions of *IKZF1*. Results were validated using Fluorescent *in situ* hybridization (FISH) (Mullighan et al, 2008 & Martinelli et al, 2009). Reverse Transcriptase PCR (RT-PCR) was also used to detect the expression pattern of the various isoforms produced by *IKZF1* using RNA specimens (Mullighan et al, 2008 &Iacobucci et al, 2009).

In order to introduce a diagnostic test for *IKZF1* analysis several options must be considered: FISH and DNA microarrays are a big financial investment. Furthermore, DNA microarray data can be technically challenging to analyse. Multiplex ligation probe dependent amplification (MLPA) is a method that can be used to detect focal deletions of up to 50 targets simultaneously in a single reaction using a DNA specimen. MLPA can detect small copy number changes of targets in all exons of specific genes such as *IKZF1*. MLPA is not as technically challenging to analyse as DNA microarrays and has also been used successfully in the diagnostic setting to detect copy number aberrations of several genes (Jeuken et al, 2009; Murugan et al, 2010; Lake, 2011 & Stuppia et al, 2012). MLPA was used to confirm results obtained from DNA microarrays (Jennings et al, 2011; Combaret et al, 2012 & Hills et al, 2010) and karyotyping (Vorstman et al, 2006 & Boormans, 2010) in other studies. It has also been suggested that MLPA should be used to detect copy number alterations in developing countries because it is cheaper than other techniques such as DNA microarrays. Furthermore MLPA has been successfully used in recent studies to detect *IKZF1* deletions (Waanders et al, 2010 & Moorman et al, 2012).

MLPA would be an ideal method to test for *IKZF1* deletions and the prevalence of *IKZF1* gene aberrations has not been determined in South African ALL patients. An *IKZF1* deletion could be a critical factor to consider when stratifying ALL patients into a risk group.

1.7 STUDY OBJECTIVES

Main objective: To assess diagnostic test methods for *IKZF1* deletions and make a recommendation. To determine the prevalence of *IKZF1* gene alterations in a cohort of South African ALL patients (adults and minors) using DNA and RNA obtained from a blood or bone marrow diagnostic specimens received at the Somatic Cell genetics unit.

Specific Objectives:

To use Reverse Transcriptase- PCR (RT-PCR) to analyse RNA and the expression of the various *IKZF1* isoforms.

To use MLPA to interrogate *IKZF1* exon copy number.

To use FISH on the DNA specimens to confirm the MLPA results, to use Sanger sequencing on RNA (cDNA) to validate the RT-PCR results.

To compare *IKZF1* copy number to *IKZF1* isoform expression.

To compare MLPA results to cytogenetics findings after analysis.

Chapter 2. Methodology and Subjects

2.1 Overview of specimens and techniques used

Newly collected and retrospective samples were used for this study. RNA specimen was used for RT-PCR of *IKZF1* and the amplified PCR products were confirmed with Sanger sequencing. MLPA was used to detect *IKZF1* with probes specific for all 8 exons of *IKZF1* and copy number changes were confirmed with FISH.

2.1.1 Prospective samples

Prospective specimen were obtained from the paediatric unit (ages 0 -16) at the Chris Hani Baragwanath hospital after informed patient/guardian consent (Appendix A). A pathologist or qualified phlebotomist on site took the specimen and stored it in a Paxgene Blood RNA tube (Qiagen) at -20°C. Prospective ALL specimens (cases 1-13) were aliquoted into 2 tubes, one tube was used for DNA extraction and the other tube was used for RNA extraction. Successful extraction of both nucleic acids from the same specimen using a Paxgene tube was obtained previously (Kruhøffer et al, 2007). Leukemic patients have an excess number of lymphoblasts or lymphocytes in the bone marrow and peripheral blood. Hence, both types of specimens were used for *IKZF1* analysis. DNA was used for MLPA analysis while RNA was used for RT-PCR.

2.1.2 Retrospective DNA samples for MLPA

There were 28 retrospective DNA patient samples (cases 14 -41) that were obtained from the diagnostic division of the laboratory (Somatic cell genetics unit). These specimens previously had chromosomal analysis in the diagnostic division of the laboratory. The remainder of the cell culture pellet (~500ul) was stored in methanol and acetic acid at -4°C. DNA was extracted from these samples using the Dneasy kit (Qiagen®) for MLPA analysis. Together with the retrospective samples, there were a total of 41 leukemic samples used for MLPA analysis. The majority of the cohort included ALL patients (32) with 23 paediatric ALL cases and 9 adult cases of which 15 were Ph+. There were also 5 AML and 5 CML cases used for MLPA analysis in this study.

2.1.3 Retrospective RNA samples for RT-PCR

There were 11 retrospective ALL patient samples (case 43-53) that were obtained from the diagnostic division of the laboratory. RNA from these samples was previously extracted for diagnostic purposes and included seven Philadelphia positive (Ph+) cases and four Philadelphia negative (Ph-) cases. Together with the 13 prospective ALL samples, a total of 24 ALL samples were used for RT-PCR.

2.1.4 Control samples

There were 10 prospective control samples (A-J) taken from healthy individuals. Control samples used in this study were obtained from South African black (7) and Caucasian (3) individuals. DNA and RNA were extracted from these samples using the Dneasy kit (Qiagen) and Paxgene (Qiagen®) extraction methods respectively. RT-PCR and MLPA analysis of *IKZF1* was performed on both types of specimen. A summary of the total number of samples used for each test is shown in table 2.1.

Table 2.1 Summary of control and patient samples used in this study

	Patient samples	Control samples	Total
MLPA analysis using DNA	41 (cases 1-41)	10 (A-J)	51
RNA analysis using RT-PCR	24 (cases 1-13, 43-53)	10 (A-J)	34

The results detected by MLPA were correlated with karyotyping data available for most cases that were tested with MLPA in this study. Patient information for these cases are shown in table 2.2 which includes the patient's diagnosis, age, gender and their Ph status. In some cases the stage of the disease was also known.

The patients enrolled in this study were diagnosed on bone marrow morphology, immunophenotyping and karyotyping. A breakdown of all samples analysed for each test is shown in table 2.3.

Table 2.2 Patient information of cases used for MLPA analysis

Case Number	Diagnosis	Gender	Age	Ph Status
1	ALL	M	12	Ph-
2	ALL	M	5	Ph-
3	ALL	M	11	Ph-
4	ALL Presentation	M	6	Ph-
5	ALL Presentation	M	3	Ph+
6	ALL Presentation	M	2	Ph-
7	ALL	M	8	Ph-
8	ALL	M	7	Ph+
9	ALL	F	10	Ph+
10	ALL	F	12	Ph+
11	AML	F	7	Ph-
12	AML	M	2	Ph-
13	ALL	F	3	Ph-
14	ALL	M	69	Ph+
15	ALL	M	18	Ph+
16	ALL Relapse	M	14	Ph-
17	ALL	M	5	Ph+
18	CML	F	40	Ph+
19	ALL Relapse	M	9	Ph-
20	AML Relapse	F	50	Ph-
21	ALL	M	3	Ph-
22	CML Presentation	M	13	Ph+
23	ALL Presentation	F	57	Ph+
24	ALL	M	16	Ph+
25	ALL	F	28	Ph+
26	ALL	F	7	Ph+
27	CML	M	74	Ph+
28	ALL	F	58	Ph+
29	ALL	M	36	Ph+
30	AML	M	28	Ph+
31	ALL	M	3	Ph+
32	ALL	M	30	Ph+
33	ALL	M	16	Ph-
34	AML	M	55	Ph+
35	ALL	F	20	Ph-
36	ALL Relapse	M	16	Ph-
37	ALL	M	5	Ph-
38	CML	F	26	Ph+
39	CML	M	44	Ph+
40	ALL	F	5	Ph-
41	ALL	M	4	Ph-

Table 2.3 Summary of patients tested for either MLPA or RT-PCR if DNA or RNA was available

Case Number	DNA –MLPA	RNA – RTPCR	Case Number	DNA –MLPA	RNA – RTPCR
1	√	√	27	√	
2	√	√	28	√	
3	√	√	29	√	
4	√	√	30	√	
5	√	√	31	√	
6	√	√	32	√	
7	√	√	33	√	
8	√	√	34	√	
9	√	√	35	√	
10	√	√	36	√	
11	√	√	37	√	
12	√	√	38	√	
13	√	√	39	√	
14	√		40	√	
15	√		41	√	
16	√		43		√
17	√		44		√
18	√		45		√
19	√		46		√
20	√		47		√
21	√		48		√
22	√		49		√
23	√		50		√
24	√		51		√
25	√		52		√
26	√		53		√

√ - sample was tested for either MLPA and/or RT-PCR.

The number of prospective samples used for this project was limited by the number of ALL patients who presented at the Chris Hani Baragwanath hospital over the 18 month period of the study. Other studies that had investigated *IKZF1* deletions in Ph+ ALL patients had a similar incidence of 10 – 15 Ph+ ALL patients (Mullighan et al, 2008 & Iacobucci et al, 2008b) over a one year period.

One sample (case p) used in this study (not listed in table 2.2 or 2.3) was obtained from an ALL patient who had a known deletion of exons 4-7 of *IKZF1* detected with DNA microarray analysis (SNP6) in an independent laboratory. There was both DNA and RNA specimen available for case p, thus, this case was used as a positive control for both RT-PCR and MLPA analysis.

2.2. DNA and RNA isolation

Approximately 5ml of patient specimen (peripheral blood or bone marrow) was collected by a doctor or phlebotomist in a Paxgene Blood RNA tube (Qiagen®) (cases 1-13). Part of the specimen was aliquoted into a separate eppendorf tube (300ul) in order to extract DNA using the DNeasy blood kit (Qiagen) (Appendix B1). RNA was extracted from the remaining specimen (4.5ml) using the Paxgene kit as per package insert (Appendix B2).

2.2.1 DNA and RNA evaluation

DNA and RNA quantity and quality was determined using the nanodrop 1000 spectrophotometer (Thermo Scientific®). A minimal yield of 20ng (5ng/ul) of DNA and RNA was required to proceed with MLPA and RT-PCR. The integrity of DNA and RNA was further evaluated on an agarose gel with Gel Red (Biotium®) (Appendix C1).

2.3 PCR

PCR, developed in 1983 (Saiki et al, 1985) is an *in vitro* molecular technique used to exponentially amplify a fragment of DNA within a heterogenous collection of DNA sequences. A PCR reaction required specific reagents and a thermocycler to amplify a specific segment of DNA. The reagents that were used in PCR include: *Thermusaquaticus* (Taq) polymerase, deoxynucleotide triphosphates (dNTPs), magnesium chloride (MgCl2),

buffer, primers and molecular grade water. A pair of oligonucleotide primers (18-25 base pairs long) flanking the region of interest was used to amplify the specific DNA sequence of interest.

2.3.1 Reverse transcription

Prior to PCR, total RNA extracted from the blood or bone marrow specimen was reverse transcribed into complementary DNA (cDNA) (1ug) using specific thermocycling conditions (Appendix D1). All reagents used to synthesize cDNA were supplied by the Scientific group.

The cDNA was then PCR amplified for the housekeeping gene involved in glycolysis, Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*). This was done in order to confirm that RNA was not degraded. The expression of *GAPDH* is ubiquitous and it is often used to normalise quantitative real time PCR data (Barber et al, 2005). The reagents and thermocycling conditions used for amplification can be seen in Appendix D2. PCR of *IKZF1* isoforms was then carried out.

2.3.2 Conventional PCR of *IKZF1* isoforms

Conventional PCR was used to amplify all possible *IKZF1* isoforms (Ik). The length of each *IKZF1* isoform ranges from 255 (Ik6) to 945 (Ik1) base pairs (Iacobucci et al, 2008b). The sizes of each isoform were verified using the BLAST tool in NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi#409253393 – date accessed 10/5/2013). Each *IKZF1* isoform contains exons 2 and 8. A two-step nested PCR was performed to amplify isoforms of *IKZF1*. The primer sequences used for both rounds of PCR are shown in table 2.4.

Table 2.4 PCR primers used for amplification of *IKZF1*

Nested PCR step	Primer sequence	Reference
1st round	Ik1 : 5_CACATAACCTGAGGACCATG-3 Ik2: 5_-AGGGCTTTAGCTCATGTGGA-3	Iacobucci et al, 2008b
2nd round	Ik3: 5_-ATGGATGCTGATGAGGGTCAAGAC-3 Ik4: 5_-GATGGCTTGGTCCATCACGTGG-3	Iacobucci et al, 2008b

Thus, a forward primer specific for exon 2 and a reverse primer specific for exon 8 were used for PCR which ensured that all possible *IKZF1* isoforms were amplified. Figure 2.1 illustrates the positions of the primers used in the nested PCR reaction to amplify *IKZF1*.

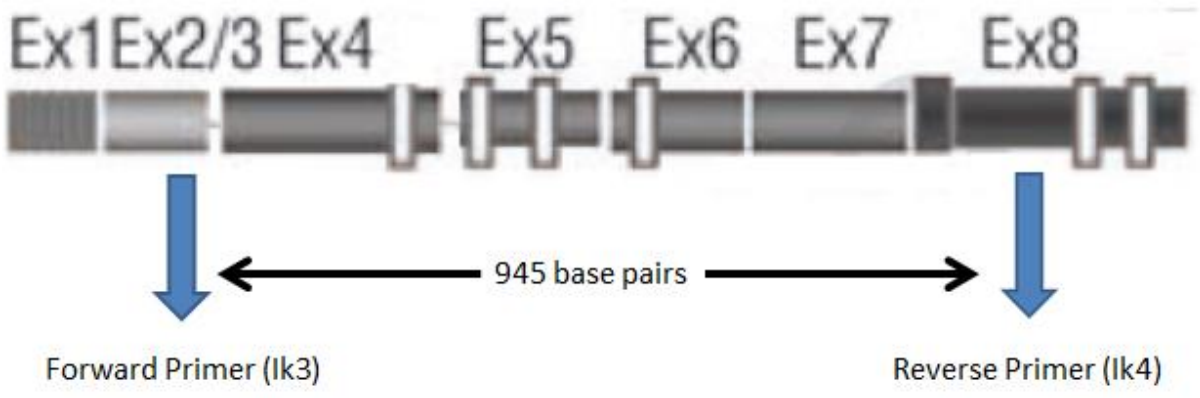

Figure 2.1 A representation of the full *IKZF1* transcript (isoform 1) including all exons
The blue arrows indicate the binding site of the forward and reverse primers in exons 2 and 8 respectively, enabling all *IKZF1* isoforms to be amplified. Ex: exon. White bars: position of zinc fingers in *IKZF1*.

The PCR cycling conditions and reagent concentrations used to amplify *IKZF1* isoforms are shown in table 2.5.

Table 2.5 PCR reaction parameters used for-PCR of *IKZF1*

PCR reagents	*IKZF1* 1st round	*IKZF1* 2nd round
†Buffer (X10)	1X	1X
DNTPs	2.5 mM	2.5 mM
MgCl2	25 mM	25 mM
Primers (Forward and Reverse)	25 pmol	25 pmol
DNA	1ug	0.5ul of first round PCR product
ᴪTaq Polymerase	1 unit	1 unit
Thermocycling conditions	95°C – 5 minutes	95°C – 5 minutes
	95°C – 30 seconds	95°C – 30 seconds
	57°C – 30 seconds 35 cycles	62°C – 30 seconds 35 cycles
	72°C – 90 seconds	72°C – 30 seconds
	72°C – 7 minutes 4°C Hold	72°C – 7 minutes 4°C Hold

†Made up of 500mM KCl, 100mM Tris-HCl (pH 8.3) and 25mM MgCl2

ᴪ (Bioline)

2.3.3 Verification of PCR amplicons

After *IKZF1* PCR, amplicons were separated by gel electrophoresis (Appendix C1). The size of each *IKZF1* isoform was compared with a 100 base pair ladder (Thermo scientific).

2.3.4 Purification of PCR products

PCR of *IKZF1* produced up to 10 bands on an agarose gel. Thus, some of the bands were excised from the gel, purified (Hangzou Bioer Technology®) and used as starting material for the DNA sequencing process to verify *IKZF1* isoforms (Appendix B3). A concentration of 30ng of DNA (nanodrop 1000) was used for DNA sequencing for each sample.

2.3.5 DNA sequencing

Samples (control and patient) were sequenced using the Sanger sequencing chain termination chemistry (Sanger et al, 1977) in order to verify that the banding patterns observed after *IKZF1* PCR were correct.

2.3.6 Cycle Sequencing

Cycle sequencing was carried out in a thermocycler using a series of fluctuating temperatures in order to amplify and fluorescently label DNA. Both strands (forward and reverse) were sequenced in order to validate each other. The forward and reverse primers used for cycle sequencing were the same used for the second round of nested PCR of *IKZF1* as described in table 2.4.

The thermocycling conditions and reagents that were used for the cycle sequencing reaction are shown in table 2.6.

Table 2.6. The protocol used for cycle sequencing of *IKZF1*

PCR component	Final concentration (1 reaction)	Volume (1 reaction)
Purified PCR product	30ng	3µl
Primer (forward (F2) or reverse (R2))	3.2pmol	6.4µl
5X sequencing Big dye Buffer	1.15X	4.6µl
ddH2O	N/A	4µl
BigDye3.1	1ul/250 base pairs	2µl
Total		20µl
Thermocycling conditions	96°C – 1 minute 96°C – 10 seconds 50°C – 5 seconds 60°C – 4 minutes 25 cycles 4°C – Hold	

2.3.7 Purification of Cycle Sequencing products

A second purification step was carried out in order to eliminate unincorporated ddNTPs, enzyme and buffer before sequencing the products. In a 96 well plate (Applied Biosystems®), 20ul of cycle sequencing product was mixed with 80µl of 75% isopropanol (Merck®) and left to stand in the dark for 30 minute. The plate was then centrifuged (Eppendorf®) at 2000 relative centrifugal force (rcf) for 45 minutes then gently flipped over onto absorbent paper. The plate was then spun for 1 minute at 700rcf to remove the

isopropanol and unwanted ddNTPs and then left to air dry for five minutes to remove residual isoproponol. Finally the DNA pellet was resuspended with 10µl of Highly deionised (Hi-Di) formamide (Applied Biosystems), an injection solvent for the DNA pellet and was sequenced using the ABI 3500 genetic analyser (Applied Biosystems).

2.4 Multiplex ligation probe dependant amplification (MLPA)

Multiplex ligation probe dependant amplification (MLPA) is a PCR based technique. MLPA provides DNA copy number information for a target region in a quantitative way. The kit for *IKZF1* (SALSA P335-B1ALL-IKZF1 probemix) contains a mix of 57 probes to assess copy number of different genes simultaneously.There were 8 probes specific for all 8 exons of *IKZF1* and 13 probes for reference genes known to have a relatively stable copy number in ALL patients. The rest of the probes targeted genes known to be aberrantly expressed in ALL patients, these included: *PAX5* (6 probes), *ETV6* (6 probes), *RB1* (5 probes), *BTG1* (4 probes), *EBF1* (4 probes), *CDKN2A*-*CDKN2B* (3 probes). There were also probes specific for exons within the pseudoautosomal region (PAR1) region on Xp22.33/Yp11.32: *IL3RA, SHOX area*, *CRLF2*, *CSF2RA, P2RY8*, *ZFY* and *IL3RA* genes (Appendix E).

MLPA uses a thermocycler with a heated lid and the amplified products are resolved by capillary electrophoresis on a sequence analyser. MLPA data was analysed by using macros enabled software that creates ratios (dosage quotient values) and uses statistics to quantify MLPA targets.

2.4.1 Principle of MLPA

The MLPA mix contains several probes that are target-specific for exons of certain genes of interest. Each probe is made up of two oligonucleotides that hybridise to a DNA sequence of interest through complementary base pairing. Both oligonucleotides contain a universal primer sequence. One of the oligonucleotides is longer because it contains a stuffer sequence (M13-derived oligonucleotide) that varies in length for each probe. The stuffer sequence allows each probe to have a unique length which allows them to separate according to their unique size (Schouten *et al*, 2002). An illustration of an MLPA probe and its components is shown in Figure 2.2.

Figure 2.2 MLPA probe design.
Both oligonucleotides contain a universal sequence at the 5' (sequence Y) and the 3' end. Primer sequence: X. Stuffer sequence: shown in green allows each probe to have a unique length. Blue: hybridization sequence. Illustration taken and adapted from Schouten et al, 2002.

The two oligonucleotides hybridize to the target DNA, then ligation completes the MLPA probe if the target DNA sequence is present in the sample that is being tested (Figure 2.3).

Figure 2.3 Ligation step
MLPA probe attaches to target DNA sequence through complementary base pairing; followed by a ligation step which completes the MLPA probe). Black: universal primer sequence. Blue: hybridization sequence. Red: target DNA. Light blue: hybridization between hybridization sequence and target DNA. Green: stuffer sequence. Illustration taken from Schouten *et al*, 2002.

Each ligation product has a length of between 50-70 nucleotides. Ligation products are then amplified by PCR. All 57 probes have a set of identical universal primers on their 5' (Y) and 3' (X) ends enabling all ligation products to be amplified in a single reaction using one pair of universal primers (Figure 2.4). All probes differ in length by 5-10 nucleotides.

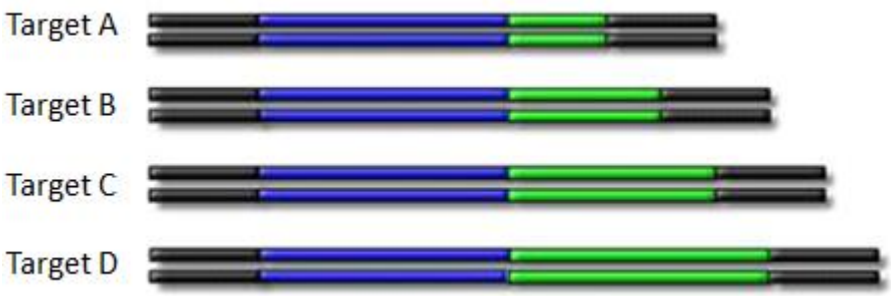

Figure 2.4 PCR of MLPA probes
Each probe has a unique length which makes it possible to separate them by size. Illustration taken and adapted from Schouten *et al*, 2002. Black: universal primer sequence. Blue: hybridization sequence. Green: stuffer sequence.

Figure 2.4 shows the different sizes of each MLPA fragment. The stuffer sequence enables each fragment to have a unique size which can be separated on the 3500 genetic analyser (Applied Biosystems) by capillary electrophoresis through the liquid polymer POP7 (Applied Biosystems).

The 5' end universal primer on each MLPA fragment is fluorescently labelled with a fluorophore (6-FAM dye set) which fluoresces as a blue colour when it excites the laser during capillary electrophoresis. This allows it to be detected during capillary electrophoresis.

2.4.2 Sizing of each MLPA fragment

A size standard, ROX500 (Applied Biosystems), was resolved simultaneously with each MLPA PCR product. ROX500 was made up of 16 fragments of known length ranging from 35 - 500 bases in length and was used to determine the size of each MLPA fragment. ROX500 was labelled with the TAMRA dye set and fluoresced as a red colour as it passed through the laser during capillary electrophoresis. This enabled it to be distinguished from the blue colour of the MLPA probes.

2.4.3 MLPA reaction parameters

A thermocycler (Applied Biosystems) with a heated lid was used for all MLPA reactions. MLPA were obtained after 24 hours. DNA (150ng) was used as starting material for each MLPA reaction and the MLPA parameters and cycling conditions that were used are shown in table 2.7 (Insert MLPA® General Protocol, MRC Holland).

Table 2.7 MLPA protocol used to amplify all 57 probes in a multiplex PCR reaction

Day 1: Hybridization reaction		
MLPA reagent	**MLPA volume (1X)**	**Thermal cycling conditions**
Sample DNA (150ng)	5µl	95°C – 1 minute 60°C – 20 hours 4°C Hold
MLPA buffer	1.5µl	
MLPA probemix	1.5µl	
Day 2: Ligation		
MLPA reagent	**MLPA volume (1X)**	**Thermal cycling conditions**
Hybridization reaction	8µl	54°C – 15 minutes 98°C – 5 minutes 4°C Hold
Ligase-35 buffer A and B	6µl	
Ligase-65	1µl	
ddH2O	25µl	
Day 2: MLPA PCR reaction		
MLPA reagent	**MLPA volume (1X)**	**Thermal cycling conditions**
Ligation product	10µl	95°C – 30 seconds 60°C – 30 seconds 72°C – 60 seconds 35 cycles 72°C – 20 minutes 15°C Hold
SALSA PCR-primers	2µl	
SALSA PCR Polymerase	0.5µl	
SALSA PCR buffer	4µl	
ddH2O	26µl	

After the MLPA PCR reaction 0.7µl of MLPA PCR product was mixed with 9µl of Hi-Di and 0.3µl of ROX500 size standard (Applied Biosystems) and resolved on the ABI 3500 genetic analyser (Applied Biosystems).

2.4.4 Quality control of MLPA

All of the steps during the MLPA reaction mentioned above were validated using two sets of control fragments throughout the MLPA procedure: Q and D fragments. In addition to

validating each MLPA step, the MLPA peak profile for each sample was also analysed to ensure that the MLPA reaction was successful.

2.4.4.1 Quality control probes and guidelines for successful MLPA run

MLPA generated Q and D fragments during denaturation, ligation and PCR of the MLPA procedure. These control peaks indicated whether or not the quantity of DNA used as starting material was sufficient and whether or not the denaturation and ligation steps were successful during the MLPA reaction.

The Q-fragments consisted of a set of four fragments that were generated after the MLPA reaction: 64, 70, 76 and 82 nucleotides in length. The Q-fragments indicated if there was adequate DNA used in the reaction and whether or not the ligation step worked. If the peaks heights and areas of the Q-fragments were greater than the rest of the MLPA probes, it meant that there was insufficient sample DNA or the ligation step during MLPA was unsuccessful (Insert MLPA General Protocol, MRC Holland). This type of sample was rejected and repeated.

The denaturation fragments (D-fragments) was another set of internal quality control probes that were generated during the MLPA reaction. The D-fragments targeted sample DNA of 88 and 96 nucleotides in length and validated the denaturation step in the MLPA reaction. The D-fragments are sequences that have strong CpG content, known to be difficult to denature because of their strong hydrogen bond content. The two probes specific for the D-fragments anneal to sequences with strong CpG content. If the D-fragments peak height and areas were <40% compared to the rest of the MLPA probes it meant that denaturation was not successful in the MLPA reaction. Incomplete denaturation does not give reliable results when analysing MLPA results (Insert MLPA General Protocol, MRC Holland). These were considered as failed reactions and the results were rejected and the MLPA run was repeated.

After validating the D and Q-fragments, each sample's MLPA peak profile was analysed further using Genemapper 4.1 software (Applied Biosystems).

2.4.4.2 Filtering of MLPA data using Genemapper 4.1

Raw data from the ABI 3500 data collection software was filtered in the Genemapper 4.1 software (Applied Biosystems). The morphology of each MLPA peak profile was analysed. It was important to ensure that there were no split peaks, unusual broad peaks or spikes present in each sample. The non-template control (NTC) was a sample in which water was used instead of DNA as starting material. The NTC ensured that there was no contamination in the MLPA run. MLPA produced minute non-specific peaks next to the peaks specific to probes of the MLPA kit. These spurious peaks, referred to as background noise were removed before proceeding with MLPA analysis. Filtering of MLPA data using Genemapper 4.1 can be found in Appendix F1.

In addition to checking the MLPA peak profile for each sample, each MLPA run had 3-5 control samples (from healthy individuals) that was used for normalising each target-specific probe in the patient samples.

After filtering raw MLPA data, the quantitative data on the peak height of each probe was exported into a Microsoft Excel® 2010 document. MLPA peak heights, measured in relative fluorescent unit (RFU), had to range between 1000 – 25 000 RFUs. RFUs that fall outside of this range give unreliable MLPA results (personal communication – Applied Biosystems). If the RFU value did not fall in this range the sample was rejected and subsequently rerun.

Each MLPA fragment length was not absolute. There was a difference of +/- 3bps between the actual lengths of each MLPA probe and the length that was detected during capillary electrophoresis. Thus, the length that got detected during capillary electrophoresis was adjusted to match the exact length of each MLPA probe. This was set out manually in Microsoft Excel (Microsoft Excel® 2010). After these changes were manually adjusted, the data was then normalised using the Coffalyser® version 9.4 (http://coffalyser.software.informer.com/9.4/ [Date accessed 14/7/2013]) software to determine target copy number. Each MLPA run also included known positive and negative control samples: a sample with a monoallelic deletion of *IKZF1* and an ALL sample which did not have any *IKZF1* exon copy number changes. All of the above factors were used to validate each MLPA run.

2.5 MLPA analysis

There were a few Excel based software programs besides Coffalyser that were available to analyse MLPA data using DQ values. The *IKZF1* kit manufacturers MRC Holland developed a Coffalyser. Manchester University developed a program called Manchester analysis sheets and Softgenetics® developed a program called Genemarker which has to be purchased. The Manchester sheet analysis software was not used for this project because there was no template available for the P335-A4 MLPA kit used in this study. The other software package, Genemarker®, normalises MLPA data using only one unaffected healthy control sample. It is recommended that more than one healthy control sample should be used to normalise MLPA data (Coffa & van den Berg, 2011). Thus, it was decided to use Coffalyser 9.4 for MLPA analysis in this study.

2.5.1 Normalisation of MLPA data using Coffalyser 9.4

MLPA data was analysed using signals obtained from the peak heights (relative fluorescent units - RFUs) from each ligation product to calculate ratios (dosage quotient (DQ)) values. First the peak height signal of each target-specific probe was divided by the peak height signal of every reference probe in that same sample. This generated several ratios for each target-specific probe. Next, the median of all the produced ratios per target-specific probe was taken, this is known as the probe's Normalisation Constant. Subsequently, the Normalisation Constant of each target-specific probe in that sample was divided by the average Normalisation Constant of that probe in all the healthy reference samples (Insert SALSA MLPA probemix P335-B1 ALL-IKZF1, MRC Holland). This method also assumes that all of the reference probes in the healthy control samples are normal (2 copies) (Coffalyser, 2011). If MLPA target copy number was normal (2 copies), than the DQ value was close to 1. A duplication was called if the ratio was > 1.3 and a deletion was called if the ratio was <0.67. The sample was assumed to have a monoallelic deletion if the ratio lied between 0.54 – 0.67 (Coffa & van den Berg, 2011). These calculations were done in Coffalyser using the "Tumour analysis (LMS) "method. A step by step procedure to choose the correct commands in the Coffalyser software can be found in Appendix F2.

2.5.2 Validation of MLPA after using Fluorescent in situ hybridization (FISH)

FISH can be used for detecting chromosomal abnormalities such as gene fusions, aneuploidy or loss of a chromosomal region. FISH uses complementary base pairing of DNA sequences between a fluorescently labelled probe and a DNA specimen obtained from metaphase or interphase cells (Bishop, 2010). The FISH probe is made from a bacterial artificial chromosome (BAC) clone that contains a DNA sequence of a specific gene of interest. FISH probes (Abbott®) are fluorescently labelled with a specific fluorochrome which emits long wavelength energy in the form of visible light when viewed under a fluorescent microscope.

The *ETV6/RUNX1* (Runt-related transcription factor) dual fusion translocation probe was used in FISH to detect translocations involving the *ETV6* and *RUNX1* gene on chromosomes 12p13 and 21q22 respectively. *ETV6* is labelled with spectrum green and *RUNX1* is labelled with spectrum orange. The *ETV6* probe binds to *ETV6* on exons 3-5 and extends about 350 kb toward the telomere on chromosome 12. The *RUNX* gene is approximately 500kb in length and covered the entire *RUNX1* gene (Abbot Molecular Vysis Catalogue, 2011). The *ETV6/RUNX* dual fusion translocation probe was routinely tested in the diagnostic division of the laboratory. In a normal cell with 2 copies of *ETV6* and *RUNX1* a specific amount of signal indicates how many copies of each gene were present. The fluorescent signal pattern for a normal cell that had been hybridized with the LSI *ETV6/RUNX* ES Dual Colour Translocation probe was: two orange (*RUNX1*), two green (*ETV6*) (2O2G) signal pattern. Thus, if one copy of *ETV6* was deleted, there was only 1 green signal under the fluorescent microscope (Olympus®).

The MLPA kit could detect copy number of exons 1A, 1B, 3, 5 and 8 of the *ETV6* gene. Known patient samples V1, V2 and V3 that were detected by FISH to harbour a monoallelic deletion of *ETV6* were also tested using MLPA. Both tests used DNA from the same samples as starting material. In this way MLPA was validated using FISH. The entire FISH procedure was done on a bone marrow smear or peripheral blood smear on a glass slide (2.54cm X 7.62cm).

The FISH procedure had 3 steps: Denaturing, hybridization and washing. All 3 steps were carried out on DNA obtained from interphase cells that were fixed on a glass slide.

2.5.3 Denaturation and hybridization of double stranded DNA with fluorescent probe

Deionized formamide (Merck ®) was used to denature the hydrogen bonds between DNA to make it single stranded. Formamide reduces the melting temperature between the nucleic acids (McConaughy, 1969). Slides were immersed in deionized formamide at 76°C in a Thermobrite (Abbott®) for 5 minutes. During this 5 minute wait, the *ETV6/RUNX* dual fusion translocation FISH probe (in a closed eppendorf 1.5ml tube) was denatured in a heated water bath (Optolabor) at 76°C for 5 minutes. After denaturation of probe and sample, 10µl of probe was streaked onto a cover slip. The cover slip was then placed onto the slide with the sample DNA. The probe and sample DNA were then left overnight in the dark at 37°C in an incubator (Pro-lab®) and allowed to hybridize to each other.

2.5.4 Wash step

The FISH wash step was carried out the following day after hybridization. This procedure was done to in order to remove unbound probe which can cause background signals when viewing the slide under the fluorescent microscope.

The coverslips were removed and the slides were washed 3 times in formamide solution at 45°C for 10 minutes, followed by 5 minutes in saline-sodium citrate (SSC) (Sigma-Aldrich®) and finally 5 minutes in SSC and Tween (Merck®) (Appendix C2) at 45°C. The slides were then counter stained with one drop of DAPI (Merck®). DAPI is a counter stain that stained AT rich areas blue on chromosomes when viewed under UV light. One drop (~50µl) of Vectashield (Vector®) was placed on a new cover slip then placed onto the slide. Vector shield is an antifade solution, used to mount the coverslips and allow for visualization under the fluorescent microscope.

Chapter 3. Results

3.1. RNA extraction and yields

RNA extraction from peripheral and bone marrow specimens in Paxgene tubes yielded >50ng in all 13 samples. The spectrophotometer ratio of 260:280nm was > 1.8 which an indication of the purity of the RNA and that there was no protein contamination in the samples.

3.1.1 RT-PCR of *GAPDH*

RNA was reverse transcribed into cDNA and was successfully amplified for the *GAPDH* housekeeping gene in all 13 ALL samples, indicating that the RNA was not degraded (figure 3.1).

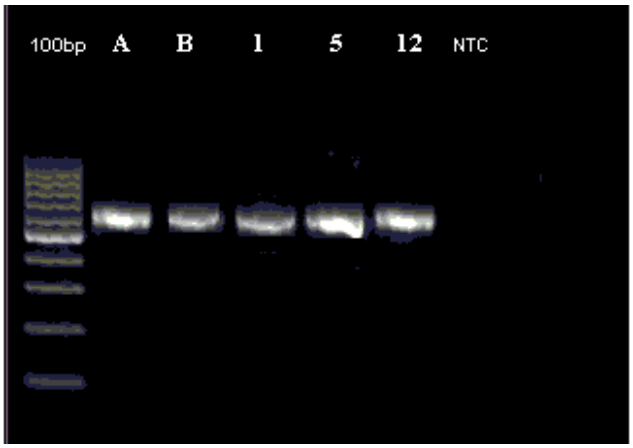

Figure 3.1. A representation of the banding patterns observed after PCR amplification of *GAPDH*
The extracted RNA was not degraded in all of the samples. Sample A and B are control samples from healthy individuals. Cases 1,5 and 12 are ALL patient samples. The NTC is a non-template control, the NTC does not contain any RNA, and it only contains the reagents used for the PCR. The absence of bands in the NTC lane ensured that there was no contamination in the PCR reaction.

3.2 RT-PCR of *IKZF1*

In addition to the 13 prospective ALL samples, there were also 11 retrospective ALL samples and 10 healthy control samples that were PCR amplified for *IKZF1* isoforms.

3.2.1 RT-PCR of *IKZF1* isoforms in normal cells

A nested PCR reaction using two sets of primers ensured that all possible isoforms of *IKZF1* were amplified in all patient and healthy control samples. The 10 healthy control samples (A-J) displayed a consistent pattern of bands. The functional isoforms: Ik1, Ik2, Ik3, and IkX, were present in all control samples, samples A and C are shown in figure 3.2.

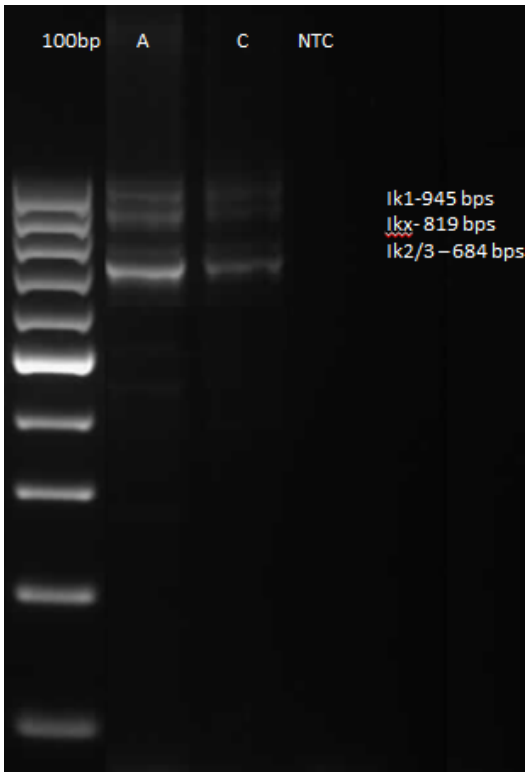

Figure 3.2. A representation of the banding patterns observed after PCR of *IKZF1* in 2 control samples (A and C) from healthy individuals
Ik2 and Ik3 are both 684 base pairs in length. Bps: base pairs. NTC: non-template control contains all PCR reagents without sample ensuring that there was no contamination present after PCR.

The majority of the control samples (8/10) had a banding pattern displayed in figure 3.2. However, control samples B and E also expressed dominant negative isoforms: Ik4A, Ik5, Ik6 and Ik8 in addition to the functional isoforms of *IKZF1*.

3.2.2 RT-PCR of *IKZF1* isoforms in ALL patients

The ALL samples showed a less consistent pattern of isoform expression when compared to the control samples. In ALL patient samples (Ph+ and Ph-), samples expressed additional

dominant negative *IKZF1* isoforms. The banding pattern of 2 ALL patients and control sample B are shown in figure 3.3.

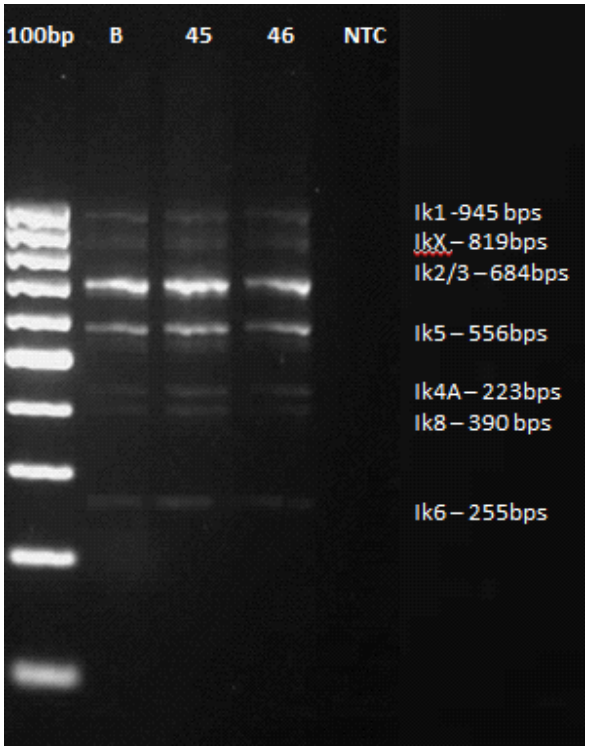

Figure 3.3 A representation of the banding patterns observed after PCR of IKZF1in ALL samples
Isoforms detected in 2 ALL samples and control sample B. Isoforms Ik1, Ik2, Ik3, Ik4A, Ik5, Ik6, Ik8 and IkX were present in these 3 samples. Cases 45 and 46 were Ph+ samples. NTC: non-template control contains all PCR reagents without sample ensuring that there was no contamination present after PCR.

Although most of the ALL patients expressed functional isoforms, there were three cases (48, 51 and p) that did not express any of the longer functional isoforms of *IKZF1*. These three cases only expressed one dominant negative isoform, Ik6 (figure 3.4).

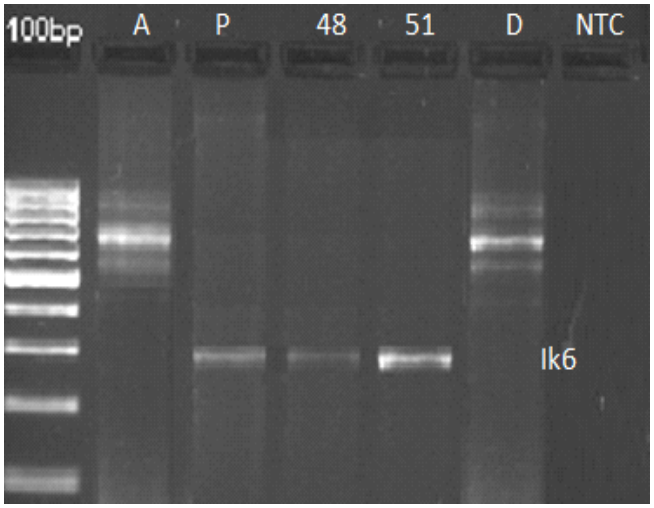

Figure 3.4 A representation of exclusive expression of Ik6 in cases p, 48 and 51
This band correlates with a size length of 255bps which corresponds to *IKZF1* isoform Ik6. Control A and control D were also run to ensure that the PCR was efficient. NTC (non-template control) shows that there was no contamination in the PCR reaction. NTC: non-template control contains all PCR reagents without sample ensuring that there was no contamination present after PCR.

The sole expression of Ik6 has been detected in All patients that have a genomic deletion of exons 4-7 of *IKZF1* (Mullighan et al, 2008 & Iacobucci et al, 2009). Sample P was a control sample which had a known deletion of exons 4-7 of *IKZF1*. A summary of all the *IKZF1* isoforms detected in all 24 ALL patients tested in this study are shown in table 3.1.

Table 3.1 *IKZF1* isoforms detected in 24 ALL patients by RT-PCR

There were 12 cases that were Ph- and 12 cases that were Ph+ for the m-bcr breakpoint of the *BCR/ABL1* fusion gene.

Case number	Ph Status (m-bcr)	*IKZF1* Isoforms detected
1	Ph-	Ik2, Ik3, Ik4, IkX
2	Ph+	Ik2, Ik3, Ik4, Ik6, IkX
3	Ph-	Ik2, Ik3, Ik4, Ik6, IkX
4	Ph-	Ik2, Ik3, Ik4, Ik6, IkX
5	Ph+	Ik1, Ik2, Ik3, Ik4, Ik4A, Ik5, Ik6, Ik7, Ik8, IkX
6	Ph-	Ik2, Ik3, Ik4, Ik6, IkX
7	Ph-	Ik2, Ik3, Ik4, Ik6, IkX
8	Ph+	Ik1, Ik2, Ik3, Ik4, Ik4A, Ik5, Ik6, Ik7, Ik8, IkX
9	Ph+	Ik2, Ik3, Ik4, Ik6, IkX
10	Ph+	Ik2, Ik3, Ik4, Ik6
11	Ph-	Ik2, Ik3, Ik4, Ik6
12	Ph-	Ik2, Ik3, Ik4, Ik6
13	Ph-	Ik2, Ik3, Ik4, Ik6
43	Ph+	Ik1, Ik2, Ik3, Ik4, Ik4A, IkX
44	Ph+	Ik1, Ik2, Ik3, IkX
45	Ph+	Ik2, Ik3, Ik4A, Ik6
46	Ph+	Ik2, Ik3, Ik4A
47	Ph+	Ik1, Ik3, Ik4, Ik6, IkX
48	Ph+	Ik6
49	Ph-	Ik1, Ik3, Ik4, IkX
50	Ph+	Ik1, Ik3, Ik4, Ik6, IkX
51	Ph-	Ik6
52	Ph-	Ik1, Ik2, Ik3, Ik4, Ik4A, Ik8, IkX
53	Ph-	Ik1, Ik2, Ik3, Ik4, IkX
Case P	Ph-	Ik6

Table 3.1 shows that the dominant negative isoforms (especially Ik6) was detected more frequently in ALL patient cases (Ph+ and Ph-) when compared to the isoforms present in healthy controls (A-J). It was also observed that 17/24 (71%) of ALL patients expressed the dominant negative isoform Ik6, in contrast to the healthy control samples in which only 20% (2/10) expressed Ik6. The functional *IKZF1* isoform Ik1 was present in 22/24 ALL cases and was expressed in all 10 healthy control samples. Ik1 contains exons 2-8 of *IKZF1* and has been shown to be present in normal healthy cells (Mullighan et al, 2008). Sanger sequencing of Ik1 was done in order to verify that this band correlates with Ik1. Cases 48, 51 and p did

not express any functional isoforms of *IKZF1* and only contained a single band which correlated with Ik6. This band was also sequenced to confirm the presence of Ik6.

3.2.3 DNA sequencing of *IKZF1* isoform Ik1 and Ik6

The PCR product from Ik1 from control sample A was extracted from an agarose gel and used as starting material for DNA sequencing. Initially gel extractions of Ik1 gave extremely low concentrations of 1-10 ng of DNA. For this reason it was attempted to increase the yield of the Ik1 PCR product. Thus, amplification of *IKZF1* from sample A was set up in quad druplicate in order to increase the amount of PCR product of Ik1. All four reactions were then combined and resolved on a 2% agarose gel and a PCR product that correlated with Ik1 (945bps) was excised and sequenced. Part of this DNA sequence for Ik1 is shown in figure 3.5.

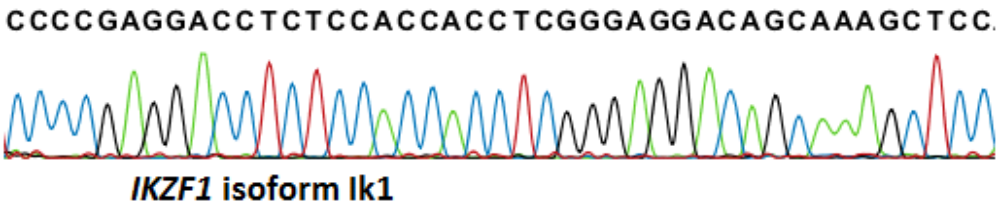

Figure 3.5 An electropherogram of a section of the DNA sequence of Ik1
A is shown in green, **C** in blue, **T** in red and **G** in black.

Control sample A was shown to align 100% with exons 2-8 of *IKZF1* and confirmed the 945bp product of Ik1. The sequence obtained was aligned to the *IKZF1* human reference sequence (NM_006060) obtained from Genbank (http://www.ncbi.nlm.nih.gov/nuccore/nm_006060 -[Date accessed 15/7/2013]). This reference sequence is the wild type sequence for the full *IKZF1* transcript (Ik1). A segment of Ik1 (exon 5) aligned to the reference sequence is shown in figure 3.6.

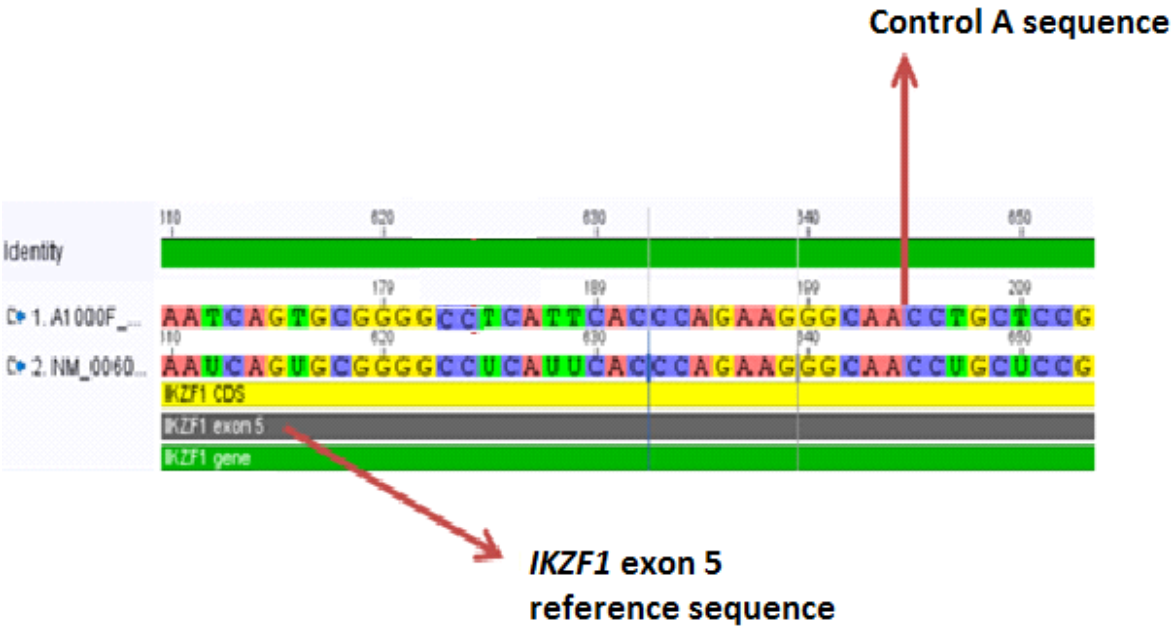

Figure 3.6 DNA sequence alignment of *IKZF1* isoform Ik1 of control sample A with the reference sequence (NM_006060)
Only a portion of exon 5 is shown in this figure. "A1000F" represents the forward sequence obtained from Control A; NM_00600 represents the reference sequence of *IKZF1* exon 5 (shown in grey). The solid green line indicates that the 2 sequences align perfectly with each other; if they did not align there would be gaps instead of a solid green line.

Figure 3.6 depicts a portion of exon 5 aligned with the reference sequence. Alignment of Ik1 with the reference sequence was done using the Geneious® software package (www.geneious.com – [Date accessed: 13/7/2013]). Exons 2-8 of Ik1 from control sample A aligned perfectly with the reference sequence of *IKZF1*. The entire DNA sequence of Ik1 for control sample A can be found in Appendix G.

The single band of 255 bps expressed in cases p, 48 and 51 was also sequenced to confirm the presence of Ik6. The DNA sequence obtained from the single band produced from case 48 is shown in figure 3.7.

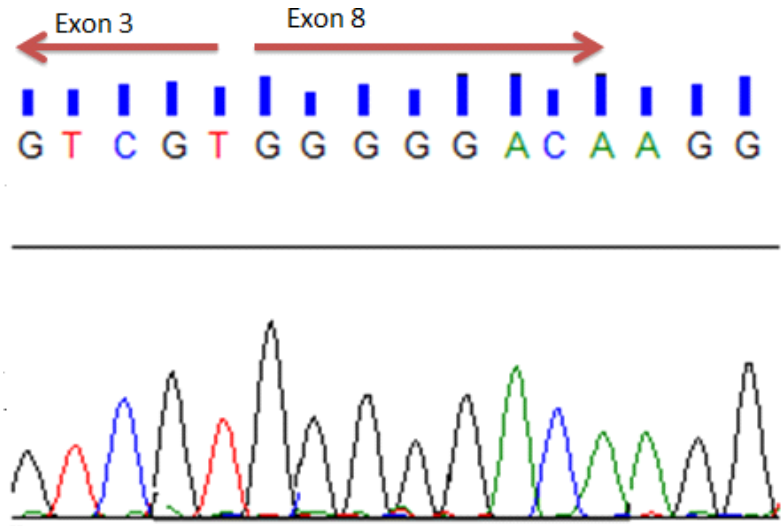

Figure 3.7 DNA sequencing of Ik6
Ik6 has exon 3 juxtaposed next to exon 8 of *IKZF1*. **A** is shown in green, **C** in blue, **T** in red and **G** in black.

Ik6 does not contain exons 4-7 which results from juxtaposition of exon 3 next to exon 8 of *IKZF1*. Cases p and 51 were also sequenced (not shown) and it was also confirmed that Ik6 was present in both samples. Case p is a known ALL patient who had a deletion of exons 4-7 of *IKZF1*, this deletion was also confirmed by MLPA in this study (shown later) and with DNA SNP6 microarrays in another laboratory.

3.4 DNA extraction and yields

DNA was successfully extracted using the DNeasy kit (Qiagen) from all 51 (patient and control) samples and were used for MLPA analysis. The yields obtained were sufficient for MLPA (ranging from 16µg – 160 µg). The DNA was also resolved on a 2 % agarose gel to check their integrity before MLPA analysis. An illustration of 3 control samples (A, B and C) and 3 ALL cases (1, 20 and 35) resolved on a 2% agarose gel is shown in figure 3.8.

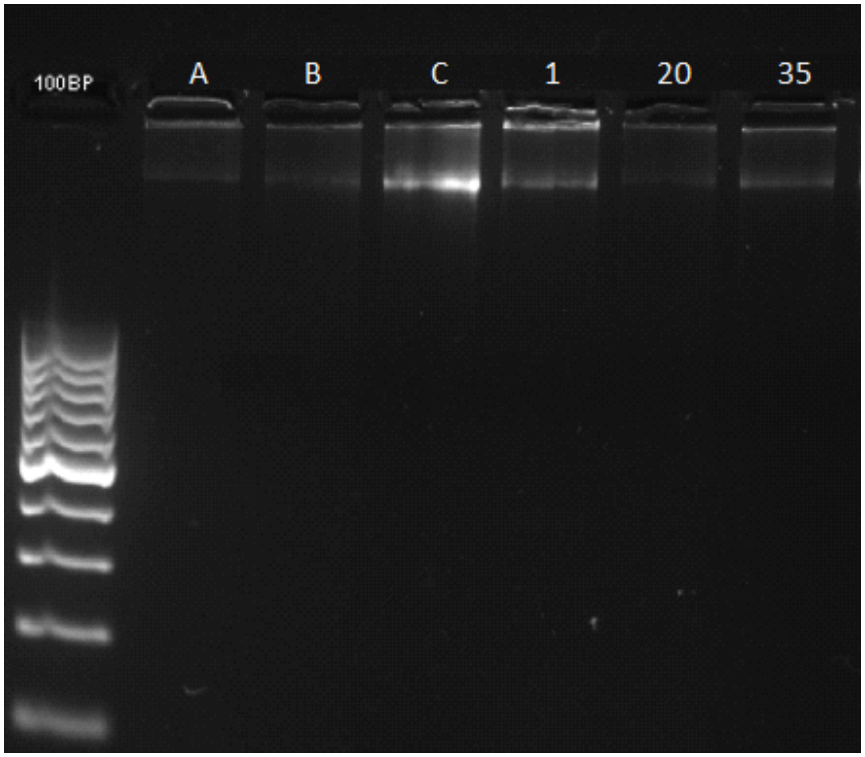

Figure 3.8 Genomic DNA after extraction
Healthy control samples (A, B and C) and ALL samples (1, 20 and 35) had DNA with good integrity shown with the presence of bands >1000bps.

Figure 3.8 is a depiction of the genomic DNA extracted from healthy control and ALL samples. After successful extraction of DNA, MLPA was then used to test for *IKZF1* deletions.

3.5 MLPA

Copy number screening of *IKZF1* exons was carried out using MLPA. The SALSA P-335 MLPA kit from MRC Holland was used on 41 leukemic samples and 10 healthy controls (A-J) in this study.

3.5.1 MLPA Quality control

MLPA was shown to be a very sensitive technique. Concentration of sample DNA was an important factor for MLPA analysis. All 51 samples that were tested using MLPA were made up to a concentration ranging between 30ng/µl and 40ng/µl (150ng and 200ng respectively) of DNA, this concentration of DNA was found to produce good MLPA peak profiles which were used to determine MLPA probe copy number. The Q-fragments (64, 70, 76 and 82 nucleotides in length) gave an indication of the DNA ligation step during the MLPA reaction.

The D-fragments (88 and 96 nucleotides in length) gave an indication in the denaturation step of the MLPA reaction. The Q and D-fragments from control sample A is shown in figure 3.9.

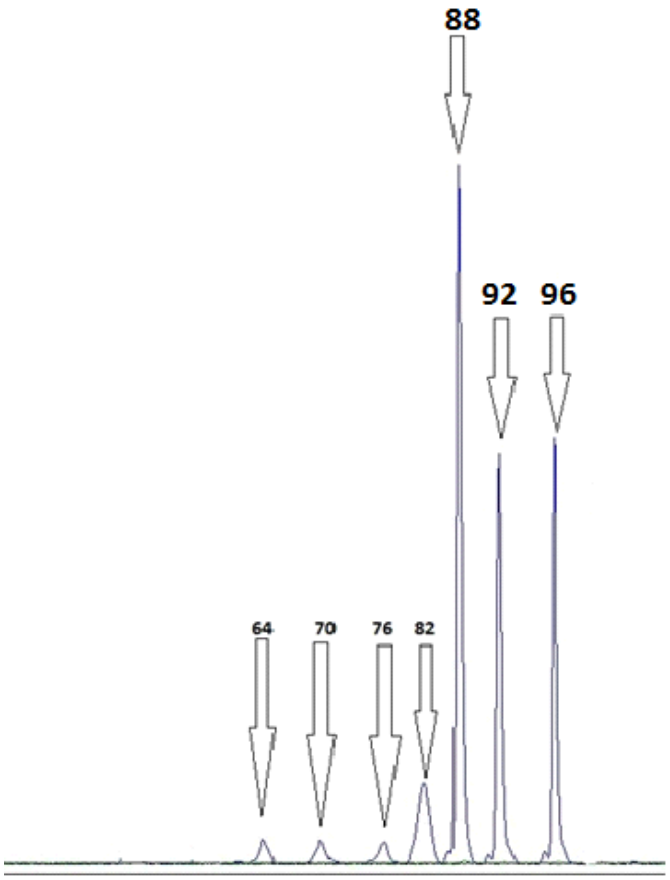

Figure 3.9 MLPA profile of internal control probes for control sample A
MLPA profile of control sample depicting the internal quality control fragments: 92bp ligation-dependent probe, Q-fragments (64, 70, 76 and 82 bps) and D-fragments (88 and 96 bps). The peak heights from the control fragments indicated that sufficient DNA (150ng) was used in this MLPA reaction (Q-fragments) and that denaturation was complete (D-fragments). The 92bp ligation probe indicated that ligation was successful.

If any of the internal control fragments did not display similar peaks as those shown in figure 3.9 then that sample was repeated for the MLPA reaction. All of the samples (controls and patient samples) displayed successful D and Q-fragment profiles such as the profile shown in figure 3.9. The rest of the MLPA probes ranging in size between 124bps – 503bps had peaks that were similar to this 92bp ligation probe shown in figure 3.9.

Generally there was a tendency of all samples (control and patient) to display a MLPA profile with a sloped appearance, also known as the skiing or sloping effect, figure 3.10

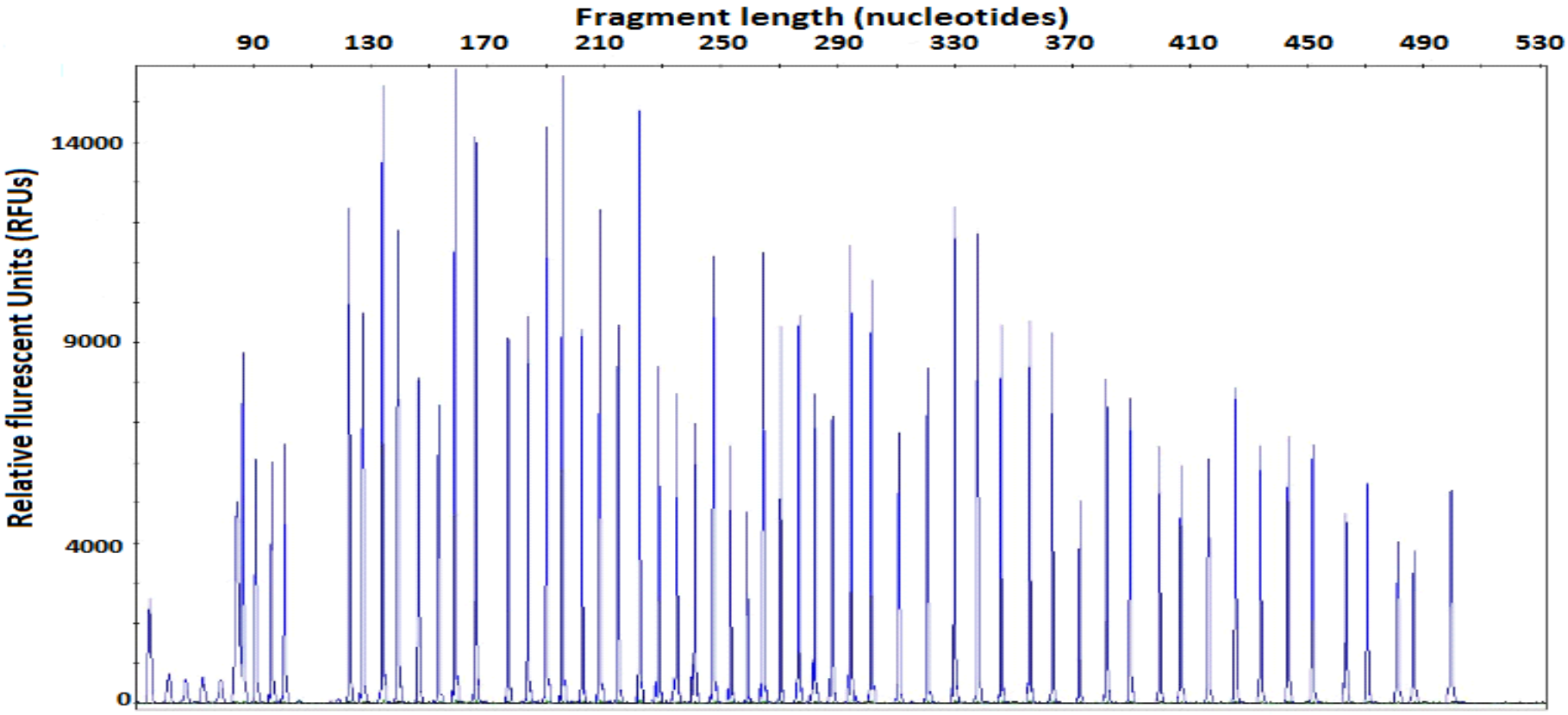

Figure 3.10 A representation of the MLPA profile of control sample A

The sloping effect of MLPA peaks. Smaller fragments have higher peak signals compared to larger fragments which have lower peak signals. This peak profile was seen in most of the patient and control samples.

The sloping effect in figure 3.10 can be attributed to the preferential movement of smaller fragments over larger fragments through the capillary. Initially the MLPA peak profiles were visualised for each sample and they all displayed a slope like pattern similar to the pattern in figure 3.10. From the 41 samples tested using MLPA, 39 showed an MLPA profile similar to the one in figure 3.10 and were used further for MLPA analysis. However, on rare occasions some samples displayed an irregular peak profile. These samples failed due to an unsuccessful ligation step or insufficient DNA. Irregular peak profiles of 2 scenarios (cases 4 and 16) are shown in figure 3.11 (a) and 3.11 (b).

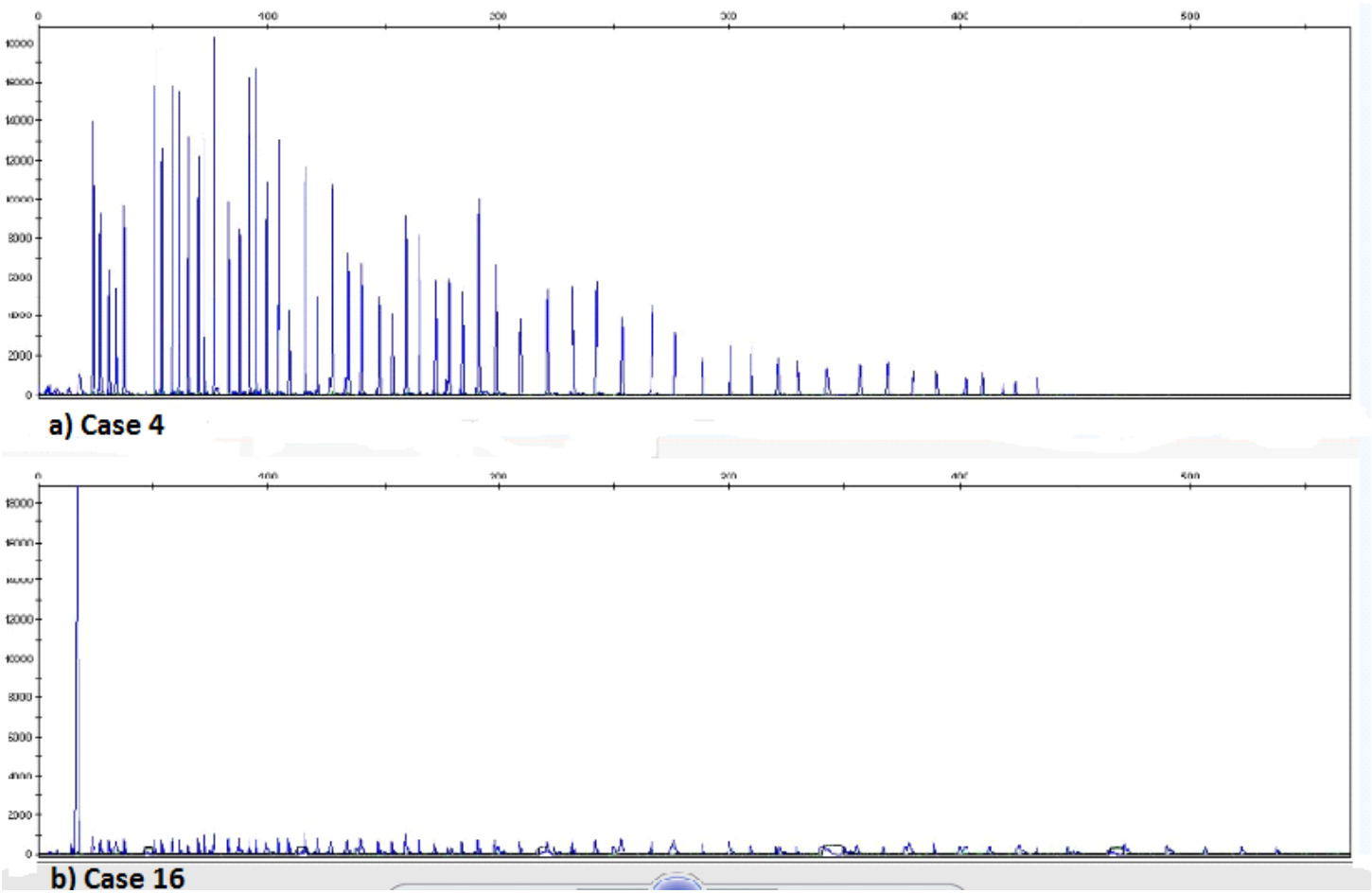

Figure 3.11 Unusual MLPA profiles

Case 4 (a) had a drastic drop in peak height as shown in. Some peak profiles were very low shown in figure 4 (b).These samples were rejected.

Although MLPA peaks generally displayed a slope or reduced peak height for longer probes, the peak heights do not drop as sharply as in the MLPA profile seen in figure 3.11 (a). Samples with these types of MLPA profiles with a drastic decrease in peak heights were not used for further MLPA analysis. These samples were repeated for the MLPA reaction or the samples were rerun on the 3500 Genetic analyser.

In another scenario, some samples displayed very high peaks with signals of >25000 relative fluorescent units (RFUs). Some samples displayed very low peaks with signals of <1000 RFUs as shown in figure 3.11 (b). A cut off value between 1000 - 25000 RFUs for each MLPA peak profile was used to accept or reject the MLPA result (personal communication Applied Biosystems).

Samples that did not meet the cut off RFU range could have resulted from other factors. There could have been incomplete hybridization in these samples or too much PCR product was added to the 96 well plate before loading it onto the genetic analyser. It is recommended that 0.7ul of MLPA PCR product, 0.3ul of ROX 500 standard and 9ul of HiDi formamide are loaded onto the genetic analyser. In both cases where there were extreme peak heights, the PCR product quantities were adjusted by increments of 0.5ul before they were run on capillary electrophoresis. Thus, a valid MLPA result was only accepted if the D and Q values and the peak profile for each sample were met the criteria mentioned above.

3.5.2 Summary of MLPA results

After MLPA analysis it was observed that 19 cases showed no aberrations in all genes tested. In the remaining 22 cases a total of 156 copy number changes were detected (7.1 aberrations per sample). There were 74 copy number gains and 82 deletions detected across these cases. Most of the copy number gains and deletions were monoallelic. The scope of all copy number gains and deletions found across the 22 leukemic samples that had copy number alterations are shown in the pie chart in figure 3.12.

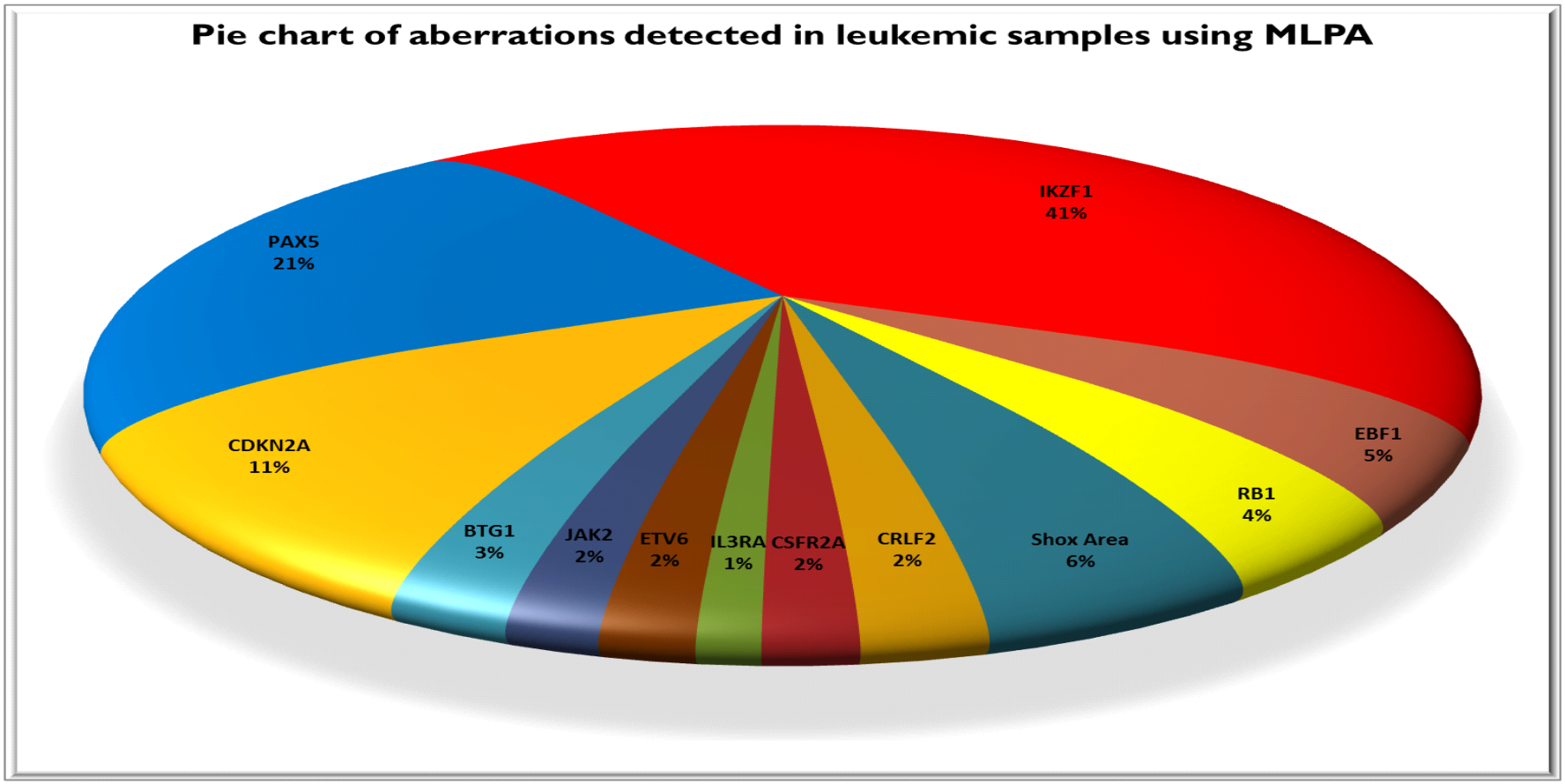

Figure 3.12 Pie chart depicting the breakdown of the aberrations (copy number gains and deletions) detected in the leukemic samples tested
The most common aberration in all 22 leukemic cases involved *IKZF1* (41%). *IKZF1*, PAX5 and *CDKN2A* accounted for the majority (73%) of aberrations in the entire cohort.

The pie chart in figure 3.12 shows that the majority of aberrations involved *IKZF1*. *PAX5* showed the second most aberrations and *CDKN2A* showed the third most aberrations. *PAX5* showed amplifications in most of the 6 exons in 6 cases that were tested and 2 cases showed deletions of all 6 exons of *PAX5*. All 3 exons of *CDKN2A* were found to be amplified in 4 cases and deleted in 3 cases.

3.5.3 Nature of *IKZF1* deletions

From the 22 cases that had copy number alterations, 15 samples had *IKZF1* aberrations: 12 samples had a deletion of *IKZF1* and 3 samples had a copy number gain of *IKZF1.* MLPA showed that exons 3-7 of *IKZF1* were the most frequently deleted exons (66-83%) and that 11 of the 12 cases that contained *IKZF1* deletions were ALL patients, one AML patient (case 20) had an *IKZF1* deletion. Case 19 was the only case that had a biallelic deletion of *IKZF1* that involved exons 4-8. DQ values were close to 0 (0.06 – 0.09) for these 5 exons in case 19. A section of the MLPA profile of case 19 is shown in figure 3.13.

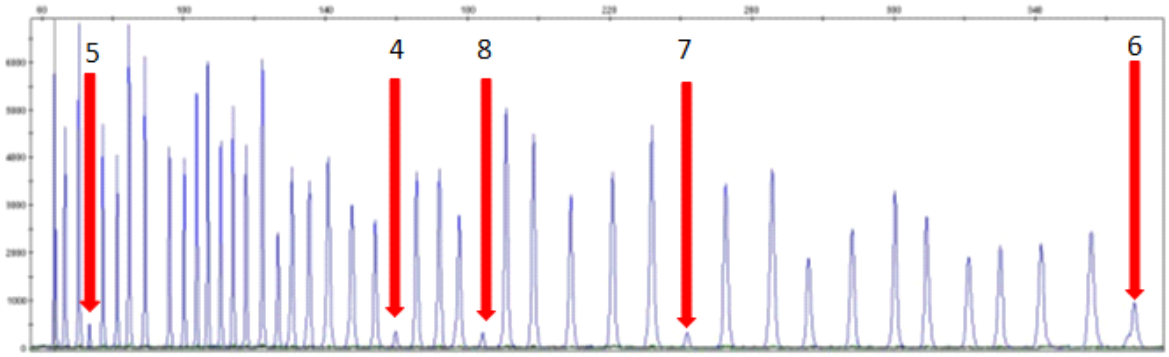

Figure 3.13 MLPA electropherogram of case 19
Case 19 had a biallelic deletion of exons 4-8 of *IKZF1*. The red arrows indicate which peak belongs to which exon of *IKZF1*.

One can easily eyeball the 5 low peaks that correspond to exons 4-8 of *IKZF1* shown in figure 3.13. In this type of MLPA profile one can easily spot the lower peaks because this case had a biallelic deletion of these exons of *IKZF1*. However, for monoallelic copy number changes one cannot use a visual interpretation of the MLPA electropherogram. For this reason, statistical analysis using regression lines and reference samples were used in order to determine copy number; this was done using the Coffalyser software.

3.5.4 MLPA and RT-PCR of *IKZF1*

The *IKZF1* exons that were deleted in each of the ALL cases after MLPA are shown in figure 3.14, they were also compared with *IKZF1* isoform expression after RT-PCR (where available).

	***IKZF1* exon number deleted (MLPA)**								***IKZF1* isoforms expressed (RT-PCR)**
	1	2	3	4	5	6	7	8	
Case Number									
5				■	■	■	■	■	Ik1, Ik2, Ik3, Ik4, Ik4A, Ik5, Ik6, Ik7, Ik8, IkX
8		■	■		■			■	Ik1, Ik2, Ik3, Ik4, Ik4A, Ik5, Ik6, Ik7, Ik8, IkX
9			■		■				Ik2, Ik3, Ik4, Ik6, IkX
16		■		■					
17	■	■	■	■					
19					■	■	■	■	
20	■	■	■	■	■	■	■	■	
23				■	■	■	■		
25	■	■	■	■	■	■	■	■	
29			■			■			
32			■	■	■	■	■		
37			■				■		
P				■	■	■	■		Ik6

Figure 3.14 Specific exons of *IKZF1* that were deleted after MLPA analysis
The red blocks indicate the particular exon that was deleted after MLPA and were compared to *IKZF1* isoform expression pattern from RT-PCR that was available for cases 5,8 9 and p.

It was possible to compare MLPA and RT-PCR for 4 cases in which both DNA and RNA specimens were available. Literature has shown that an *IKZF1* deletion that involves exons 4-7 results in exclusive expression of Ik6 (Mullighan et al, 2008; Iacobucci et al, 2009 & Dupuis et al, 2012), as shown with case p in this study. There has been no direct correlation found between any other combination of deleted exons of *IKZF1* and *IKZF1* isoform expression pattern (Mullighan et al, 2008; Iacobucci et al, 2009 & Dupuis et al, 2012), the same observation was made with cases 5, 8 and 9.

3.5.5 Copy number changes in relapsed and Ph+ ALL cases

One of the main focuses of this study was to investigate high risk ALL patients: Ph+ and relapse ALL patients. There were 15 Ph+ and 3 relapse ALL cases (adult and paediatric) tested in this study. The aberrations found in these ALL patients are shown in table 3.2.

Table 3.2 Copy number gains (green) and deletions (red) detected in Ph+ ALL patients

Ph+ ALL (cases 5, 8, 9, 10, 14, 15, 17, 23, 24, 25, 26 28, 29, 31 and 32) and relapsed ALL patients (cases 16, 19 and 36).

Case Number	Ph status	No Mutations	BTG1	CDKN2A	CRLF2	CSF2RA	EBF1	ETV6	IKZF1	IL3RA	Jak2	PAX5	RB1	SHOX area
Case 5	Ph+													
Case 8	Ph+													
Case 9	Ph+													
Case 10	Ph+													
Case 14	Ph+													
Case 15	Ph+													
Case 16	Ph-													
Case 17	Ph+													
Case 19	Ph-													
Case 23	Ph+													
Case 24	Ph+													
Case 25	Ph+													
Case 26	Ph+													
Case 28	Ph+													
Case 29	Ph+													
Case 31	Ph+													
Case 32	Ph+													
Case 36	Ph-													

Black: No copy number changes. Green: copy number gains. Red: deletions

Table 3.2 illustrates that *IKZF1* deletions were the most frequent deletion that was detected (61%) in Ph+ or relapsed ALL cases, with 8/15 (53%) patients having the Ph chromosome. The three other patients who had *IKZF1* deletions, although Ph-, were relapse ALL patients. The *IKZF1* deletions in these patients accounted for the majority of the *IKZF1* deletions (11/12) in the entire study. The frequency of copy number changes of other genes that were detected together with an *IKZF1* deletion is shown in figure 3.15.

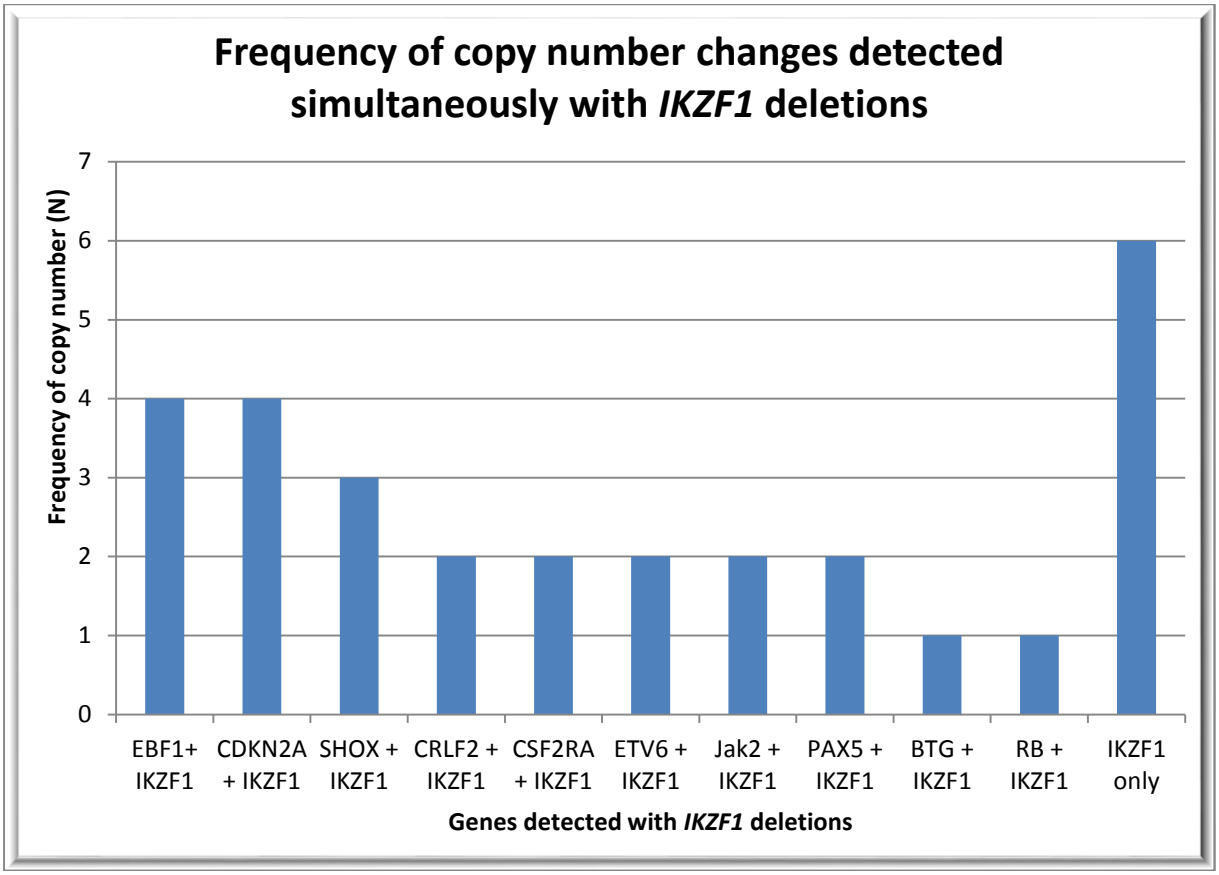

Figure 3.15 Copy number changes of genes that occurred together with an *IKZF1* deletion

Half of the cases only had an *IKZF1* deletion with no other aberrations in any of the other genes that were tested after MLPA. The other half of the cases had an *IKZF1* deletion along with copy number changes of other genes at a frequency of 8-30%. Literature has shown that copy number changes of these genes occur at a similar frequency (16-30%) in high risk ALL patients (Mullighan et al, 2009 & Iacobucci et al, 2012).

3.5.6 Copy number changes in Ph- ALL, CML and AML cases

The rest of the ALL (Ph-) samples tested in this study did not have any *IKZF1* deletions but had copy number changes that involved exons of other genes. There were also 5 CML and 5 AML cases that were tested using MLPA. All these cases are shown in table 3.3.

Table 3.3 Amplifications (green) and deletions (red) detected in the rest of the Ph- ALL, CML and AML cases

CML patients (cases 18, 22, 27, 38 and 39) and AML patients (cases 11, 12, 20, 30 and 34) are also shown below.

Case Number	Ph status	No Mutations	BTG1	CDKN2A	CRLF2	CSF2RA	EBF1	ETV6	IKZF1	IL3RA	Jak2	PAX5	RB1	SHOX area
Case 1	Ph-													
Case 2	Ph-													
Case 3	Ph-													
Case 4	Ph-													
Case 6	Ph-													
Case 7	Ph-													
Case 11	Ph-													
Case 12	Ph-													
Case 13	Ph-													
Case 18	Ph+													
Case 20	Ph-													
Case 21	Ph-													
Case 22	Ph+													
Case 27	Ph+													
Case 30	Ph+													
Case 33	Ph-													
Case 34	Ph+													
Case 35	Ph-													
Case 37	Ph-													
Case 38	Ph+													
Case 39	Ph+													
Case 40	Ph-													
Case 41	Ph-													

Black: No copy number changes. Green: copy number gains. Red: deletions

Table 3.3 shows that the majority (57%) of the leukemic patients did not have any copy number changes after MLPA analysis. The rest of the cases showed a random pattern of aberrations. Table 3.3 also shows that 1 of the 5 AML cases (case 20), had a deletion of *IKZF1*. There were no *IKZF1* deletions detected any of the CML cases that were tested. However, cases 18 and 38 both showed copy number gains of *IKZF1*, *CDKN2A* and *PAX5*.

Most of the copy number changes detected after MLPA analysis were also compared to ploidy karyotyping which explained some of the results detected by MLPA.

3.6 MLPA and Ploidy karyotyping

Ploidy karyotyping has been shown to be comparable with MLPA results (Boormans, 2010 & Vorstman, 2006). Results from this study also showed that ploidy karyotyping could be correlated with MLPA and was available for most of the cases that were tested for MLPA; these cases are shown in table 3.4. The chromosomal position for each MLPA probe is listed in the parenthesis and the number of cells analysed is shown in the square parenthesis. The nomenclature used for defining a cytogenetic karyotyping is stated according to the ISCN (International system of Human Cytogenetics) 2009 where a "?" denotes that the identification of the exact region on a chromosome or band is questionable. The regions on which the MLPA probes target are also shown in parenthesis.

Table 3.4 Comparison of Diploid karyotyping and MLPA results

Case Number	Karyotyping	MLPA	
		Copy number gain	Copy number deletion
Case 5	46,XX,t(9;22)(q34;q11) [16]		*IKZF1* (7p12.2)
Case 7	26,X,+X,del(3)(p?22),t(17;?)(p11.2;?),+18,+21,inc[cp18]/43~52,idem,del(1)(q21),add(13)(q34),add(16)(q24),+20,+22[cp7]/46,XX[4]		*ETV6* (12p13.2)
Case 8	46,XY,t(9;22)(q34;q11) [18]		*IKZF1* (7p12.2)
Case 9	46,XY,t(9;22)(q34;q11.2) [15]		*IKZF1* (7p12.2)
Case 16	33,X,-Y,+1,+5,+6,+del(9)(q?31),+10,+11,+18,+der(19)t(8;19)(q13;p13.2),+21,+22 [15]/46,XY [5]	*CDKN2A* (9p21), *EBF1* (5q33.3), *PAX5* (9p13.2)	*IKZF1* (7p12.2)
Case 17	47,XY,+X,t(9;22)(q34;q11)[14]/46,XY[6]	*CDKN2A* (9p21), *CRLF2* and *IL3RA* (Xp22.3/Yp11.3)	*IKZF1* (7p12.2)
Case 18	46,XX,t(9;22)(q34;q11) [7]	*CDKN2A* (9p21), *IKZF1* (7p12.2), *PAX5*(9p13.2)	
Case 19	46,XY,t(9;10)(p24;q?24) [19]/46,XY [1]	*SHOX*(Xp22.33)	*CDKN2A* (9p21), *CRLF2*, *IKZF1* (7p12.2), *EBF1*(5q33.3)
Case 20	45,XX,add(1)(p36.3),t(3;12)(q?26,q?13),-7,del(14)(q?)[17]		*IKZF1* (7p12.2)
Case 23	46,XX,t(9;22)(q34;q11) [21]		*IKZF1* (7p12.2)
Case 25	46,XX,t(9;22)(q34;q11.2) [7]		*IKZF1* (7p12.2)
Case 26	46,XX,t(9;22)(q34;q11),inv(14)(q31q32)[4]/46,XX,del(1)(q42),der(9)t(1;9)(q31;p22)t(9;22)q(34;q11),inv(14)(q31q32) [13]/46,XX [3]	*CDKN2A* (9p21)	
Case 29	46,XY,t(5;7)(q35;q11.2),t(9;22)(q34;q11.2)[15]		*IKZF1* (7p12.2), *EBF1* (5q33.3)
Case 31	46,XY,t(9;22)(q34;q11.2) [10]		*CRLF2*, *IKZF1* (7p12.2), *Jak2* (9p24.1), *PAX5* (9p13.2)
Case 33	47,XY,+8,der(19)t(1;19)(q23;p13) [14]/46,XY [1]	*CDKN2A* (9p21), *PAX5* (9p13.2)	
Case 36	45~46,X,-Y,i(7)(q10),der(9)del(9)(p23)add(9)(q34),t(20;?)(q?13.2;?),+add(21)(p13),+i(21)(q10)[cp 14]/46,XY [2]		*IKZF1* (7p12.2), *CRLF2* and *IL3RA* (Xp22.3/Yp11.3), *PAX5* (9p13.2), *RB1*(13q14.2)
Case 37	46,XY[20]	*CDKN2A* (9p21), *PAX5* (9p13.2)	*Jak2* (9p24.1)
Case 38	46,XX,t(9;22)(q34;q11) [15]	*CDKN2A* (9p21), *IKZF1* (7p12.2), *PAX5* (9p13.2)	
Case 41	46,XY [8]		*RB1* (13q14.2)

Green: Copy number gain. Red: Deletion

3.6.1 Ploidy karyotyping and MLPA analysis in ALL samples

It was observed that there were advantages and disadvantages of MLPA when compared to cytogenetics.

From table 3.4 *IKZF1* deletions could only be detected with MLPA. *IKZF1* deletions were detected in Ph+ ALL patients with no additional chromosomal aberrations (cases 5, 8, 9, 23, 25 and 31) and in cases which had additional chromosomal aberrations (cases 16, 17, 19, 29 and 36) with the Ph chromosome. There were two cases in which IKZF1 deletions were detected with rare translocations: Case 19 had a rare translocation involving chromosomes 9 and 10 and case 29 had rare translocation involving chromosomes 5 and 7.

There were four additional cases (19, 29, 37 and 41) in which MLPA identified deletions or copy number gains which could not be detected by chromosomal analysis. Cases 37 and 41 were cytogenetically normal yet MLPA detected micro deletions and copy number gains in these cases. MLPA also detected micro deletions of *CDKN2A, CRLF2* and *EBF1* in case 19. All three genes are located on chromosomes which had normal copy number according to cytogenetics. These cases showed the higher sensitivity of MLPA over cytogenetics.

However, MLPA could be correlated with certain aneuploidies detected with cytogenetics. Case 16 had an extra copy of chromosome 5 shown by its karyotype. MLPA detected copy number gains involving all exons of *EBF1*, which is located on chromosome 5. Case 17 had copy number gains of *CRLF2* and *IL3RA,* these genes are located on the PAR of chromosome X and Y. Cytogenetics showed that case 17 had an extra chromosome X which also explains the copy number gains detected by MLPA. Case 36 had a deletion of *CRLF2* and *IL3RA* which could be attributed to loss of the Y chromosome from its karyotype.

There was a disadvantage of MLPA compared to cytogenetics. Cases 7 and 16 had severe ploidy changes. Reference genes that are used to determine MLPA copy number are located on chromosomes that were not diploid. Although MLPA detected copy number changes in case 16, case 16 has an additional chromosome 1 and 11 on which reference genes are located to normalise MLPA ratios. This could skew MLPA normalisation.

3.6.2 Ploidy karyotyping and MLPA analysis in AML and CML patients

Case 20 was the only AML patient had an *IKZF1* deletion and chromosomal analysis showed that the patient only had one chromosome 7, the same chromosome on which *IKZF1* is located.

3.7 MLPA validation using FISH

MLPA results were validated using FISH. FISH detects copy number of ETV6 using the (t(12:21)) Vysis probe. *ETV6* is located on chromosome 12 and *RUNX* is located on chromosome 21. The FISH probe specific for *ETV6* is orange and the FISH probe specific for *RUNX* is green. The MLPA kit has 6 probes specific for exons 1 (two probes), 2, 3, 5 and 8 of ETV6 (Insert SALSA MLPA probemix P335-B1 ALL-IKZF1, MRC Holland). Three known patient samples (V1, V2 and V3) with monoallelic deletions of *ETV6* and an unaffected control with 2 copies of *ETV6* were tested using MLPA and FISH. MLPA confirmed FISH results: FISH detected a single copy of the *ETV6* gene (90% of 200 cells) and MLPA detected a monoallelic deletion of the *ETV* exons in the above mentioned samples. MLPA and FISH also showed that the unaffected control had 2 copies of ETV6 (figure 3.16).

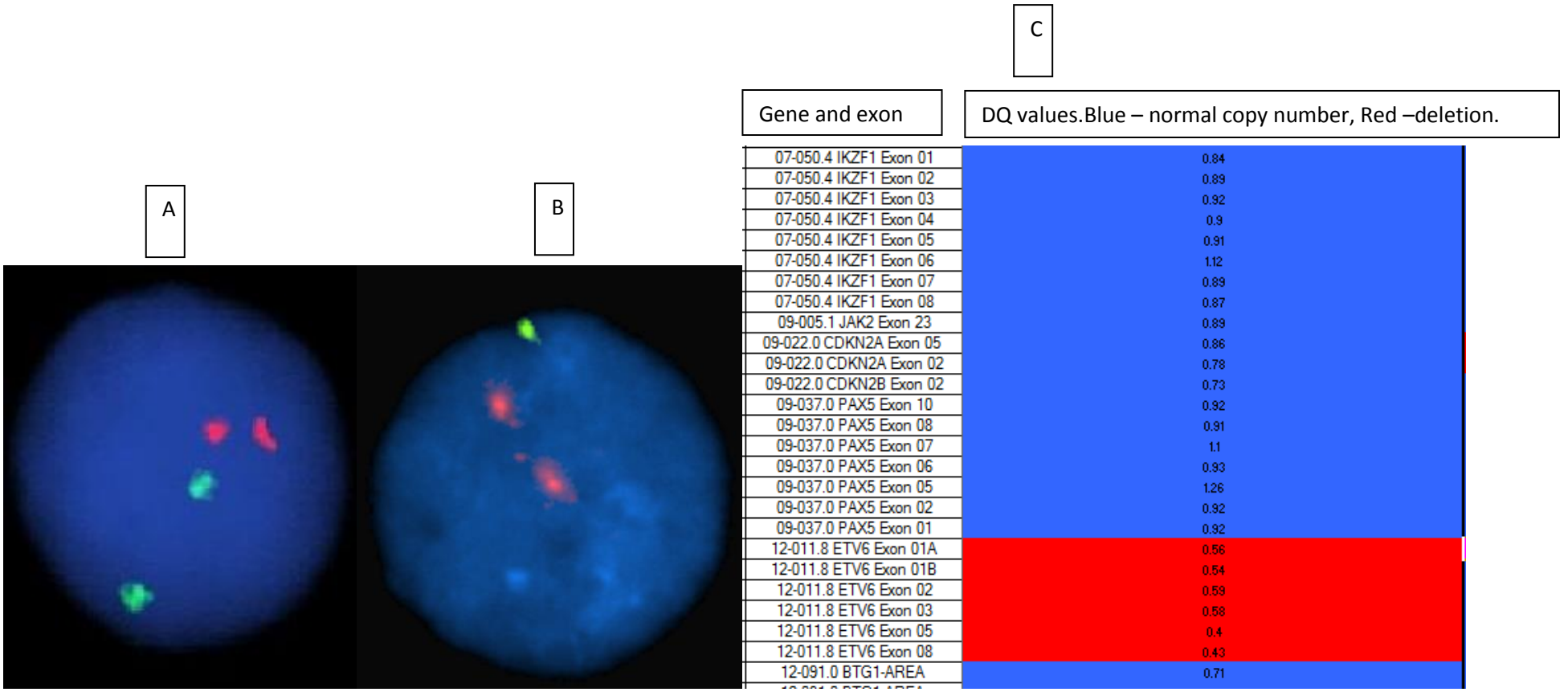

Figure 3.16 A FISH image using the Vysis LSI ETV6\RUNX dual colour probe

Normal interphase cell (A) from a control sample A. This cell shows two green signals (*ETV6*) and two red signals (*RUNX*) inferring that both copies of *ETV6* and *RUNX* are present in control A. The abnormal cell V1 (B) shows that this patient had only one copy of *ETV6* represented by the one green signal. C is the MLPA result for patient V1. MLPA ratios for *ETV6* indicates that V1 has a monoallelic deletion of exons 1A, 1B, 2, 3, 5 and 8 (ratio below 0.7). Copy number for some of the genes also shown here, *IKZF1, CDKN2A, BTG1-AREA* and *PAX5* are shown to have exon ratios close to 1 which are normal copy number.

Figure 3.16b shows that V1 had a deletion of *ETV6* with the presence of only one green signal. MLPA also shows a monoallelic deletion of *ETV6* with DQ values between 0.4 - 0.59 for the same patient (figure 3.16c). MLPA and FISH were tested on the same 3 patients using DNA and both were found to harbour the same deletion. In this way FISH validated MLPA.

Chapter 4. Discussion

Over the past 4 years *IKFZ1* deletions have been established as the hallmark of Ph+ ALL cases and are found in ~72% of these patients (Mullighan et al, 2008; Iacobucci et al, 2009; Kuiper et al, 2010; Martinelli et al, 2009; Harvey et al, 2010; Mi et al, 2012; Caye et al, 2013 & Dupuis et al, 2012). Studies have also shown that there is an excess of dominant negative *IKZF1* isoforms, especially Ik6 detected in ALL patients. It has also been established that *IKZF1* deletions are preserved from diagnosis until relapse and are found in 54 – 75% of ALL patients who relapse. Relapse is the main reason of treatment failure in ALL and it has been difficult to predict patients that relapse in low, medium and high risk groups. Thus, *IKZF1* deletions could be used to predict relapse in these patients (Mullighan et al, 2008; Iacobucci et al, 2009 & Kuiper et al, 2010).

In this project we sought to determine the frequency of *IKZF1* deletions in a cohort of South African ALL patients (Ph+ and Ph-) using MLPA and RT-PCR to determine the *IKZF1* expression pattern. Both methods were validated using FISH and Sanger sequencing respectively.

4.1 RT-PCR of *IKZF1*

IKZF1 isoform Ik6 is the most common dominant negative isoform detected in Ph+ and relapsed ALL patients (Mullighan et al, 2008; Iacobucci et al, 2009 & Kuiper et al, 2010). In this study 12 Ph+ and 12 relapsed ALL patients (adult and paediatric) were tested using RT-PCR. Nineteen out of twenty four ALL patients (79%) expressed the dominant negative isoforms (Ik 4A, 5, 6, 7 and 8), of which 17/19 (94%) expressed the isoform Ik6 (table 3.1). Ik6 was present in 71% (17/24) of all the ALL samples that were tested. Dominant negative isoforms were more frequently detected in these patients compared to healthy individuals in which only two patients (20%) expressed dominant negative isoforms.

Ten healthy individuals (A-J) were analysed for *IKZF1* isoforms expression. Although the 10 healthy individuals all expressed the functional *IKZF1* isoforms, two healthy controls (cases B and E) also expressed dominant negative isoforms Ik4A, Ik5, Ik6 and Ik8 in addition to functional isoforms although in low quantity. These dominant negative *IKZF1* isoforms were

shown by the presence of faint bands after gel electrophoresis (figure 3.3). Earlier studies (Sun et al, 1999a & Yagi et al, 2002) did not detect any dominant negative isoforms in healthy cells after RT-PCR and most of the recent studies (Mulligan et al, 2008; Martinelli et al, 2009 & Kuiper et al, 2010) did not list all the *IKZF1* isoforms expressed in healthy individuals. However two studies (Payne et al, 2001 & Iacobucci et al, 2009) detected dominant negative isoforms (including Ik6) of *IKZF1* in healthy individuals at low levels, as in this study.

4.2 RT-PCR and MLPA analysis of *IKZF1*

In the study presented here there were 13 cases (cases 1 – 13) in which RNA and DNA specimens were available for analysis of *IKZF1* using RT-PCR and MLPA respectively. Cases 5, 8 and 9 were Ph+ ALL patients. All three cases had a monoallelic deletion that involved different exons of *IKZF1* on MLPA analysis, RT-PCR showed that the same three cases expressed functional and dominant negative isoforms of *IKFZ1*.

Cases 5 and 8 expressed the most dominant negative isoforms simultaneously, 10 isoforms in total (Ik1, Ik2, Ik3, Ik4, Ik4A, Ik5, Ik6, Ik7, Ik8, IkX). Genomic deletions involving various exons of *IKZF1* have shown to produce an excess number of different dominant negative isoforms simultaneously (Iacobucci et al, 2009; Martinelli et al, 2009 & Mullighan et al, 2008), as was seen in these two cases.

4.2.1 MLPA and RT-PCR of *IKZF1* in case p

Case p showed a direct correlation between the MLPA and RT-PCR results. This case was a known positive sample which had an *IKZF1* deletion of exons 4-7 on SNP6 array analysis (personal communication). MLPA also detected a monoallelic deletion involving the same exons and RT-PCR revealed that case p only expressed the isoform Ik6. Studies have shown a concordance between genomic deletions that involve exons 4-7 of the *IKZF1* gene and the exclusive expression of Ik6 (Mullighan et al, 2008; Iacobucci et al, 2009 & Dupuis et al, 2012). It has been suggested that a recombination event may be the mechanism responsible for the deletion of exons 4-7. The breakpoints of this deletion occur in introns 3 and 7 of the *IKZF1* gene sequence (Mullighan et al, 2008). This breakpoint region is flanked by highly

conserved recombination sequences (RSS). These RSS sequences are recognised by recombination activating gene (RAG) enzymes, most likely deoxynucleotidyl transferase (TdT) which facilitates variable diverse joining (VDJ) recombination. During RAG-mediated recombination exons 4-7 of *IKZF1* are deleted (Mullighan et al, 2008).

Case 48 and case 51 also showed the exclusive expression of Ik6. Based on the results found from case p in this study and in the literature we can assume that exclusive expression of Ik6 resulted from a genomic deletion of *IKFZ1* exons 4-7. Unfortunately there was no DNA available from these two patients for testing MLPA. DNA sequencing of cases p, 48 and 51 confirmed the presence of Ik6 with exon 3 juxtaposed next to exon 8 (figure 3.7).

4.2.2 Alternate mechanism of *IKZF1* isoform expression

Although focal deletions have been established as the main mechanism of *IKZF1* deletions in ALL patients, two studies have shown that mutations that cause insertions of *IKZF1* lead to aberrant alternate splicing of *IKZF1*. Aberrant alternate splicing of *IKZF1* resulted in expression of additional *IKZF1* isoforms: Ik2(ins), Ik4(ins), Ik4A(ins) and Ik5A(ins). Consequently, the additional isoforms function as the dominant negative isoforms and impair *IKZF1* function (Iacobucci et al, 2008a and Sun et al, 1999b). These aberrant isoform sizes were not detected in this study after RT-PCR and sequencing of *IKZF1*.

4.2.3 Summary of RT-PCR of *IKZF1*

Multiplex nested RT-PCR using a single set of primers amplified all possible isoforms generated by *IKZF1* in healthy and patient samples. The common deletion that involve exons 4-7 of *IKZF1*results in exclusive expression of Ik6 (Mullighan et al, 2008; Iacobucci et al, 2009 & Zhou et al, 2011), this was also confirmed in this study. Deletion 4-7 was the only *IKZF1* that showed concordance when testing DNA and RNA from the same patient.

Data from this study showed that *IKZF1* isoform expression in Ph+ and relapsed ALL patients differs compared to healthy individuals. The frequency of the *IKZF1* dominant negative isoforms (especially Ik-6) detected in ALL patients were more frequently detected compared to isoforms detected in healthy cells, although it was also shown that healthy individuals also express dominant negative isoforms less frequently and in smaller quantities.

4.3 *IKZF1* deletions in Ph+ and relapse ALL patients after MLPA

A total of 41% (65/158) of all copy number changes involved *IKZF1*, the most common copy number change detected by MLPA. There were 11 *IKZF1* deletions detected in 18 ALL cases, five cases (5, 8, 9, 16, 17, 19 and 31) were paediatric ALL Ph+ patients and three cases (23, 25, 29 and 36) were adult Ph+ patients. *IKZF1* deletions were detected at presentation and at relapse in 11/18 cases by MLPA.

Out of 15 Ph+ cases, 53% (8) had *IKZF1* deletions. The frequency of *IKZF1* deletions in Ph+ ALL (paediatric and adult patients) was 72% (n=186) in total from different studies (Mullighan et al, 2008; Iacobucci et al, 2009 & Martinelli et al, 2009). The results of this study differ from published data, as the frequency of *IKZF1* deletions was not as high as described in the literature. This could be due to the fact that the cohort of Ph+ ALL patients used in this study was much smaller than published data; various studies tested patients over a period of 15-25 years (Mullighan et al, 2008; Iacobucci et al, 2009 & Martinelli et al, 2009). Retrospective and prospective samples tested in this study were only available from the past 5 years. However, it is clear that the Ph+ (adult and paediatric ALL patients) had more *IKZF1* deletions compared to Ph- ALL patients because the majority of *IKZF1* deletions (8/11) were detected in Ph+ ALL patients.

4.3.1 Nature of *IKZF1* deletions

All the *IKZF1* deletions detected by MLPA were monoallelic except for one case, case 19 which had a biallelic deletion involving exons 4-8. Biallelic deletions of *IKZF1* although rare, have been detected previously (Mullighan et al, 2008). Deletions involving almost any *IKZF1* exon (2-8) result in a null mutation (Kastner et al, 2013; Iacobucci et al, 2008b & Martinelli et al, 2009). All the *IKZF1* deletions detected in this study involved at least 2 exons. The majority of *IKZF1* deletions reported thus far involve deletions of exons 2-7, 4-7 or 1-8 (Mullighan et al, 2008; Martinelli et al, 2009; Kuiper et al, 2010 & Dupuis et al, 2012). In this study two cases had a deletion of exons 1-8 and one case had a deletion of exons 4-7 of *IKZF1*, the rest of the cases had deletions of exons in various other combinations (figure 3.14). As mentioned above, this could also be due to the fact that the cohort of Ph+ ALL patients used in this study was much smaller than published data.

4.3.2 Prognostic significance of *IKZF1* deletions

IKZF1 is a transcription factor which is critical during lymphoid differentiation and development (Nakase et al, 2000 & Hosokawa et al, 2000) and in the generation of B cell precursors (Georgopoulos et al, 1997). *IKZF1* deletions are associated with a poor progonosis in leukemic patients (Mullighan et al, 2008; Iacobucci et al, 2009 & Martinelli et al, 2009).

Currently there is a large contingent (50%-60%) of medium risk paediatric ALL patients who relapse. Relapse is the main reason for treatment failure in ALL and diagnostic criteria have not been sufficient to predict relapse in these patients, especially in the medium risk group, although relapse also occurs in low and high risk ALL patients (Waanders et al, 2010). *IKZF1* deletions together with minimal residual disease (MRD) targeting rearrangements between Immunoglobulin and T cell-receptor genes were able to predict relapse in 79% of ALL patients (Waanders et al, 2010) supporting the idea that *IKZF1* deletions should be considered for stratifying risk groups in ALL patients. Consequently intensifying treatment in high risk ALL patients may improve event-free survival (EFS). Data from this study also showed that certain cases, if tested for *IKZF1* deletions at presentation could have had an impact on risk stratification and EFS. Examples of how *IKZF1* deletions could have assisted with prognosis in this study will be described in two cases.

Case 8, an 8 year old male patient presented with a normal karyotype and FISH did not detect a Ph chromosome or any rearrangements involving the *MLL* gene (associated with a poor prognosis) or a fusion between *ETV* with *RUNX* (associated with a good prognosis). RQ-PCR detected a Ph chromosome 6 months later, the same time an *IKZF1* deletion was detected. This patient may have been stratified as a high risk ALL patient if he had been tested for *IKZF1* copy number changes at presentation in the absence of other prognostic markers.

Case 51, a 19 year old male ALL patient presented with a Ph chromosome in 2008. He was in remission for 4 years (2009 – 2012). During this remission period (in 2011) chromosomal analysis, FISH and RQ-PCR of *BCR/ABL1* excluded the presence of the Ph chromosome; however *IKZF1* was found to be deleted. The patient then relapsed in 2013 with the same

population of Ph+ cells (shown by immunophenotyping). It has been established that *IKZF1* deletions are almost always preserved from presentation until relapse or may be acquired at the time of relapse (Mullighan et al, 2008, Yang et al, 2008 & Kuiper et al, 2010). The presence of an *IKZF1* deletion in this patient while in remission was an important finding. Had it been known that *IKZF1* was deleted during remission, this patient could have also been identified early as a high risk ALL patient.

4.3.3 Copy number changes of other genes that accompanied an *IKZF1* deletion

IKZF1 deletions were detected in 6 cases which also had copy number changes of other genes: *CDKN2A, CRLF2, EBF1, IL3RA,* Jak2, *SHOX, PAX5* and *RB1.* These genes have been detected at different frequencies with *IKZF1*. Copy number changes of *PAX5* and *CDKN2A* have been detected in ~30% of high risk ALL patients (Mullighan et al, 2009 & Iacobucci et al, 2012). *CRLF2* (16%) alterations has been detected at a lower frequency and *IL3RA* (7%) and *Jak2* (10%) copy number changes at much lower frequencies (Mullighan et al, 2009 & Iacobucci et al, 2012). The frequencies obtained from these studies have used cohorts of up to 221 high risk ALL patients (Mullighan et al, 2009). Although a much smaller cohort of ALL patients were tested in this study, *PAX5* and *CRLF2* were the most frequently detected copy number changes by MLPA, consistent with the literature (Mullighan et al, 2009 & Iacobucci et al, 2012).

It was also shown that *IKZF1* deletions can occur with chromosomal aberrations other than the Ph chromosome. Case 29, a 36 year old male patient had a Ph chromosome and a rare translocation involving chromosomes 5 and 7, with the following karyotype : 46,XY,t(5;7)(q35;q11.2),t(9;22)(q34;q11.2). There has only been one report of a patient that had the same karyotype in the literature (Ross et al, 1998). This fusion involves the platelet-derived growth factor receptor (*PDGFbR*) gene on chromosome 5q33 with the Huntingtin protein1 (HIP1) gene on chromosome 7q11. Functional studies showed that the *HIP1/PDGFbR* fusion transcript contributes to leukemogenesis and is associated with a poor prognosis (Ross et al, 1998).

4.3.4 Copy number changes of genes in ALL (Ph-), CML and AML patients after MLPA

Most of the Ph- ALL cases did not have any copy number changes after MLPA. However, *CDKN2A* and *PAX5* were also the most common copy number alteration in these patients. In addition to the Ph- ALL cases, there were also 5 AML patients and 5 CML patients tested using MLPA.

There have been no reports of *IKZF1* deletions in chronic phase CML, but *IKZF1* deletions have been identified in blast phase CML patients (Mullighan et al, 2008). There were no blast phase CML patients available at the time of this study. The five Ph+ CML patients (chronic phase) who were tested did not have any *IKZF1* deletions.

In the AML cases, four cases, 11, 12, 30 (Ph+) and 34 (Ph+) did not contain any *IKZF1* copy number changes. Case 20, a 50 year old female relapsed AML patient had a monoallelic deletion of all 8 exons of *IKZF1*. However, this deletion can be attributed to aneuploidy, revealed by cytogenetics which will be discussed below.

Although only a small cohort of CML and AML patients were tested, it gives some indication that *IKZF1* deletions are usually not seen in chronic phase Ph+ CML patients and AML patients.

4.3.5 Summary of *IKZF1* deletions detected using MLPA

It can be deduced from this study that *IKZF1* deletions can be detected at presentation and at relapse in paediatric and adult ALL patients (Ph+ and Ph-).

All of the landmark studies have used DNA microarrays as the primary technique to detect *IKZF1* deletions and only confirmed *IKZF1* deletions involving exons 4-7 with RT-PCR (Mullighan et al, 2008; Iacobucci et al, 2009 & Zhou et al, 2011). This study used MLPA and also made a correlation between the same deletion and RT-PCR with the exclusive expression of isoform Ik6.

IKZF1 deletions were not only detected exclusively in Ph+ ALL patients but were also accompanied with copy number changes involving other genes and in one case a rare translocation.

IKZF1 deletions should be tested in all newly diagnosed ALL patients which could assist with prognosis.

4.4 MLPA and Cytogenetics

MLPA results were also compared to chromosomal analysis for certain cases. When comparing MLPA to cytogenetics it was found that MLPA had advantages and disadvantages over cytogenetics and MLPA also correlated with cytogenetics in certain cases.

4.4.1 Advantages of MLPA over cytogenetics and other techniques

MLPA proved to be more informative over cytogenetics in several cases. The major advantage of MLPA over cytogenetics was the resolution at which it detected copy number changes, most importantly *IKZF1* deletions. Recent advancements in technology have allowed small copy number changes to be detected in certain genes. *IKZF1*, *PAX5, CDKN2A*, *ETV6, RB1, Jak2*, and *EBF1* all had copy number alterations by MLPA. These copy number changes could not be picked up with cytogenetics due to its lower sensitivity and all of the copy number alterations detected by MLPA involved genes that play a significant role in leukemogenesis (Mullighan et al, 2007).Except for the high frequency of *IKZF1* deletions in Ph+ ALL patients, all of the other genes have not shown to have overlap with any cytogenetic abnormalities (Mullighan et al, 2007 & Moorman et al, 2012). Furthermore cases that were shown to be cytogenetically normal had copy number changes of genes detected by MLPA. Cases 37 and 41 were cytogenetically normal but MLPA detected deletions of *Jak2* and *RB1* respectively.

MLPA is a reliable method that can be used to detect copy number using genomic DNA. In recent years the use of MLPA has become more widespread in a diagnostic setting. This is mostly due to the fact that it can detect copy number in 50 targets (including reference control genes) in a single multiplex reaction. Furthermore, MLPA only requires a thermocycler and a genetic analyser which are in most diagnostic laboratories (Stuppia et al, 2012).

MLPA has also shown to be cost effective and results can be obtained quickly (within 24 hours) compared to cytogenetics. Cytogenetics and can take up to 15 days in a diagnostic

laboratory and cost 47% more than MLPA (Boormans et al, 2010). Although MLPA is not as sensitive as DNA microarrays, it has been shown that MLPA costs ~90% less than arrays and is not as technical to analyse (Regier et al, 2010). It is possible to investigate chromosomal regions using FISH technique, however many FISH probes would have to be used in separate reactions to obtain a result for one patient. FISH is also not appropriate to detect small regions of the genome such as single exons and FISH is also costly and time consuming. MLPA detects copy number changes in contrast to DNA sequencing. Although RQ-PCR can also detect copy number, it cannot be used in a multiplex reaction in contrast to MLPA. Because of these reasons MLPA has become a fast growing technique in recent times (Stuppia et al, 2012).

4.4.2 Concordance between MLPA and cytogenetics

MLPA could detect aneuploidies in certain cases which correlated with cytogenetics.

Case 16, a 14 year old male relapsed ALL patient had several extra chromosomes; one of them was chromosome 5. MLPA detected copy number gains involving all exons of *EBF1,* which is located on chromosome 5, which correlated with MLPA results.

There was one AML patient (case 20) which had a monoallelic deletion involving all exons of *IKZF1*. Case 20, was a 50 year old relapsed AML female patient. This patient had the following karyotype: 45,XX,add(1)(p36.3),t(3;12)(q?26,q?13),-7,del(14)(q?). *IKZF1* is found on chromosome 7, which is a chromosome that is missing in this patient. Despite this patient having a poor prognosis and having a relapse of the disease, the deletion of chromosome 7 could have been the main reason why all the exons of *IKZF1* were deleted.

Chromosomal analysis of case 17 revealed that the patient had an extra copy of the X chromosome: 47,XY,+X,t(9;22)(q34;q11)[14 cells]/46,XY[6 cells]. MLPA also showed copy number gains of *CRLF2* and *IL3RA,* these genes are located on the X chromosome which validates the presence of the extra X chromosome.

Case 36 had deletions of *CRLF2*, *IL3RA, CDKN2A*a and *Jak2* by MLPA. Chromosomal analysis revealed the following karyotype in this case: 45~46,X,-Y,i(7)(q10),der(9)del(9)(p23)add(9)(q34),t(20;?)(q?13.2;?), +add(21)(p13),+i(21)(q10) [cp

14]/46,XY. *CDKN2A* is located on chromosomal region 9p21.3 and *Jak2* is found on 9p24.1 which flanks *CDKN2A* on the telomeric side. The fact that both these genes are deleted is an indication that the deletion on chromosome 9 is large. The deleted section of chromosome 9 in this patient explains why *CDKN2A* and *Jak2* were deleted after MLPA analysis. MLPA also detected deletions of exons on *CRLF2* and *IL3RA* located on the sex chromosomes in the PAR1 region in case 36. Cytogenetics showed that there was a loss of the Y chromosome in case 36. Genes located in the PAR1 region can be deleted simultaneously. The extent of the deletion in the PAR region start from *CRLF2* and extends to intron 1 on *P2RY8*, this area covers five genes in total: *CRLF2-CSF2RA-IL3RA-SLC25A6- P2RY8*, indicating that the extent of this deletion in the PAR1 region is large. Deletion of these genes in the PAR1 region leads to constitutive Jak-Stat activation which drives leukemogenesis (Mullighan et al, 2009).

4.4.3 Disadvantages of MLPA

Although MLPA could be correlated with cytogenetics, normalisation of MLPA ratios which is used to determine copy number can be skewed in certain cases that have aneuploidy.

Case 16 had a complex karyotype:
33,X,Y,+1,+5,+6,+del(9)(q?31),+10,+11,+18,+der(19)t(8;19)(q13;p13.2),+21,+22[15]/46,XY.

MLPA normalises target-specific probes using reference genes assumed to have normal copy number (2) in the same sample and in healthy reference samples. Some of the extra chromosomes (1 and 11) detected in case 16 have reference genes that are used for normalising MLPA data. MLPA data would be skewed and copy number changes detected in other MLPA targets would not be reliable.

MLPA uses DNA quantity to determine copy number. Rearrangements result in a change in the order of the DNA sequences, but not in the DNA quantity. Thus, MLPA cannot distinguish between a diploid or tetraploid genome and cannot detect balanced translocations (Insert MLPA® General Protocol, MRC Holland). Furthermore, it is difficult to detect a deletion of a gene if the sample DNA has <50% of cancer cells. The presence of normal cells can mask the cancer cells (mosaicism). Despite some drawbacks of MLPA, it is recommended that it should be used to detect *IKZF1* copy number changes when compared to other techniques (Caye et al, 2012 & Volejnikova et al, 2013).

4.5 MLPA technical discussion

It was important to ensure that the amount of DNA used from each sample (patient and control) for each MLPA run are all diluted to the same concentrations as close as possible. There was variation in MLPA peak heights if 50ng of sample was used as opposed to 120-200ng in the same MLPA run, this affected normalisation.

There was one probe in particular that showed an inconsistent peak height after most of the MLPA runs. This probe is 166bps in length and targets exon 10 of the *CSF2RA*, a region located on the X chromosome. The peak height for this probe was less (in control and patient samples) in comparison to the other MLPA probes. One could sometimes spot the reduced peak height of this probe relative to the other probes from an MLPA peak height profile. This probe was discarded from MLPA analysis (in control samples) in certain cases to prevent incorrect statistical analysis when normalising MLPA ratios for copy number. This has not been the first time that a probe within an MLPA probe has been reported to show reduced amplification relative to the other probes in the kit (personal communication with other labs).

4.5.1 Using control samples to normalise MLPA ratios

It is important that controls from healthy individuals are thoroughly assessed before they are used as controls to normalise MLPA ratios. These samples should not be degraded. If control samples showed severe waning off towards the end it will affect normalisation, these controls should be not be used.

The manufacturers of the MLPA kit recommend that 3 control samples be used for every MLPA run but more control samples are statistically better for normalisation (Insert SALSA MLPA probemix P335-B1 ALL-IKZF1, MRC Holland). In this study 5 control samples were used for normalisation. Control samples derived from different ethnic groups did not affect normalisation in this study.

It was attempted to use control samples from different MLPA runs to normalise patient samples, however there is inter run variation between each MLPA run. The peak heights obtained from healthy control samples from two different MLPA runs could not be used to

normalise MLPA ratios. This means that each MLPA run has to use healthy control samples. From this study it would be recommended that each MLPA run should include 5 healthy controls and all of them should have normal copy number for all MLPA targets after the run. These control samples can be run together with 20-25 patient samples so that MLPA reagents do not get wasted.

Chapter 5. Conclusion and future studies

In this study *IKZF1* deletions were detected in ALL patients at presentation and at relapse. MLPA has been shown to be a reliable technique to detect microdeletions such as *IKZF1*, although MLPA ratios can be unreliable if reference genes are involved in aneuploidy. However, *IKZF1* was mostly detected in Ph+ ALL patients and these patients usually do not have aneuploidy as shown in this study and other studies that have detected *IKZF1* deletions (Mullighan et al, 2008; Iacobucci et al, 2009). MLPA is also a good alternative to use over other techniques because of its waiting time and cost and it would be recommended that it be used in a diagnostic setting.

It was also shown that *IKZF1* deletions could have had a prognostic impact in certain cases over other markers. Thus, *IKZF1* deletions should be screened for in newly diagnosed ALL patients; this would assist in stratifying ALL patients into different risk group. Consequently, identification of an *IKZF1* deletion might assist doctors to intensify current treatment with a view to prevent relapse in these patients and improve EFS.

Currently there is no therapy that targets *IKZF1* deletions. Furthermore, the prognostic impact of the different types of *IKZF1* deletions, whether the entire gene or specific exons of *IKZF1* is deleted has not been determined (Kastner et al, 2013).

It is worthy to note that point mutations which lead to premature stop codons have recently been detected in *IKZF1* in genomic DNA and mRNA. Recently, point mutations were most frequently detected when using deep sequencing in ALL patients (Roberts et al, 2012). Studies that did not detect any point mutations in large cohorts of ALL patients (Mulligan et al, 2008; Iacobucci et al, 2009 & Yang et al, 2008) did not use deep sequencing. This implies that deep sequencing should be used in future studies in an attempt to detect point mutations of *IKZF1*.

It would also be worthy to determine the mechanisms that lead to *IKZF1* deletions. It has been sought to determine the strong association between *IKZF1* deletions and the presence of the Ph chromosome in ALL patients. The JAK-STAT pathway could be a common link between these two aberrations. It is known that *BCR-ABL1* from the Ph chromosome

activates the JAK-STAT pathway and *IKZF1* deletions have been found more frequently in leukemia patients with activated JAK-STAT. It is thought that *IKZF1* may play a role in repressing this pathway (Harvey et al, 2010 & Roberts et al, 2012 & Chen et al, 2012). However, the mechanism in which *IKZF1* represses the JAK-STAT pathway has yet to be determined.

Appendices

Appendix A

M10

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG
Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Mr Mishalan Moodley

CLEARANCE CERTIFICATE	M10M101172
PROJECT	Optimizing a Genetic Test that Detects Ikaros Gene Aberrant Expression/Intragenic Deletion in Acute Lymphoblastic Leukemic Patients
INVESTIGATORS	Mr Mishalan Moodley.
DEPARTMENT	Molecular Medicine & Haematology
DATE CONSIDERED	26/11/2011
DECISION OF THE COMMITTEE*	Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE 14/02/2011 CHAIRPERSON .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable
cc: Supervisor : Dr P Willem

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...

Appendix B

Appendix B1

Nucleic acid extractions

DNA extraction from whole blood or bone marrow specimen using the QIAamp DNAeasy blood Mini kit (QIAGEN)

Blood or Bone marrow sample collection*:*

Step 1: The participant delivers blood or bone marrow specimen into a Paxgene vial.

Step 2: The vial is closed. The vial contains an enzyme that preserves and keeps RNA stable. The Paxgene vial can be stored at -70°C for up to two years. From the same vial 300ul of specimen was aliquoted into a separate 1ml eppendorf tube to use for DNA extraction.

Protocol for DNA extraction using the QIAamp DNAeasy blood Mini kit (QIAGEN)

- 20 µl proteinase K was pipetted into a 1.5ml eppendorf tube with 300µl blood or bone marrow specimen. The volume was adjusted to 520 µl with Phosphate buffer saline (PBS)
- Added 200µl buffer AL , this was mixed thoroughly by vortexing for 5-10 seconds until mixture is homogenous. The specimen and AL buffer was then incubated at 56°C for 10 min.
- 200µl ethanol (100%) was then added to the sample and mixed thoroughly by vortexing for 5-10 seconds.
- This mixture was then aliquoted into the DNeasy Mini spin column (supplied with kit) and placed in a 2 ml collection tube.
- The collection tube was then centrifuged at 6000 Relative Centrifugal Force (rcf) for 1 minute and the flow through was discarded.
- The DNeasy Mini spin column is placed into a new 2 ml collection tube and 500ul of buffer AW1 wasadded, this was then centrifuged for 1 minute at 6000 rcf.
- The DNeasy Mini spin column is placed into a new 2 ml collection tube and 500ul of buffer AW2 wasadded, this was then centrifuged for 3 minutes at 20 000 rcf to dry the DNeasy membrane.
- The DNeasy Mini spin column was placed in a clean 1.5 ml eppendorf tube and 200µl Buffer AEwas added to the DNeasy membrane.
- The column was incubated at room temperature for 1 min, and then centrifuge for 1 min at 6000 rcf to elute.

Appendix B2

RNA extraction from whole blood or bone marrow specimen using the Paxgene blood Mini kit (Preanalytix)

Protocol for RNA extraction using the Paxgene blood Mini kit (Preanalytix)

- PAXgene Blood RNA Tubes were centrifuged for 10 minutes at 5000 rcf using a swing-out rotor (Eppendorf 5840R).
- The supernatant was removed and discarded. 5 ml RNase-free water was added to the pellet and the tube was closed.
- The pellet was resuspended by vortexing and then centrifuged for 10 min at 5000 rcf.
- The supernatant was discarded and the pellet was resuspendin 360 µl buffer BR1 by vortexing. The sample was placed into a 1.5 ml eppendorf tube and 300ul of buffer BR2 and 40ul of proteinase K were added.
- This was then mixed by vortexing, and then incubated for 10 minutes at 55°C using a shaker–incubator (Eppendorf) with a heating block.
- The tube were then centrifuged for 3 minutes at 22 000 rcf and the supernatant was transferred into a new 1.5 ml eppendorf tube.
- 350 µl of 100% ethanol was added and mixed by vortexing, then centrifuged for 5 seconds at 100 rcf.
- 700 µl sample was added to the Paxgene column resting in a 2 ml tube
- This was centrifuged for 1 minute at 8000 rcf.
- The PAXgene column was then placed in a new 2 ml tube.
- 700 µl of buffer BR3 was added to the Paxgene column and centrifuged for 1 minute at 8000 rcf.
- The Paxgene column in then placed into a new 2 ml tube, and the flow through from the old 2 ml tube was discarded.
- 500 µl of buffer BR4 was added to the Paxgene column, and centrifuged for 1 minute at 8000 rcf.
- The Paxgene column was then placed into a new 2 ml tube, and the flow through from the old 2 ml tube was discarded.
- 500 µl of buffer BR4 was again added to the Paxgene column and centrifuged for 3 minutes at maximum speed to dry the Paxgene column membrane.
- The Paxgene column was placed into a 1.5 ml elution tube and 40 µl buffer BR5 was added directly onto the Paxgene columnmembrane. The RNA was eluted by centrifuging the tube for 1 minute at 8000 rcf.

Appendix B3

Purification of PCR product from agarose gel using the Bioflux extraction kit (Hangzou Bioer Technology)

Protocol for DNA extraction

- DNA fragment excised from the agarose gel under UV light using sterile scalpel.
- Weigh excised DNA fragment and place it into a clean 1ml Eppendorf tube.
- Add 300µl extraction buffer for each 100 mg agarose gel to lyse the gel.
- Incubate samples at 50°C for 10 minutes, vortexing every 2 minutes until the gel is completely dissolved.
- Load the samples to a spin column, centrifuge at 6000rcf for 1 minute. Discard the flow through.
- Add 500µl extraction buffer to spin column, centrifuge for 1 minute at 12 000rcf. Discard the flow-through.
- Add 750µl wash buffer to spin column, centrifuge for 1 minute at 12 000rcf. Discard the flow-through.
- Centrifuge for an additional 1 minute at 12 000rcf and transfer column to 1.5ml Eppendorf collection tube.
- Add 30µl water to spin column and incubate for 1 minute.
- Centrifuge for 1 minute at 12 000rcf to elute the sample.

Appendix C

Reagents used in the study

Appendix C1

2% Agarose Gel

- 2g agarose mixed with 100ml 1xTBE and boil until agarose is completely dissolved. Allow to cool for 25 minutes. Add 10µl (10 000X) Gel Red to the solution. Pour into gel casting tray to set.

10X TBE (Tris/Borate/EDTA)

- Dissolve 108 g of Tris base and 55 g boric acid in 800 ml of dH2O. Add 40 ml of 0.5 M EDTA (pH 8.0) and make up to 1 litre. The pH of this stock should be 8.3

1X TBE

- Mix 100 ml of 10X TBE stock and 900 ml of dH2O to make 1 litre of 1X TBE solution

Approximately 5µl of DNA, RNA or PCR product was mixed with 3µl of ficoll dye and were subsequently loaded on a 2% agarose gel with a molecular weight ladder and resolved at 90 volts for 60 minutes in a gel tank with TBE buffer

Appendix C2

FISH Wash buffers

50% formamide

- Mix 20ml 20x SSC with 80ml distilled water and 100ml formamide. The pH should be made up 7 with concentrated HCL
-

20x SSC

- Mix 3M NaCl (SMM Chemicals) with 0.3M sodium citrate (SMM Chemicals) and adjust the pH to 7 with concentrated HCL. Autoclave and store at room temperature

DAPI (4', 6-diamino-2-phenylindole)

- Add 0,2mg/ml DAPI stock solution (Serva) with 2xSSC

2x SSC

Mix 5ml of 20x SSC with 45ml distilled water

Appendix D

Appendix D1

cDNA synthesis

Table 1.RT-PCR components and thermocycling conditions used for making cDNA from 1ug RNA.

RT-PCR reagents	Concentration (1X)	Thermocycling conditions
MMLV-RT	200 U	25°C – 10 minutes
RT buffer	5X	42°C – 45 minutes
dNTP	1mM	99°C – 3 minutes
DTT	10mM	4°C –hold (store)
Random primers	25um/µl	
RNAase OUT ribonuclease inhibitor	40 U/µl	
dH_2O	2ul	

Appendix D2

PCR of *GAPDH*

Table 2.PCR components for amplification of *GAPDH*.

PCR reagents	Concentration (1X)
Forward Primer 5’ cccttcattgacctcaactacatg 3’	10pmol/µl
Forward Primer 5’ catgccagtgagcttcccgttcag 3’	10pmol/µl
dNTP	100 mM
PCR Buffer	10 X
$MgCl_2$	10mM
dH_2O	9.4 µl

Thermocycling conditions: 95°C – 5 minutes, (94°C for 30 seconds, 56°C for 1 minute, 72°C for 1 minute) X 30 cycles, 72°C for 5 minutes, 4°C hold (store).

Appendix E

Table3.SALSAMLPAP335-B1ALL-IKZF1probemix (Taken from MRC Holland MLPA insert package)

Length (nt)	SALSAMLPAprobe	Chromosomalposition location reference region (hg18)
64-70-76-82	Q-fragments: DNAquantity;onlyvisiblewithless than100ngsampleDNA	
88-92-96	D-fragments:Lowsignalof88 or96ntfragmentindicatesincompletedenaturation	
100	X-fragment:SpecificfortheXchromosome	
105	Y-fragment:Specific fortheYchromosome	
118*¬	**ZFY**probeS0135-L16766	Yp11.31 Y-002.889246
124*	Referenceprobe15370-L13762	7q11.23 07-075.448402
130*	Referenceprobe13867-L15385	16p13.2 16-008.765461
136	**CRLF2**probe13889-L15427	Xp22.33-PAR region X-001.281274
142	**IKZF1** probe13872-L15390	7p12.2 07-050.417749
148	**SHOX-AREA**probe05648-L06218	Xp22.33-PAR region X-000.770580
153¬	**JAK2** probe07452-L09964	9p24.1 09-005.112998
160¥	**PAX5**probe12501-L22020	9p13.2 09-037.024272
166*	**CSF2RA**probe13892-L16221	Xp22.33-PAR region X-001.374311
172¥	**PAX5**probe14647-L15394	9p13.2 09-037.010669
178*	**BTG1-AREA**probe18021-L22630	12q21.33 12-090.657641
184*	**P2RY8**probe17837-L15740	Xp22.33-PAR region X-001.545095
190*	Referenceprobe06941-L06521	11q12.3 11-061.483953
196*	**ETV6**probe17838-L22035	12p13.2 12-011.935935
202*	**PAX5**probe17839-L22036	9p13.2 09-036.956555
208	**IKZF1** probe14056-L15654	7p12.2 07-050.329206
214*	Referenceprobe13265-L15166	1p21.1 01-103.203651
220	**RB1** probe01782-L01346	13q14.2 13-047.821213
226	**EBF1**probe12509-L13559	5q33.3 05-158.459187
232	**BTG1**probe12542-L15913	12q22 12-091.063286
238	**CDKN2B**probe10337-L15914	9p21.3 09-021.995813
244	**ETV6**probe13874-L17160	12p13.2 12-011.883341
251	**CDKN2A**probe10333-L15916	9p21.3 09-021.964958
258*	Referenceprobe04534-L22019	2q24.3 02-166.606098
263	**IKZF1** probe13873-L15917	7p12.2 07-050.411797
269±	**IKZF1** probe13877-L15918	7p12.2 07-050.315024
274*	**PAX5**probe17840-L22037	9p13.2 09-036.871952
282	**PAX5**probe13870-L15920	9p13.2 09-036.992699
288*	**IKZF1** probe17109-L20256	7p12.2 07-050.436281
295*	Referenceprobe10435-L22110	9q34.3 09-136.850306
301±	**ETV6**probe14058-L15656	12p13.2 12-011.694211
308*	**CDKN2A**probe17814-L22631	9p21.3 09-021.957858
315¥	**RB1** probe01789-L22025	13q14.2 13-047.851428
324¥	Referenceprobe03918-L20270	15q21.1 15-046.584530
330¥	**BTG1**probe12553-L22632	12q22 12-091.061700
337*	**PAX5**probe17841-L22038	9p13.2 09-036.913345
343¥	**IKZF1** probe13869-L22296	7p12.2 07-050.426948
350¥	**IL3RA**probe13907-L22294	Xp22.33-PAR region X-001.415699
358¥	**RB1** probe01792-L22295	13q14.2 13-047.928372
364*	Referenceprobe14675-L16327	3p25.3 03-010.142307
371¥	**EBF1**probe14059-L22017	5q33.3 05-158.137017
379*	**IKZF1** probe15427-L22113	7p12.2 07-050.337845
386¥±	**ETV6**probe14060-L22634	12p13.2 12-011.694327
393*	**PAX5**probe17842-L22633	9p13.2 09-036.829905
400¥	**ETV6**probe13875-L22014	12p13.2 12-011.796696
409*	**BTG1AREA**probe18022-L22363	12q21.33 12-091.005964
418	**RB1** probe01797-L01360	13q14.2 13-047.945205
427*	Referenceprobe07924-L22013	20p13 20-001.867470
436¥	**EBF1**probe13868-L22053	5q33.3 05-158.071810
445	**RB1** probe01799-L01362	13q14.2 13-047.949488

454*	Referenceprobe11533-L20700	19q13.43	19-061.386348
462 ¥	EBF1 probe 12527-L22010	5q33.3	05-158.057828
470 ¥	IKZF1 probe 14061-L22112	7p12.2	07-050.422502
477 *	Reference probe 14846-L22111	3q12.2	03-099.782925
484 ¥	ETV6 probe 13871-L22009	12p13.2	12-011.913644
493 *	Reference probe 15203-L16978	3p12.2	03-081.774613
503 *	Reference probe 09870-L19465	2p16.1	02-061.126370

*NewinversionB1(fromlot B1-0412onwards).
¥ChangedinversionB1(fromlotB1-0412onwards).Smallchangeinlength,no changeinsequencedetected.
±Thisprobeislocatedwithin, orcloseto,averystrongCpGisland.Alowsignalofthisprobecan bedueto incompletesampleDNAdenaturation, e.g.dueto thepresence ofsaltinthesample DNA.
¬Flankingprobe:Includedtofacilitatethedeterminationoftheextentofadeletion/duplication. Copynumber alterationsofflankingandreferenceprobesareunlikelyto berelatedto theconditiontested.

Appendix F

Software commands for MLPA analysis

Appendix F1

Genemapper 4.1 commands

- Open Genemapper, open a “new project” and select AFLP
- Open Panel manager
- In Panel manager create a new kit
- Create a new bin set
- Create a new panel
- Select a new panel and create a new marker, in the marker columns select marker name, min peak size, analysis type
- Add reference data
- Change analysis method to “MLPA default”, select the panel created above and select appropriate size standard
- Analyse data
- Go back to “Add reference data” in Panel manager and choose the current project with samples
- Add the samples you would like to use as reference samples (A-J)
- Select the marker created in your new panel
- Select a reference sample from the list in the bottom left window. Peaks will appear in the window on the right
- Creates bins for each peak from the “Bins” drop down menu. Click on each peak and name it accordingly to the P335 MLPA kit package insert
- Exit Panel manager and analyse the samples using the new bin set
- Click on “Genotype” to view the peak height of each peak
- Select “Display plots”, all the selected runs will be displayed underneath each other
- Deselect the other labels (except FAM (blue)). Select “Sizing table”. Sizing table should only display columns for “Dye/sample peak, Sample file name, Size, Seight, Area, Data point”
- Select “File” in the drop down menu and “Export table”.
- The exported table can now be imported to Microsoft Excel® and Coffalyser as a “Genemapper txt” file.

Appendix F2

Coffalyser 9.4 commands

- Start-up Coffalyser and enable macros
- Select “P335-B1 MLPA mix” under “Load mix”
- Click “Start MLPA analysis”
- Select “import Genemapper txt file (multiple runs)”
- In the “data filtering window”, select the healthy control samples and click on “>>”
- Select “Save all mix changes” and click on the play button
- After data is filtered then check the QC by checking if “ok” is displayed under the “number of probes, Q fragments and D fragments” columns
- Select “analysis options” and choose “median of all imported samples”
- Select “LMS” method, click the play button
- Results page will open, ratios between 0.7-1.3 will be highlighted in a blue colour, deletions will be highlighted in a red colour (<0.7) and duplications in a green colour (>1.3)
- Select “create and save all sample reports” to save the MLPA run analysis so it can be viewed at a later stage

Appendix G

Entire *IKZF1* sequence (Ik1)

Exon 2

AUAACCUGAGGACCAUGGAUGCUGAUGAGGGUCAAGACAUGUCCCAAGUUUCAG

Exon 3

GGAAGGAAAGCCCCCCUGUAAGCGAUACUCCAGAUGAGGGCGAUGAGCCCAUGCCGAUCCCCGAGGACC
UCUCCACCACCUCGGGAGGACAGCAAAGCUCCAAGAGUGACAGAGUCGUGG

Exon 4:

CCAGUAAUGUUAAAGUAGAGACUCAGAGUGAUGAAGAGAAUGGGCGUGCCUGUGAAAUGAAUGGGGA
AGAAUGUGCGGAGGAUUUACGAAUGCUUGAUGCCUCGGGAGAGAAAAUGAAUGGCUCCCACAGGGACC
AAGGCAGCUCGGCUUUGUCGGGAGUUGGAGGCAUUCGACUUCCUAACGGAAAACUAAAGUGUGAUAUC
UGUGGGAUCAUUUGCAUCGGGCCCAAUGUGCUCAUGGUUCACAAAAGAAGCCACACUG

Exon 5:

GAGAACGGCCCUUCCAGUGCAAUCAGUGCGGGGCCUCAUUCACCCAGAAGGGCAACCUGCUCCGGCACA
UCAAGCUGCAUUCCGGGGAGAAGCCCUUCAAAUGCCACCUCUGCAACUACGCCUGCCGCCGGAGGGACG
CCCUCACUGGCCACCUGAGGACGCACUCCG

Exon 6:

UUGGUAAACCUCACAAAUGUGGAUAUUGUGGCCGAAGCUAUAAACAGCGAAGCUCUUUAGAGGAACAU
AAAGAGCGCUGCCACAACUACUUGGAAAGCAUGGGCCUUCCGGGCACACUGUACCCAG

Exon 7:

UCAUUAAAGAAGAAACUAAUCACAGUGAAAUGGCAGAAGACCUGUGCAAGAUAGGAUCAGAGAGAUCU
CUCGUGCUGGACAGACUAGCAAGUAACGUCGCCAAACGUAAGAGCUCUAUGCCUCAGAAAUUUCUUG’

Exon 8:

GGGACAAGGGCCUGUCCGACACGCCCUACGACAGCAGCGCCAGCUACGAGAAGGAGAACGAAAUGAUGA
AGUCCCACGUGAUGGACCAAGCCAUCAACAACGCCAUCAACUACCUGGGGGCCGAGUCCCUGCGCCCGCU
GGUGCAGACGCCCCCGGGCGGUUCCGAGGUGGUCCCGGUCAUCAGCCCGAUGUACCAGCUGCACAAGCC
GCUCGCGGAGGGCACCCCGCGCUCCAACCACUCGGCCCAGGACAGCGCCGUGGAGAACCUGCUGCUGCUC
UCCAAGGCCAAGUUGGUGCCCUCGGAGCGCGAGGCGUCCCCGAGCAACAGCUGCCAAGACUCCACGGACA
CCGAGAGCAACAACGAGGAGCAGCGCAGCGGUCUCAUCUACCUGACCAACCACAUCGCCCCGCACGCGCG
CAACGGGCUGUCGCUCAAGGAGGAGCACCGCGCCUACGACCUGCUGCGCGCCGCCUCCGAGAACUCGCAG
GACGCGCUCCGCGUGGUCAGCACCAGCGGGGAGCAGAUGAAGGUGUACAAGUGCGAACACUGCCGGGU
GCUCUUCCUGGAUCACGUCAUGUACACCAUCCACAUGGGCUGCCACGGCUUCCGUGAUCCUUUUGAGUG
CAACAUGUGCGGCUACCACAGCCAGGACCGGUACGAGUUCUCGUCGCACAUAACGCGAGGGGAGCACCG
CUUCCACAUGAGCUAA

Reference List

Abdul-Hamid, G. 2011. Classification of Acute Leukemia, Acute Leukemia - The Scientist's Perspective and Challenge, Prof. Mariastefania Antica (Ed.), ISBN: 978-953-307-553-2, InTech, DOI: 10.5772/19848. Available from: http://www.intechopen.com/books/acute-leukemia-the-scientist-s-perspective-and-challenge/classification-of-acute-leukemia. [Accessed 13/7/2013].

Argiropoulos, B., & Humphries, RK. 2007. Hox genes in hematopoiesis and leukemogenesis. *Oncogene* 26:6766-6776.

Armstrong, SA., Staunton, JE., Silverman, LB., Pieters, R., den Boer, ML., Minden, MD., Sallan, SE., Lander, ES., Golub, TR., & Korsmeyer, SJ. 2002. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nature Genetics* 1:41-7.

Aspland, SE., Bendall, HH., & Murre, C. 2001. The role of E2A-PBX1 in leukemogenesis.*Oncogene* 20:5708-5717.

Barber, RD., Harmer, DW., Coleman, RA., & Clark BJ. 2005. GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiological genomics* 21(3):389-395.

Bennett, JM., Catovsky, D., Daniel, MT., Flandrin, G., Galton, DA., Gralnick, HR., Sultan, C. 1976. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *British journal of haematology* 33(4):451-458.

Bennett, JM., Catovsky, D., Daniel, MT., Flandrin, G., Galton, DA., Gralnick, HR., Sultan, C. 1985. Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. *Annals of internal Medicine* 103(4):620-625.

Bishop, R. 2010. Applications of fluorescence in situ hybridization (FISH) in detecting genetic aberrations of medical significance. *Bioscience Horizons* 3: 85-95.

Boerma, EJG., Siebert, R., Kluin, PM., & Baudis, M. 2009. Translocations involving 8q24 in Burkitt Lymphoma and other malignant lymphomas – a historical revew of cytogenetics in the light of today's knowledge. *Leukemia* 23: 225-234.

Boormans, EM., Birnie, E., Oepkes, D., Galjaard, RJ., Schuring-Blom, GH., &van Lith, JM. 2010. Comparison of multiplex ligation-dependent probe amplification and karyotyping in prenatal diagnosis. *Obstetrics and gynecology* 115(2 Pt 1): 297-303.

Caye, A., Beldjord, K., Mass-Malo, K., Drunat, S., Soulier, J., Gandemer, V., Baruchel, A., Bertrand, Y., Cavé, H., &Clappier, E. 2013. Breakpoint-specific multiplex polymerase chain reaction allows the detection of IKZF1 intragenic deletions and minimal residual disease monitoring in B-cell precursor acute lymphoblastic leukemia. *Haematologica* 98 (4):597-601.

Combaret, V., Iacono, I., Bréjon, S., Schleiermacher, G., Pierron, G., Couturier, J., Bergeron, C., & Blay., JY. 2012. Analysis of genomic alterations in neuroblastoma by multiplex ligation-dependent probe amplification and array comparative genomic hybridization: a comparison of results. *Cancer Genetics* 205 (12): 657-664.

Conter, V., Bartram, CR., Valsecchi, MG., Schrauder, A., Panzer-Grümayer, R., Möricke, A., Aricò, M., Zimmermann, M., Mann, G., De Rossi, G., Stanulla, M., Locatelli, F., Basso, G., Niggli, F., Barisone, E., Henze, G., Ludwig, WD., Haas, OA., Cazzaniga, G., Koehler, R., Silvestri, D., Bradtke, J., Parasole, R., Beier, R., van Dongen, JJ., Biondi, A., & Schrappe, M. 2010. Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study. *Blood* 115(16): 3206-3214.

Daly. 2010. Understanding Acute Lymphoblastic Leukaemia (ALL) in Children: A booklet, published by Leukemia Foundation. Queensland. Available from: https://www.google.co.za/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&ved=0CDsQFjAA&url=http%3A%2F%2Fwww.leukaemia.com%2Fweb%2Fresources%2Fpdf%2F2010_ALL_Booklet.pdf&ei=3fsJUpuCIuw7Qbny4DYDA&usg=AFQjCNFy7BBIaoa0mKbQMAQyPoKXeFI_MQ&sig2=x5ItKI7XCzcgUF5bK-CDSw&bvm=bv.50500085,d.ZWU. [Accessed 13/7/2013].

Druker, BJ.,& Lydon, NB. 2000. Lessons learned from the development of an abl tyrosine kinase inhibitor for chronic myelogenous leukemia. *Journal of Clinical Investigation* 105:3-7.

Dupuis, A., Gaub, MP., Legrain, M., Dreno u, B., Mauvieux, L., Lutz, P., Herbrecht, R., Chan., S., & Kastner, P. 2012. Biclonal and biallelic deletions occur in 20% of BALL cases with IKZF1 mutations. *Leukemia* 27(2): 503-507.

Eisenman, RN. 2001. Deconstructing Myc. *Genes & Development* 15:2023-2030.

Forestier, E., Izraeli, S., Beverloo, B., Haas, O., Pession, A., Michalová, K., Stark, B., Harrison, CJ., Teigler-Schlegel, A., Johansson, B. 2008. Cytogenetic features of acute lymphoblastic and myeloid leukemias in pediatric patients with Down syndrome: an iBFM-SG study. *Blood* 111(3):1575-1583.

Garcia M., Jemal, A., Ward, EM., Center, MM., Hao, Y., Siegel, RL., Thun, MJ. In: Global Cancer Facts & Figures 2007. Atlanta, GA: American Cancer Society, 2007.

Gaynon, PS., Qu, RP.,& Chappell, RJ. 1998. Survival after relapse in childhood acute lymphoblastic leukemia: Impact of site and time to first relapse—The Children's Cancer Group Experience. *Cancer* 82: 1387–95.

Georgopoulos, K., Winandy, S., & Avitahl, N. 1997. The role of the Ikaros gene in lymphocyte development and homeostasis.*The Annual Review of Immunology* 15:155–76.

Harvey, RC., Mullighan, CG., Chen, IM., Wharton, W., Mikhail, FM., Carroll, AJ., Kang, H., Liu, W., Dobbin, KK., Smith, MA., Carroll, WL., Devidas, M., Bowman, WP., Camitta, BM., Reaman, GH., Hunger, SP., Downing, JR., & Willman, CL. 2010. Rearrangement of CRLF2 is associated with mutation of JAK kinases, alteration of IKZF1, Hispanic/Latino ethnicity, and a poor outcome in pediatric B-progenitor acute lymphoblastic leukemia. *Blood* 115(26):5312-5321.

Hills, A., Ahn, JW., Donaghue, C., Thomas, H., Mann, K., & Ogilvie, CM. 2010. MLPA for confirmation of array CGH results and determination of inheritance.*Molecular cytogenetics* 3(19):1-7.

Hosokawa, Y., Maeda, Y., Ichinohasama, R., Miura, I., Taniwaki, M., & Seto M. 2000. The Ikaros gene, a central regulator of lymphoid differentiation, fuses to the BCL6 gene as a result of t(3;7)(q27;p12)translocation in a patient with diffuse large B-cell lymphoma. *Blood* 95(8): 2719-2721.

Hunger, SP., Raetz, EA., Loh, ML., & Mullighan, CG. 2011. Improving Outcomes for High-Risk ALL: Translating New Discoveries Into Clinical Care. *Pediatric blood & cancer* 56: 984-93.

Iacobucci, I., Lonetti, A., Cilloni, D., Messa, F., Ferrari, A., Zuntini, R., Ferrari, S., Ottaviani, E., Arruga, F., Paolini, S., Papayannidis, C., Piccaluga, PP., Soverini, S., Saglio, G., Pane, F., Baruzzi, A., Vignetti, M., Berton, G., Vitale, A., Chiaretti, S., Müschen, M., Foà, R., Baccarani, M., & Martinelli, G. 2008a. Identification of different Ikaros cDNA transcripts in Philadelphia-positive adult acute lymphoblastic leukemia by a high-throughput capillary electrophoresis sizing method. Haematologica 93(12):1814-1821.

Iacobucci, I., Lonetti, A., Messa, F., Cilloni, D., Arruga, F., Ottaviani, E., Paolini, S., Papayannidis, C., Piccaluga, PP., Giannoulia, P., Soverini, S., Amabile, M., Poerio, A., Saglio, G., Pane, F., Berton, G., Baruzzi, A., Vitale, A., Chiaretti, S., Perini, G., Foà, R., Baccarani, M., &Martinelli G. 2008b. Expression of spliced Ikaros oncogenic isoforms in Philadelphia-positive acute lymphoblastic leukemia patients treated with tyrosine kinase inhibitors: implications for a new mechanism of resistance. *Blood* 112:3847-3856.

Iacobucci, I., Storlazzi, CT., Cilloni, D., Lonetti, A., Ottaviani, E., Soverini, S., Astolfi, A., Chiaretti, S., Vitale, A., Messa, F., Impera, L., Baldazzi, C., D'Addabbo, P., Papayannidis, C., Lonoce, A., Colarossi, S., Vignetti, M, Piccaluga, PP., Paolini, S., Russo, D., Pane, F., Saglio, G., Baccarani, M, Foa`, R., & Martinelli., G. 2009. Identification and molecular characterization of recurrent genomic deletions on 7p12 in the IKZF1 gene in a large cohort of BCR-ABL1–positive acute lymphoblastic leukemia patients: on behalf of Gruppo Italiano Malattie Ematologiche dell'Adulto Acute LeukemiaWorking Party (GIMEMAALWP). Blood 114 (10): 2159-2167.

Iacobucci, I., Papayannidis, C., Lonetti, A., Ferrari, A., Baccarani, M., & Martinelli, G. 2012. Cytogenetic and molecular predictors of outcome in acute lymphocytic leukemia: recent developments. *Current hematologic malignancy reports* 7(2): 133-143.

Jankovic, M., Robbiani, DF., Dorsett, Y., Eisenreich, T., Xu, Y., Tarakhovsky, A., Nussenzweig, A., & Nussenzweig, MC. 2009. Role of the translocation partner in protection against AID – dependent chromosomal translocations. *Immunology* 107 (1): 187-192.

Jennings, CD., & Foon, KA. 1997. Recent advances in Flow Cytometry: Application to the Diagnosis of Hematologic Malignancy. *Blood* 90:2863-2892.

Jennings, LJ., Yu, M., Fitzpatrick, C., & Smith, FA. 2011. Validation of multiplex ligation-dependent probe amplification for confirmation of array comparative genomic hybridization. *Diagnostic Molecular Pathology* 20(3): 166-174.

Jeuken, J, Sijben, A., Alenda, C., Rijntjes, J., Dekkers, M., Boots-Sprenger, S., McLendon, R., & Wesseling, P. 2009. Robust Detection of EGFR Copy Number Changes and EGFR Variant III: Technical Aspects and Relevance for Glioma Diagnostics. *Brain Pathology* 19:661-671.

Kastner, P., Dupuis, A., Gaub, MP., Herbrecht, R., Lutz, P., & Chan, S. 2013. Function of Ikaros as a tumor suppressor in B cell acute lymphoblastic leukemia. *American journal of blood research* 3(1): 1-13.

Kruhøffer, M., Dyrskjøt, L., Voss, T., Lindberg, RL., Wyrich, R., Thykjaer, T., & Orntoft, TF. 2007. Isolation of microarray-grade total RNA, microRNA, and DNA from a single PAXgene blood RNA tube. *The Journal of molecular diagnostics: JMD* (4):452-458.

Kuiper, RP., Waanders, E., van der Velden, VHJ., van Reijmersdal, SV., Venkatachalam, R., Scheijen, B., Sonneveld, E., van Dongen, JJM., Veerman, AJP., van Leeuwen, FN., van Kessel, AG., & Hoogerbrugge, PM. 2010. IKZF1 deletions predict relapse in uniformly treated pediatric precursor B-ALL. *Leukemia* 7: 1258-1264.

Kurzrock, R., Kantarjian, HM., Druker, BJ., &Talpaz, M. 2003. Philadelphia Chromosome-Positive Leukemias : From Basic Mechanisms to Molecular Therapeutics. *Annals of Internal Medicine* 138: 819-830.

Lake, SL., Damato, BE., Dopierala, J., Baudo, MM., Taktak, AF., & Coupland, SE. 2011. Multiplex Ligation-Dependent Probe Amplification Analysis of Uveal Melanoma with

Extraocular Extension Demonstrates Heterogeneity of Gross Chromosomal Abnormalities. *Investigative Ophthalmology & Visual Science* 52(8): 5559-5564.

Mangs AH., & Morris, BJ. 2007. The Human Pseudoautosomal Region (PAR): Origin, Function and Future. *Current Genomics* 8(2):129-136.

Martinelli, G., Iacobucci, I., Storlazzi, CT., Vignetti, M., Paoloni, F., Cilloni, D., Soverini, S., Vitale, A., Chiaretti, S., Cimino, G., Papayannidis, C., Paolini, S., Elia, L., Fazi, P., Meloni, G., Amadori, S., Saglio, G., Pane, F., Baccarani, M., & Foa`, R. 2009. IKZF1 (Ikaros) Deletions in BCR-ABL1-Positive Acute Lymphoblastic Leukemia Are Associated With Short Disease-Free Survival and High Rate of Cumulative Incidence of Relapse: A GIMEMA AL WP Report. *Journal of Clinical Oncology* 31: 5202-5207.

McConaughy, BL., Laird, CD., & McCarthy BJ. 1969. Nucleic acid reassociation in formamide. *Biochemistry* 8(8): 3289-3295.

Mi, JQ., Wang, X., Yao, Y., Lu, HJ., Jiang, XX., Zhou, JF., Wang, JH., Jiao, B., Shen, SH., Tang, JY., Gu, LJ., Jiang, H., Ma, LY., Hao, SG., Chen, FY., Xiong, SM., Shen, ZX., Chen, Z., Chen, B., & Chen, SJ. 2012. Newly diagnosed acute lymphoblastic leukemia in China (II): prognosis related to genetic abnormalities in a series of 1091 cases. *Leukemia* 26(7): 1507-1516.

Moorman, AV., Schwab, C., Ensor, HM., Russell, LJ., Morrison, H., Jones, L., Masic, D., Patel, B., Rowe, JM., Tallman, M., Goldstone, AH., Fielding, AK., & Harrison, CJ. 2012. IGH@ translocations, CRLF2 deregulation, and microdeletions in adolescents and adults with acute lymphoblastic leukemia. *Journal of Clinical Oncology* 30(25): 3100-3108.

Mullighan, CG., Goorha, S., Radtke, I., Miller, CB., Coustan-Smith, E., Dalton, JD., Girtman, K., Mathew, S., Ma, J., Pounds, SB., Su, X., Pui, CH., Relling, MV., Evans, WE., Shurtleff, SA., & Downing, JR. 2007. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature 446: 758-764.

Mullighan, CG., Miller, CB., Radtke,I., Phillips, LA., Dalton, J., Ma., J, White., D Hughes., TP., Le Beau, MM., Pui, CH., Relling, MV., Shurtleff, SA., & Downing, JR. 2008. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature* 453: 110-114.

Mullighan, CG., Su, X., Zhang, J., Radtke, I., Phillips, LA., Miller, CB., Ma, J., Liu, W., Cheng, C., Schulman, BA., Harvey, RC., Chen, IM., Clifford, RJ., Carroll, WL., Reaman, G., Bowman, WP., Devidas, M., Gerhard, DS., Yang, W., Relling, MV., Shurtleff, SA., Campana, D., Borowitz, MJ., Pui, CH., Smith, M., Hunger, SP., Willman, CL., & Downing, JR. 2009. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia.*The New England journal of medicine* 360(5):470-480.

Mullighan, CG. 2011. Genomic profiling of B-progenitor acute lymphoblastic leukemia. *Best Practice & Research Clinical Haematology* 24(4):489-503.

Murugan, S., Chandramohan, A., & Lakshmi, BR. 2010.Use of multiplex ligation-dependent probe amplification (MLPA) for Duchenne muscular dystrophy (DMD) gene mutation analysis.*The Indian journal of medical research* 132:303-311.

Nakase, K., Kita K., Shiku H., Tanaka I., Nasu K., Dohy H., Kyo T., Tsutani H., & Kamada N. 1996. Myeloid antigen, CD13, CD14, and/or CD33 expression is restricted to certain lymphoid neoplasms. *American Journal of Clinical Pathology* 105 (6):761-768.

Nakase, K., Ishimaru, F., Avitahl, N., Dansako, H., Matsuo, K., Fujii, K., Sezaki, N., Nakayama, H., Yano,T., Fukuda, S., Imajoh, K., Takeuchi, M., Miyata, A., Hara, M., Yasukawa, M., Takahashi, I., Taguchi, H., Matsue, K., Nakao, S., Niho, Y., Takenaka, K., Shinagawa, K., Ikeda, K., Niiya, K., & Harada, M. 2000. Dominant Negative Isoform of the Ikaros Gene in Patients with Adult B-Cell Acute Lymphoblastic Leukemia.*Cancer Research* 60:4062–4065.

Nguyen, K., Devidas, M., Cheng, S-C., La, Mei., Raetz., EA., Carroll, WL., Winick, NJ., Hunger, SP., Gaynon, PS., & Loh., ML. 2008. Factors influencing survival after relapse from acute lymphoblastic leukemia: A Children's Oncology Group study. *Leukemia* 22 (12): 2142-50.

Nishii, K., Katayama, N., & Shiku, H. 2002. Adult acute myeloid leukemia cells do not express non-functional Ikaros isoforms. *Blood* 100: 3436-3437.

Nowell, PC.,&Hungerford, DA. 1960. A minute chromosome in human chronic granulocytic leukemia. *Science* 132:1488-1501.

Olivero, S., Maroc, C., Beillard, E., Gabert, J., Nietfeld, W., Chabannon, C., &Tonnelle, C. 2000.Detection of different Ikaros isoforms in human leukaemias using real-time quantitative polymerase chain reaction. *British Journal of Haematology* 110: 826-830.

Onciu, M. 2009. Acute lymphoblastic Leukemia. *Hematology Oncology Clinics North America* 23:655-674.

Oster, SK., Ho, CS., Soucie, EL., Penn, LZ. 2002. The myc oncogene: MarvelouslY Complex. *Advances in cancer research* 84:81-154.

Payne, KJ., Nicolas, JH., Zhu, JY., Barsky, LW., & Crooks, GM. 2001. Cutting edge: predominant expression of a novel Ikaros isoform in normal human hemopoiesis. *Journal of immunology* 167(4):1867-1870.

Pui, CH., & Jeha S. 2007.New therapeutic strategies for the treatment of acute lymphoblastic leukaemia.*Nature reviews* 6:149-65.

Regier, DA., Friedman, JM., & Marra, CA. 2010.Value for money?Array genomic hybridization for diagnostic testing for genetic causes of intellectual disability. *American journal of human genetics* 86(5):765-772.

Reichmann, R. 2000. Acute Lymphoblastic Leukaemia: A Review. *Trinity Student Medical Journal* 1:26-28.

Roberts, KG., Morin, RD., Zhang, J., Hirst, M., Zhao, Y., Su, X., Chen, SC., Payne-Turner, D., Churchman, ML., Harvey, RC., Chen, X., Kasap, C., Yan, C., Becksfort, J., Finney, RP., Teachey, DT., Maude, SL., Tse, K., Moore, R., Jones, S., Mungall, K., Birol, I., Edmonson, MN., Hu, Y., Buetow, KE., Chen, IM., Carroll, WL., Wei, L., Ma, J., Kleppe, M., Levine, RL., Garcia-Manero, G., Larsen, E., Shah, NP., Devidas, M., Reaman, G., Smith, M., Paugh, SW., Evans, WE., Grupp, SA., Jeha, S., Pui, CH., Gerhard, DS., Downing, JR., Willman, CL., Loh, M., Hunger, SP., Marra, MA., & Mullighan, CG. 2012. Genetic alterations activating kinase and cytokine receptor signaling in highrisk acute lymphoblastic leukemia. Cancer Cell 22:153–166.

Ross, TS., Bernard, OA., Berger, R., & Gilliland, DG. 1998. Fusion of Huntingtin interacting protein 1 to platelet-derived growth factor beta receptor (PDGFbetaR) in chronic myelomonocytic leukemia with t(5;7)(q33;q11.2). *Blood* 91(12):4419-4426.

Saarinen-Pihkala, UM., Heilmann, C., Winiarski, J., Glomstein, A., Abrahamsson, J., Arvidson, J., Béka´ssy, AN., Forestier, E., Jonmundsson, G., Schroeder, H., Vettenranta, K., Wesenberg, F., & Go¨ran, G. 2006. Pathways Through Relapses and Deaths of Children With Acute Lymphoblastic Leukemia: Role of Allogeneic Stem-Cell Transplantation in Nordic Data. *Journal of Clinical Oncology* 24 (36): 5750-5762.

Saiki, RK, Sharf, S., Faloona, F., Mullis, KB., Horn, GT., Erlich, HA., & Arnheim,N. 1985. Enzymic amplification of ß-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia.*Science* 230(4732): 1350-1354.

Sanger, F., Nicklen, S., & Coulson, AR. 1977.DNA sequencing with chain-terminating inhibitors.*Proceedings of the National Academy of Sciences* 74(12): 5463–5467.

Schouten, JP., McElgunn, CJ., Waaijer, R., Zwijnenburg, D., Diepvens, F., & Pals, G. 2002. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic acids research 30(12): 1-13.

Schultz, KR., Pullen, DJ., Sather, HN., Shuster, JJ., Devidas, M., Borowitz, MJ., Carroll, AJ., Heerema, NA., Rubnitz, JE., Loh, ML., Raetz, EA., Winick, NJ., Hunger, SP., Carroll, WL., Gaynon, PS., & Camitta, BM. 2007. Risk- and response-based classification of childhood B-precursor acute lymphoblastic leukemia: a combined analysis of prognostic markers from the Pediatric Oncology Group (POG) and Children's Cancer Group (CCG). Blood 109(3): 926-935.

Secker-Walker, LM. 1998. General report on the European Union Concerted Action Workshop on 11q23, London, UK, May 1997. *Leukemia* 12(5): 776-778.

Stefan, DC. 2010. Epidemiology of childhood cancer and the SACCSG tumour registry. *Continuing Medical Education* 28: 317-19.

Stuppia L., Antonucci I., Palka G., & Gatta, V. 2012.Use of the MLPA Assay in the Molecular Diagnosis of Gene Copy Number Alterations in Human Genetic Diseases.*International journal of molecular sciences* 13(3): 3245-3276.

Sun, L., Crotty, ML., Sensel, M., Sather, H., Navara, C., Nachman, J., Steinherz, PG., Gaynon, PS., Seibel, N., Mao, C., Vassilev, A., Reaman, GH., & Uckun, FM. 1999a. Expression of Dominant-Negative Ikaros Isoforms in T-Cell Acute.*Clinical Cancer Research* 5(8): 2112-2120.

Sun, L., Heerema, N., Crotty, L., Wu, X., Navara, C., Vassilev, A., Sensel, M., Reaman, GH., & Uckun, FM. 1999b. Expression of dominant-negative and mutant isoforms of the antileukemic transcription factor Ikaros in infant acute lymphoblastic leukemia*. Proceedings of the National Academy of Sciences* 96(2): 680-685.

Szczepański, T., Beishuizen, A., Pongers-Willemse, MJ., Hählen, K., Van Wering, ER., Wijkhuijs, AJ., Tibbe, GJ., De Bruijn, MA., & Van Dongen., JJ. 1999. Cross-lineage T cell receptor gene rearrangements occur in more than ninety percent of childhood precursor-B acute lymphoblastic leukemias: alternative PCR targets for detection of minimal residual disease. *Leukemia* 13(2): 196-205.

Volejnikova, J., Mejstrikova, E., Dörge, P., Meissner, B., Zimmermannova, O., Svojgr, K., Stanulla, M., Cario, G., Schrappe, M., Stary, J., Hrusak, O., Trka, J., & Fronkova, E. 2013. Ikaros (IKZF1) alterations and minimal residual disease at day 15 assessed by flow cytometry predict prognosis of childhood BCR/ABL-negative acute lymphoblastic leukemia. *Pediatric blood & cancer* 60(3): 420-427.

Vorstman, JA., Jalali, GR., Rappaport, EF., Hacker, AM., Scott, C., & Emanuel, BS. 2006. MLPA: a rapid, reliable, and sensitive method for detection and analysis of abnormalities of 22q. Human mutation 27(8): 814-833.

Waanders E., van der Velden, VH., van der Schoot, CE., van Leeuwen, FN., van Reijmersdal, SV., de Haas, V., Veerman, AJ., van Kessel, AG., Hoogerbrugge, PM., Kuiper, RP., & van Dongen, JJM. 2010. Integrated use of minimal residual disease classification and IKZF1 alteration status accurately predicts 79% of relapses in pediatric acute lymphoblastic leukemia. *Leukemia* 2: 254-58.

Wang J-H., Nichogiannopoulou A, Wu, L., Sun, L., Sharpe, AH., Bigby, M., & Georgopoulos, K. 1996. Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. *Immunity* 5: 537-549.

Yagi, T., Hibi, S., Takanashi, M., Kano, G., Tabata, Y., Imamura, T., Inaba, T., Morimoto, A., Todo, S., & Imashuku, S. 2002. High frequency of Ikaros isoform 6 expression in acute myelomonocytic and monocytic leukemias: implications for up-regulation of the antiapoptotic protein Bcl-XL in leukemogenesis. *Blood* 99(4): 1350-1355.

Yang, JJ., Bhojwani, D., Yang, W., Cai, X., Stocco, G., Crews, K., Wang, J., Morrison, D., Devidas, M., Hunger, SP., Willman, CL., Raetz, EA., Pui, CH., Evans, WE., Relling, MV.,& Carroll, WL. 2008. Genome-wide copy number profiling reveals molecular evolution from diagnosis to relapse in childhood acute lymphoblastic leukemia. *Blood* 112 (10): 4178-4183.

Yap, WH., Yeoh, E., Tay, A., Brenner, S., & Venkatesh, B. 2005. STAT4 is a target of the hematopoietic zinc-finger transcription factor Ikaros in T cells. *Federation of European Biochemical Societies* 579: 4470-4478.

Zhou, F., Mei, H., Jin, R., Li, X., & Chen, X. 2011. Expression of Ikaros isoform 6 in chinese children with acute lymphoblastic leukemia. *Journal of Pediatric Hematology/Oncology* 33(6): 429-433.

Electronic references

Abbott. 2011. Abbott Molecular FISH Catalog. Available from: http://www.abbottmolecular.com/static/cms_workspace/img/Support/download_catalog/pdfs/EMEAI__FISHcatalog2011b.pdf. [Accessed 14/7/2013]

Coffalyser. 2008. Coffalyser v8. Available from: http://www.readbag.com/old-mlpa-coffalyser-files-pdf-2008-coffalyser-v8-quick-analysis-guide-pre-release. [Accessed 13/7/2013].

Jordy Coffa & Joost van den Berg. 2011. Analysis of MLPA Data Using Novel Software Coffalyser.NET by MRC-Holland, Modern Approaches To Quality Control, Dr. Ahmed Badr Eldin (Ed.), ISBN: 978-953-307-971-4, InTech, DOI: 10.5772/21898. Available from:

http://www.intechopen.com/books/modern-approaches-to-quality-control/analysis-of-mlpa-data-using-novel-software-coffalyser-net-by-mrc-holland