CHAPTER 1

INTRODUCTION

1.1. Acute Lymphoblastic Leukemia

Acute Lymphoblastic Leukemia (ALL) is a malignancy of immature lymphoid cells. It is characterized by clonal expansion of a transformed hematopoietic cell with progeny arrested at some stage of differentiation. ALL is the most common pediatric malignancy. According to Madras Metropolitan Tumor Registry (MMTR), Chennai, India, the age standardized incidence rate of childhood ALL during 2004 is 3.5 out of 100,000 in boys and 4.2 out of 100,000 in girls (Shanta and Swaminathan, 2006).

Depending upon the origin of lymphoid cells (T or B), ALL is classified mainly into T-ALL and B-lineage ALL. Diagnosis of ALL is based on standard French, American and British (FAB) classification and immunophenotypic criteria. ALL was one of the first malignancies to respond to chemotherapy and was among the first that could be cured in a majority of children (Cortes and Kartarjian, 1995). Modern chemotherapy protocols developed over the last three decades have changed the treatment outcome of this disease dramatically. Childhood ALL, once considered to be a fatal disease has changed to curable with a long-term survival rate of nearly 80% in Western countries and nearly 60% in developing countries. Despite the intensive chemotherapy, in 30-40% of the patients disease relapses as the treatment does not kill all clonogenic leukemic
cells. For the majority of children diagnosed with leukemia, the specific etiologic factors linked to leukemogenesis remain speculative. It has been postulated that the development of clinically overt leukemia occurs over a period of several months from the generation of the initial abnormal leukemic clone until it reaches an estimated leukemic cell burden number of $10^{12}$ cells (Margolin and Poplack, 1997).

1.1.1. Etiology of ALL

Many epidemiological studies have examined the role of genetic and environmental factors which may be linked to the development of childhood ALL, often with conflicting or non-conclusive results. Factors that have been examined in relation to leukemogenesis include: in utero exposure to radiation or parental exposure to ionizing radiation, occupational use of benzene and its derivatives, pesticides, dietary nitrates, linkage to specific HLA antigens, enzymes involved in the metabolism of xenobiotics, polymorphisms of genes encoding folate metabolizing enzymes, prenatal exposure to toxins or infectious agents, socioeconomic status etc (Margolin and Poplack, 1997; Sandler and Ross, 1997; Little, 1999; Franco et al., 2001). The autosomal recessive disorder Ataxia Telangiectasia (AT) appears to be a true etiological factor because patients with AT have an increased risk of developing lymphoid malignancies, including T-lineage ALL. Patients with genetic disorders such as Down's syndrome and Fanconi's anemia are associated with higher than normal risk of leukemia (Toledano and Lange, 1980).
Several mammalian leukemias are associated with infection with specific oncogenic viruses, but the only human leukemia in which a specific viral etiology is implicated is the rare adult T-cell leukemia/lymphoma syndrome, associated with Human T Lymphotrophic Virus-I (HTLV-I) infection (Greaves, 1997). Greaves has hypothesized that ALL develops via at least two mutations with the first mutation occurring in utero and the second mutation occurring postnatally as a result of exposure to an infectious agent (Greaves, 1993). Infection somehow plays a role in the particular B-cell type known as “common-ALL” (c-ALL), which exhibits an incidence peak at about 2 to 6 years. One hypothesis is that c-ALL increases with better hygiene and it could be a pathological response to delayed exposure to some common infectious agent or agents (Greaves et al., 1997).

1.2. Minimal Residual Disease

At diagnosis of ALL, patients generally have >90% lymphoblasts otherwise a total of \( \sim 10^{12} \) leukemic cells in the bone marrow, which can be detected easily with light microscopy (Fig 1.1A). The conventional techniques like light microscopy have a detection limit of 5% blasts. The disease is considered to be in clinical and hematological remission when the bone marrow (BM) blasts are less than 5% (Fig 1.1B). Minimal Residual Disease (MRD) is defined as the lowest level of disease detectable in patients in clinical and hematological remission by the methods available (Foroni, 1999). Though the disease is in clinical and hematological remission, patients still harbor \( \sim 10^{10} \) neoplastic cells in their BM (Campana and Pui, 1995). A major concern in the detection of MRD is
Fig 1.1: Picture illustrating the leukemia load at diagnosis and after chemotherapy of a ALL patient who has responded to treatment.
sensitivity of detecting the tumor cell burden. Hence, to detect MRD with high sensitivity, sophisticated molecular techniques are required.

1.2.1. Methods for the detection of MRD

Currently, three different methods are used to detect and quantitate the MRD with a sensitivity of detecting one leukemic cell in $10^3 - 10^6$ normal cells (Table 1.1). They are

i) Flow cytometric immunophenotyping (using aberrant or leukemia associated phenotype) [Couston-Smith, 1998; Campana and Couston-Smith, 2002]

ii) Reverse Transcriptase - Polymerase chain reaction analysis of chromosomal break point fusion genes (Van Dongen et al., 1999b) and

iii) Detection of clone-specific Immunoglobulin (Ig) and T-cell receptor (TCR) gene rearrangements by PCR (Szczepanski et al., 2002a).

1.2.1.1. MRD monitoring by Flow Cytometric Immunophenotyping

In patients with T-ALL, MRD can be monitored by detecting for cells expressing ‘Terminal deoxynucleotidyl Transferase’ (TdT) and CD3 or other cell markers in bone marrow or peripheral blood (Campana & Couston-Smith, 2002). In B-lineage ALL, one needs to identify aberrant phenotypes that are not expressed by normal hematopoietic cells. The normal hematopoietic cells may limit the immunophenotypic detection of leukemic cells (Campana and Pui, 1995; San Miguel et al., 1999).
Table 1.1: Applicability, Sensitivity, Advantages and Disadvantages of MRD Techniques in Acute Lymphoblastic Leukemia

<table>
<thead>
<tr>
<th></th>
<th>Immunophenotyping</th>
<th>PCR analysis</th>
<th>Fusion gene transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TCR/ Ig rearrangements</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Applicability</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In childhood:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T-ALL</td>
<td>&gt;95%</td>
<td>&gt;95%</td>
<td>10-20%</td>
</tr>
<tr>
<td>Precursor-B-ALL</td>
<td>60-95%</td>
<td>&gt;95%</td>
<td>30-35%</td>
</tr>
<tr>
<td>In adults:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T-ALL</td>
<td>&gt;95%</td>
<td>&gt;95%</td>
<td>10-20%</td>
</tr>
<tr>
<td>Precursor-B-ALL</td>
<td>80-90%</td>
<td>&gt;90%</td>
<td>35-40%</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td>$10^{-3} - 10^{-4}$</td>
<td>$10^{-4} - 10^{-5}$</td>
<td>$10^{-4} - 10^{-6}$</td>
</tr>
<tr>
<td><strong>Advantages</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ fast</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ relatively patient-specific</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ also information on normal cells</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ single cell analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ cell viability can be determined</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ stability of DNA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ patient-specific DNA amount</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>relatively stable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ stable target</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ no background</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- background of normal cells</td>
<td></td>
<td></td>
<td>- instability of RNA</td>
</tr>
<tr>
<td>- immunophenotypic shift</td>
<td></td>
<td></td>
<td>- variable expression levels</td>
</tr>
<tr>
<td>- subclones</td>
<td></td>
<td></td>
<td>- tumor specific (risk of contamination)</td>
</tr>
<tr>
<td>- loss of rearrangements (due to ongoing/ secondary rearrangements and / or subclones)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- labor and time consuming</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- background of normal cells</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

However, immunophenotypic MRD detection is possible in leukemic cells
due to the aberrant antigen expression (Campana and Pui, 1995; Szczepanski et
al., 2001c). One important pitfall of immunophenotypic MRD detection in
leukemic cells is the occurrence of immunophenotypic shifts during the course of
the disease (van Wering et al, 1995).

1.3. T-Cell Receptor molecules and their encoded genes

The antigen receptor on the surface of T-cells is known as T-cell receptor
(TCR). T-cells generally recognize fragments of protein antigens only when they
are bound to major histocompatibility complex (MHC)-encoded glycoproteins on
the surface of host cells. The two subclasses can be distinguished by the
expression of two surface proteins CD4+ and CD8+. CD4+ CD8– cells are helper
T-cells and CD4– CD8+ cells are cytotoxic T-lymphocytes (CTL).

The majority of mature T-lymphocytes (85 to 98%) in peripheral blood and
most lymphoid tissues express TCR αβ; a minority (2 to 15%) expresses TCR γδ.
Both types of TCR molecules are closely associated with the CD3 protein chains,
which together form the TCR-CD3 complex (Fig 1.2b) Each TCR chain consists
of a variable domain and a constant domain. The variable domain of a TCR chain
is encoded by a combination of the available variable (V) and joining (J) gene
segments in case of TCRα and TCRγ chains or by a combination of the available
V, diversity (D) and J gene segments in case of TCRβ and TCRδ chains (Davis
and Bjorkman, 1988).
Fig 1.2. Structure of Immunoglobulin and T-cell Receptor

Fig 1.2a

Fig 1.2b

TCRα  TCRβ

CD3ε  CD3ε

CD3δ  CD3γ

CD3ζ
T- and B-lymphocytes, which together comprise the adaptive arm of the vertebrate immune system, can generate specific responses to a tremendous number of antigens. At the heart of this mechanism are the B-cell receptor (BCR) and T-cell receptor (TCR) proteins; which physically bind their cognate targets and direct cellular responses to these diverse stimuli (Burnet, 1957).

1.3.1. TCR (γ,δ) and IgH Gene Rearrangement

During early T and B cell differentiation, rearrangement of TCR (α, β, γ and δ) gene complex and Ig gene loci occurs by joining germline V, (D) and J gene segments. By this process, each lymphocyte thereby acquires a specific combination of V-(D)-J segment that codes for the variable domains of TCR and Ig molecules (Tonegawa, 1983; Davis and Bjorkman, 1988; Walter et al., 1990). The schematic representation of germline TCRδ, TCRα and TCRγ gene segments are shown in Fig 1.3. For an example in the case of TCRδ chain, first D→J gene rearrangement occurs followed by V→(D)-J thereby deleting all intervening sequences. The rearranged gene segments are transcribed to precursor TCRδ mRNA then introns spliced out to produce mature TCRδ mRNA.

The uniqueness of each rearrangement further depends on random insertion of nucleotides by the enzyme TdT (Blackwell and Alt, 1989, Lewis and Gallert, 1989; Kallenbach et al., 1992) and trimming of nucleotides at the junction sites of V, (D) and J gene segments by exonuclease activity or addition of palindromic (P) nucleotides makes the junctional region of Ig and TCR gene into
Fig 1.3A. TCRα and TCRδ gene complex (#14q11)

Fig 1.3B. TCRγ gene complex (# 7p14-15)
unique "fingerprint-like" sequence (Van Dongen and Wolvers-Tettero, 1991a; Breit and Van Dongen, 1994b). The junctional regions of TCR and Ig encode for the so-called third complementarity determining region (CDR), which contributes considerably to the antigen recognition site of the variable protein domains (Sanz, 1991).

1.3.2. RAG Proteins

The gene rearrangement processes are mediated by a recombinase enzyme system, which includes the protein products of the so-called recombinase activating genes (RAG-1 and RAG-2). The recombinase complex recognizes specific recombination signal sequences (RSS). It consists of a palindromic heptamer and nonamer with respective consensus sequence of CACAGTG and AC(A)₅CC separated by spacer regions of 12 or 23 base pairs. RSS is present at 3' end of V-gene segment, at both sides of D-gene segments and 5' side of J gene segment. A gene rearrangement involves back-to-back fusion of heptamer-nonamer RSS followed by deletion of these RSS and all intervening sequences in the form of a circular excised product and by joining of V-D-J gene segments (Schatz et al., 1989; Oettinger et al., 1990; Libe et al., 1992; Jung and Alt, 2004).

1.3.3. Repertoire of TCR molecules

The combined V(D)J sequence constitutes a ‘specific signature’ of each lymphocyte (Tonegawa, 1983). Due to the clonal origin of the neoplasm, each malignant lymphoid disease will represent the expansion of a clonal population
with a specific Ig/TCR signature. By the combinatorial diversity and junctional diversity $>10^{12}$ diverse TCR or Ig molecules are produced (van Dongen and Wolvers Tettero, 1991a). The frequency of TCR gene rearrangements varies in different ethnic and geographic populations (Scrideli et al., 2004; Scrideli and Tone, 2006). TCR and Ig gene rearrangements detected at diagnosis of ALL are used as clonal markers to detect and quantitate the MRD.

Table 1.2: Antigen-recognizing repertoire of TCR molecules

<table>
<thead>
<tr>
<th></th>
<th>TCR(\alpha)</th>
<th>TCR(\beta)</th>
<th>TCR(\gamma)</th>
<th>TCR(\delta)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Germline diversity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V gene segments</td>
<td>45</td>
<td>44-47</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>D gene segments</td>
<td>---</td>
<td>2</td>
<td>---</td>
<td>3</td>
</tr>
<tr>
<td>J gene segments</td>
<td>50</td>
<td>13</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td><strong>Combinatorial diversity</strong></td>
<td>$&gt;3 \times 10^6$</td>
<td>$&gt;5000$</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Junctional diversity</strong></td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td><strong>Total repertoire</strong></td>
<td></td>
<td></td>
<td>$&gt;10^{12}$</td>
<td></td>
</tr>
</tbody>
</table>

The total number of V, D and J gene segments and estimated diversity of TCR molecules is shown in Table 1.2. The combinatorial repertoire is estimated to be $\sim 2 \times 10^6$ for Ig molecules, $\sim 3 \times 10^6$ for TCR$\alpha\beta$ molecules and $5 \times 10^3$ for TCR$\gamma\delta$ molecules. The junctional diversity further contribute to the total repertoire of Ig and TCR molecules, estimated to be $>10^{12}$ (Van Dongen et al., 2002a).
\(\text{TCR}\alpha\beta\) contains many V and J gene segments and hence PCR is difficult to be studied with different sets of primers whereas TCR\(\gamma\delta\) contain limited number of germline VJ segments, hence it is easy to detect the gene rearrangements by PCR (Breit et al., 1991a; Van Dongen et al., 1996).

1.4. T-Cell Acute Lymphoblastic Leukemia (T-ALL)

T-ALL derives from T-lymphoid precursors arrested at various stages of differentiation. It is one of the commonly occurring type of leukemia in India. The percentage of T-ALL in India is very high (20-44%) when compared to Western countries (10-20%). The incidence of T-ALL is 20.7% in west India, 31.8% in North India and is the highest (44%) in South India (Sazawal et al., 1996; Rajalekshmy et al., 1997; Magrath et al., 2005). It is associated with male gender, high WBC count, presence of mediastinal mass and CNS leukemia (Pui et al., 1990a). In patients with ALL, Hyperdiploidy (> 50 chromosomes) correlated with favorable outcome and pseudodiploidy associated with poor outcome. The hyperdiploid karyotype is more often associated with pre-B or early pre-B ALL whereas the pseudodiploid karyotype is more often associated with T-ALL (Uckun et al., 1998).

Immunological markers used for classifying T-ALL are CD1, CD2, CD5, CD7, cytoplasmic CD3, surface membrane CD3 (sm CD3), CD4 and CD8. Using CD1 and Sm CD3 at least three types of T-ALL can be identified. They are immature T-ALL (CD1−/SmCD3−), common thymocytic (cortical) T-ALL
(CD1+/SmCD3−/+). The immature T-ALLs consist of either rarely occurring prothymocytic T-ALL (pro-T-ALL) showing CD2+, CD7+ and CD5− or the immature thymocytic T-ALL (pre-T-ALL) showing markers for CD2+, CD5+ and CD7+ (Foon and Todd, 1986; Van Dongen et al., 1988; Pui et al., 1993; Bene et al., 1995). The cortical T-ALL is classified further into SmCD3− and SmCD3+ types. The expression of CD3 antigen on the cell surface is associated with TCR molecules. In T-ALL, CD3+ is approximately 35% including 20% of TCRαβ+ and 15% of TCRγδ+ and SmCD3− is 65% (Van Dongen et al., 1990).

1.4.1. TCRG Gene Rearrangements

TCR Gamma (TCRG) gene rearrangement is a preferential target for clonality studies since it is rearranged at an early stage of T-lymphoid development, probably just after TCR Delta (TCRD) in both TCRαβ and TCRγδ lineage precursors (Bloom et al., 1999). It is rearranged in >90% of T-ALL (Szczepanski et al., 1998). The human TCRG locus on chromosome 7p14 contains 14 Vγ segments, only 10 of which have been shown to undergo rearrangement (Fig 1.3B). The expressed Vγ repertoire includes only six functional Vγ genes (Vγ2, Vγ3, Vγ4, Vγ5, Vγ8 and Vγ9) but rearrangement also occurs with the ΨVγ7, ΨVγ10 and ΨVγ11 (pseudogenes) gene segments (Chen et al., 1998). TCRG rearrangements also occur in 60% of B-lineage ALL (Szczepanski et al., 1999), implying that they cannot be used for the assessment of B- vs T-cell lineage in immature proliferations.
1.4.2 TCRD Gene Rearrangements

The human TCRD gene locus is located on chromosome 14q11.2. During T-cell differentiation, TCRα rearrangement is preceded by deletion of the TCRδ genes because the major part of the TCRδ gene complex (Fig 1.3A) is located between the Vα and Jα gene segments and is flanked by the so-called δREC and ψJα gene segments (Takhira et al., 1988; Griesser et al., 1988; Isobe et al., 1988). TCRδ gene deletion is mediated via rearrangement of the δREC and ψJα gene segments (De Villartay et al., 1988; Hockett et al., 1988; Dyer, 1989; Breit et al., 1994a; Verschuren et al., 1998). TCRδ genes rearrange followed by TCRγ gene rearrangement, which might result in TCRγδ+ T-lymphocytes if the rearrangement is functional. TCRαβ+ T-lymphocytes develop via rearrangement of TCRβ prior to TCRα rearrangement (Van Dongen et al., 1991a). In TCRD gene rearrangements, although the small number of V, D and J gene segments available for recombination limits the combinatorial diversity, the extensive junctional diversity occurs due to the addition of 'N' regions, 'P' regions, and random deletion of nucleotides by recombinases. This diversity is also extended by the recombination of up to three Dδ segments and therefore four 'N' regions within the rearranged TCRD locus. That makes the TCRD as a useful target for PCR analysis and has been used most extensively as clonal markers in both T- and B-cell ALL (Breit et al., 1991b; Langlands et al., 1993).

The TCRD locus is the first of all TCR loci to rearrange during T-cell ontogeny. The first event is a Dδ2-Dδ3 rearrangement, followed by a
Vδ2-(Dδ1-Dδ2)-Dδ3 rearrangement and finally Vδ-Dδ-Jδ rearrangement. Immature rearrangements (Vδ2-Dδ3 or Dδ2-Dδ3) occur in 70% of precursor-B-ALL (and hence non-lineage restricted) [Szczepanski et al., 1999] while there is a predominance of mature rearrangements comprising incomplete Dδ2-Jδ1 and complete Vδ1, Vδ2, Vδ3, to Jδ1 found in T-ALL (Breit et al., 1993a; Schneider et al., 1997).

1.4.3. TCRγδ+ T lymphocytes in peripheral blood and lymphoid organs

At birth, the repertoire of cord blood TCRγδ+ T-cells is broad, with no apparent restriction or preferred expression of particular Vγ/Vδ combinations. During childhood, the peripheral blood TCRγδ+ T-cell repertoire is shaped so that in normal peripheral blood ~85% of TCRγδ+ T-lymphocytes express Vγ9-Jγ1.2-Cγ1 chains in combination with Vδ2-Jδ1-Cδ chains, whereas in ~60% of the thymocytes TCRγδ molecules contain Vδ1-Jδ1-Cδ chains, linked to any of the TCRγ chains (Breit et al., 1991a; Kabelitz et al., 1999; Triebel et al., 1998). TCRγδ+ T-cells are evenly distributed throughout human lymphoid tissues, but there is a preferential expression of particular Vδ segments in some organs. Most intraepithelial TCRγδ T-cells in small intestine and colon express Vδ1. Similarly, Vδ1 is expressed by TCRγδ+ T cells in normal spleen but TCRγδ+ T cells in the skin express Vδ2 gene (Deusch et al., 1991; Alaibac et al., 1993; Van Dongen et al., 2003).
1.4.4. Detection of clonal TCR Gene Rearrangements by PCR

As the different 'V' genes are grouped into families on the basis of sequence homology, primers are chosen to be complementary to sequences found in all or most members of the V gene family. To study the TCRγ gene rearrangements, four forward primers for Vγ family (VγI to VγIV) and two reverse primers for 'J' region (Jγ1.3/2.3 & Jγ1.1/2.1) are used. To study the TCRδ gene rearrangements, four forward primers (Vδ1, Vδ2, Vδ3 and Dδ2) in combination with two reverse primers (Jδ1 and Dδ3) are used. PCR products are analysed by size after agarose or polyacrylamide gel electrophoresis. The main advantages of PCR analysis are the technique is rapid, low amount of DNA required and relatively good sensitivity of 1-5% (Van Dongen et al, 2003).

1.5. B-Lymphopoiesis

Human B-lineage cells are present in multiple tissue sites in early fetal development. However, from midgestation through the eighth decade of life, BM is the exclusive site of B-lymphopoiesis (Le Bien, 2000). B-lymphocytes express unique receptors on their surface membrane called B-cell receptor (BCR) that allows the specific recognition of antigens. Antigen recognized by the BCR is mediated by membrane inserted Ig molecule of the IgM or IgD class in naïve B cells or of other Ig classes in antigen-primed cells. In addition, the BCR also contain at least two co-receptor molecules Igα and Igβ, which is essential for trans-membrane signaling in response to antigen (Rajewsky, 1996).
1.6.1. Ig molecules and their encoded genes

Ig molecule consists of two Ig heavy (IgH) chains and two Ig light (IgL), chains held together by disulfide bonds (Fig 1.2a). Each B-lymphocyte or B-lymphocyte clone expresses only one type of Ig light chain (either Igκ or Igλ). The Ig protein chains are composed of one variable domain, which is involved in antigen recognition and one constant domain in case of IgL or three to four constant domains in case of IgH ( Tonegawa, 1983).

The variable portion of an IgH chain is encoded by an exon, which consists of a combination of V, D and J gene segments. The variable domain of IgL is encoded by a combination of V and J gene segment (Early et al., 1980; Tonegawa, 1983). During B-cell differentiation, combinations of the available V, D and J gene segments are made through a process of gene rearrangement (Alt et al., 1986; Walter et al., 1990). The total number of functional V, D and J germline gene segments is shown in Table 1.3. By combinatorial diversity and Junctional diversity, a repertoire of Ig molecules is produced (Van Dongen et al., 2002a).

The rearrangement of Ig genes is an ordered process. The IgH loci are the first to rearrange during B-lymphocyte development, beginning with D to J joining followed by VH to (D) J and the complete heavy chain is assembled at DNA level (Alt, 1984). The switch from DJ to V(D)J recombination is regulated at the level of VH gene accessibility and requires the transcription factor Pax-5 and signals from the interleukin-7 receptor α (IL-7Rα) chain (Kurosaki, 2003).
Table 1.3: Repertoire of Ig molecules

<table>
<thead>
<tr>
<th></th>
<th>IgH</th>
<th>Igκ</th>
<th>Igλ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of functional</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene segments</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V gene segments</td>
<td>40–46</td>
<td>34–37</td>
<td>30–33</td>
</tr>
<tr>
<td>D gene segments</td>
<td>25</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>J gene segments</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Combinatorial diversity</td>
<td></td>
<td>2×10⁶</td>
<td></td>
</tr>
<tr>
<td>Junctional diversity</td>
<td>++</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td>Estimated total repertoire</td>
<td></td>
<td>&gt;10¹²</td>
<td></td>
</tr>
</tbody>
</table>

Mature B-lymphocytes further extend their Ig repertoire upon antigen recognition in germinal centers via somatic hypermutation that occurs on the V(D)J exon of IgH and Ig-light-chain genes in the variable domain of Ig molecule. It occurs by mutations and rare insertions or deletions of nucleotides that leads to affinity maturation of the Ig molecules (French et al., 1989). Somatically mutated Ig genes are therefore found in mature B-cell malignancies of germinal center or post germinal center origin (Kuppers et al., 1999). Similar mechanism has not been reported in TCR genes so far.
1.5.2. Formation of CDR3

The Ig heavy and light chains each contain three hyper variable region and four conserved framework areas. The CDR is the most diverse region of the antibody molecule. The CDR3 is in direct contact with antigen and is the most variable portion of Ig molecule (Tonegawa, 1983; Sanz 1991). The CDR3 sequence is unique to each rearrangement and therefore identifies individual B-cells or clonal B-cell expansion (Tonegawa, 1983). The CDR3 in Ig consist of 3' end of VH, all of D and the 5' end of JH chain. It also contains 'N nucleotides' which are randomly inserted at both VH-D and D-JH junctions by TdT (Kiyoi et al., 1992). Another source of diversity in CDR3 is random deletion by nucleases of terminal nucleotides of rearranged VH, D and JH gene segments (Campana and Pui, 1995).

![Diagram of V region and CDRs](image)

**Fig 1.4**: Organization of variable domain of Ig heavy chain
1.5.3. B-Cell Development

The $V_HD_HJ_H$ gene repertoire is restricted and developmentally regulated at early stages of differentiation (Schroeder et al., 1987; Willeford et al., 1996). In the most immature B-cell precursors- progenitor or pro-B-cells, the IgH remain germline or there is only $D_HJ_H$ joining. At the next stage, in early pre-B cells, VH genes join to the DH-JH to complete the IgH rearrangements. The stages of B-cell development are illustrated in Fig 1.5. The pre-BCR consists of functional IgH chain in association with surrogate light chains. Assembly of the pre-BCR requires the successful completion of V(D)J rearrangement in at least one IgH allele and the synthesis of a membrane inserted heavy chain of the Igμ class. Signaling through the pre-BCR stops further rearrangement on the second IgH allele, a phenomenon called ‘allelic exclusion’. So, each mature B-lymphocyte contains an antigen receptor with a single specificity. It is followed by the IgL rearrangement (either κ or λ) leading to the generation of B-lymphocytes that express an IgH/IgL antigen receptor on their surface. Due to the imprecise joining event, the rearrangement leads to the production of a functional Ig gene in only a minority of all attempts (Alt et al., 1980; Alt et al., 1984).

1.5.4. Precursor-B-ALL

Precursor-B-ALL is derived from a single transformed hematopoietic cell "frozen" at a stage where it is already committed to the B-lineage and has undergone either complete or partial VDJ recombination within the Ig heavy chain (Stankovic et al., 2000). It is classified into pro-B-ALL, pre-B-ALL, common-
Fig. 1.5. B-CELL DEVELOPMENT

- Pro-B: D-JH (D)JH
- Pre-B: V(D)JH V-(D)JH Cμ+
- Immature B: V(D)JH V-JL VJL SlgM+
- Mature B: SlgD+ SlgM+
ALL and transitional pre-B-ALL. All four forms of precursor-B-ALL are positive for TdT, CD10, pre-B cylgμ and smlgμ are important markers for discriminating the subtypes (Pui et al., 1993). The incidence of precursor-B-ALL is 45.4% in South India, 59.5% in North India and is the highest (75.2%) in West India (Magrath et al., 2005).

1.5.5. IgH Gene Rearrangements

In B-lineage cells, rearrangement of one D gene and one JH gene is followed by the addition of one of the numerous VH genes to the fused D-JH segments. If successful, the resulting VH-D-JH rearrangement is transcribed into RNA that is spliced to the constant region RNA (Cμ) before translation into an Ig heavy (μ) chain in the pre-B cell. If the initial rearrangement produces a sequence that cannot be translated (nonproductive or abortive rearrangement) then rearrangement of the IgH locus proceeds on the other allele (Alt et al., 1984).

According to clonal selection theory, allelic exclusion ensures that a B or T cell expresses a single "Specificity" of 'ag' receptor (Pernis et al., 1985). Thus each pre-B cell express only one kind of heavy chain. After successful heavy chain rearrangement in pre-B cells, cytoplasmic μ chain attaches to the endoplasmic reticulum (ER) predominantly in association with the Ig binding protein Bip (Haas and Wabl, 1983). Binding of newly formed μ chain to surrogate light chain proteins (ψ light chains) results in displacement of small amounts of μ
from the ER and in late pre-B cells, allows translocation of the resulting complex to the cell surface (Rolink et al., 1991; Lassoued et al., 1993; Melchers et al., 1993). The ψ light chain proteins are encoded by the V_{pre-B} and λ_{5/141} genes, which have significant homology with the V_λ, J_λ and light chain constant region genes (Pillai et al., 1988; Bossy et al., 1991). The association of μ the ψ light chain and two associated proteins Igα and Igβ on the pre-B cell surface forms the B-cell antigen receptor complex, which in many ways resembles its counterpart T-cell antigen receptor complex (van Noesel and van Lier, 1993; Stewart and Schwartz, 1994).

1.5.6. PCR-based Detection of IgH Gene Rearrangements

The IgH genes encoded on chromosome 14q32.2 are grouped into 7 V_H families, 7 D_H families and consist of 6 functional J_H segments based on their sequence homology and not by their location on chromosome (Matsuda et al., 1988). The IgH rearrangement involves the joining of V, D and J segments normally separated by more than several kb, when they are brought into adjacent position over approximately 320bp of DNA. It occurs in >95% of cases with B-lineage ALL and in 5-10% of T-cell ALL. The resulting sequence is known as CDR3 and can be amplified by PCR (Foroni et al., 1999). Detection of MRD is based on identification of the patient/leukemia specific clonal CDR3 in the background of normal polyclonal bone marrow cells.
rearrangements (e.g. DH-JH replacements, VH→VH replacements), which result in one or more subclones (Wasserman et al., 1992a; Steenbergen et al., 1993).

TCR gene oligoclonality is rarely seen at diagnosis in T-ALL (Beishuizen et al., 1994; Van Dongen and Wolvers-Tettero, 1991h). Studies in precursor-B-ALL reveal that 20% of cross-lineage TCR gene rearrangements are oligoclonal, which is less than that of IgH gene rearrangements (Beishuizen et al., 1991; Steward et al., 1994a; Ghali et al., 1995; Steenbergen et al., 1995a). The problem of oligoclonality at diagnosis is the uncertainty which clone is going to emerge at relapse and should be monitored with MRD-PCR techniques. Also, secondary and ongoing IgH gene rearrangement might also occur in the time period between diagnosis and relapse, resulting in loss of leukemia specific MRD targets. Patients with oligoclonal rearrangement reflect an aggressive disease and tend to have a shorter disease free survival. Usually relapse subclones are different from the subclones observed at presentation (Steward et al., 1994a; Szczepanski et al., 2002b).

Clonal MRD-PCR targets in childhood precursor-B-ALL are characterized by high stability, with ~85-90% of all targets detectable at relapse. Conversely, only 40% of the oligoclonal MRD-PCR targets are preserved at relapse (Szczepanski et al., 2002b). Despite the high frequency of immunogenotypic changes in childhood ALL at relapse, at least one rearranged IgH, TCRG and /or TCRD allele remains stable in >90% of T-ALL (Beishuizen et al., 1994; Szczepanski et al., 2002b).
1.6. Methods for Studying Clonality

1.6.1. Southern Blot analysis

For a long time, Southern blot analysis has been the gold standard technique for molecular clonality studies. Southern blot analysis focuses on the rearrangement diversity of TCR/ Ig gene segments and hence takes advantage of combinatorial repertoire. It is based on the detection of non-germline ('rearranged') DNA fragments, obtained after digestion with restriction enzymes. Digestion of DNA with well-chosen restriction enzymes results in fragments of 2-15kb and well-positioned DNA probes (particularly downstream J segment probes) allow the detection of virtually all TCR and Ig gene rearrangements (Van Dongen et al., 1991a; Beishuizen et al., 1993; Breit et al., 1993a; Moreau et al., 1999; Langerak et al., 1999b). Optimal Southern blot results for clonality assessment can be obtained with the IgH, IgK and TCRB genes, because these genes have an extensive combinatorial repertoire that can be evaluated with only one or two DNA probes (Beishuizen et al., 1993, Beishuizen et al., 1994; Langerak et al., 1999b). Whereas, TCRG and TCRD genes have a limited combinatorial repertoire, which are less optimal for discrimination between monoclonality and polyclonality via Southern blot analysis (Beishuizen et al., 1993; Van Dongen and Wolvers-Tettero, 1991b). Despite the high reliability of Southern blot analysis, it is increasingly replaced by PCR techniques, because of inherent disadvantages such as the technique is time-
consuming, need technical expertise, also requires 10-20 μg of high-quality DNA and has a limited sensitivity of 5-10% (Van Dongen et al., 1991b).

1.6.2. PCR based analysis

Discrimination between clonal (leukemia-derived) and polyclonal (normal lymphoid cell-derived) PCR products is virtually impossible by standard agarose gel electrophoresis. Hence, PCR amplified rearranged gene products need additional analysis for distinction between clonal and polyclonal PCR products. Methods that have been successfully applied for this purpose include single stranded confirmation polymorphism analysis (Davis et al., 1993; Baruchel et al., 1995), homo-heteroduplex analysis (Bottaro et al., 1994; Langerak et al., 1997), denaturing gradient gel electrophoresis (Bourguin et al., 1990), fluorescent genescan analysis (Kneba et al., 1995) and denaturing high performance liquid chromatography (Udo zur Stadt et al., 2001). Of these methods, homo-heteroduplex analysis is the simplest, reliable and cost-effective (Langerak et al., 1997).

1.6.3. Homo-Heteroduplex analysis

In Homo-Heteroduplex analysis, double stranded PCR products are denatured at 94°C for 5 min followed by renaturing at 4°C for 1 hour to induce duplex formation. This results in homoduplex (sequences with identical junctional regions) or heteroduplex (sequences with different junctional regions).
Fig 1.6. Principle of Homo-Heteroduplex Analysis

Monoclonal cells

Monoclonal cells in polyclonal background

dS DNA

Denaturation and renaturation of PCR product

Heteroduplex

Homoduplex

1, 2 → Clonal

3 → Polyclonal
Homoduplexes and heteroduplexes are subsequently separated from each other by running on a 6% non-denaturing polyacrylamide gel electrophoresis that clearly discriminates between the rapidly migrating homoduplex bands and slowly migrating heteroduplex bands (Fig 1.6).

The sensitivity of Homo-Heteroduplex analysis in the detection of clonality is 5-10%. Vγ1-Jγ1.3/2.3 rearrangement showed a sensitivity of only 10% but using VγII(Vγ9)-Jγ1.3/2.3 gene rearrangement a sensitivity of 5% was reached. The reduced sensitivity of Vγ1 probably caused by the high background of these rearrangements, occurs commonly in normal T-cells (Langerak et al, 1997).

The detection of clonality does not always imply the presence of malignancy. Some clinically benign proliferations have a clonal origin, such as many cases of CD8+/CD4+ T-lymphocytosis, benign monoclonal gammapathies, initial phases of EBV+ lymphoproliferations (frequently being oligoclonal) in immunodeficient patients, and benign cutaneous T-cell proliferations (e.g. lymphomatoid papulosis). This implies that results of clonality studies should always be interpreted in the context of the clinical, morphological and immunophenotypic diagnosis, that is in close collaboration with Hematologists, Pathologists and Immunologists (Van Dongen et al, 2003).
Chromosomal translocation is the hallmark of leukemias and Acute Lymphoblastic Leukemia in particular. Due to the large number of cell divisions that occur in the bone marrow ($10^{11}$ per day in adult human), DNA damage arising from endogenous and exogenous genotoxic factors and the lack of complete fidelity in DNA-repair leads to chromosomal translocations (Greaves and Wiemels, 2003). Translocations are mostly balanced or reciprocal. The products of balanced chromosomal translocations are fusion genes generating either a dysregulated partner gene or a chimeric fusion gene with new function either altered transcriptional regulation or constitutive kinase activity (Greaves, 1999). Chromosomal translocations involve illegitimate recombination or juxtaposing of normally separate genes. It is initiated by double-strand DNA breaks (DSB) followed by repair mechanism. Breakpoint cluster region (BCR) occurs in introns and the leukemic clone of each patient has a unique DNA breakpoint. Either that the BCR themselves are vulnerable to breaks or that DNA breaks and translocations can occur randomly throughout the gene (Greaves and Wiemels, 2003).

The presence of chimeric oncogenes is also reported in newborns or normal children. It has been shown that the TEL-AML1 fusion occurs during human embryonic development (Wiemels et al., 1999). TEL-AML1 fusion gene is present in the blood of about 1% of all newborns. Fortunately, only about one in every 100 of these newborns will develop leukemia later during childhood.
Chromosomal translocations can initiate leukemogenesis but are usually not sufficient with additional postnatal events being required (Mori et al., 2002).

1.7.1a Repair of DNA breaks

The two repair pathways in mammalian cells are Homologous recombination (HR) and non-homologous end joining (NHEJ). There is experimental evidence that simultaneous induction of DSB in two chromosomes can produce interchromosome recombinations at high frequency by combination of HR and NHEJ.

- NHEJ is the dominant repair mechanism in mammalian cells and is mediated by well-characterized set of proteins
- Two multi-protein complexes are necessary for NHEJ-DNA-dependent protein kinase/Ku, which binds and positions DSB ends and Ligase4/XRCC4, which catalyses the ligation step
- The NHEJ pathway is crucial for the normal integrity of the genome and it seems strongly to favour re-ligation of DNA ends to the correct partner following DSB (Khanna and Jackson, 2001; Jackson, 2002).

However, in the presence of coincident multiple breaks in DNA, the normal NHEJ process facilitates illegitimate recombination (Greaves and Wiemels, 2003).
1.7.1b Requirements for functional chromosomal translocations

To produce a functional leukemogenic chromosomal translocation, the following criteria have to be met (Greaves and Wiemels, 2003).

- Simultaneous DSB have to occur in two chromosomes in a single cell
- Breaks have to occur in 'relevant' introns of 'relevant' genes to generate a functional chimeric gene product
- Chromosomal translocation has to be in frame to generate a functional chimeric gene product
- The DSB that initiate the translocation cannot be lethal to the cell in which they occur

1.7.2. Methods Used for the Detection of Chromosomal Translocations

Conventional cytogenetics can detect only major translocations like BCR-ABL and E2A-PBX1; it cannot detect minor translocations such as TEL-AML1 and MLL-AF4. Also conventional karyotyping may miss out the abnormalities when the tumor cells are low. Molecular methods used for detection of chromosominal translocations are Southern blotting, Fluorescence in situ hybridization (FISH), RT-PCR and Real-time PCR.

Southern blot needs a large amount of DNA and use of radioactivity for the study and the technique is laborious for routine diagnostic use. FISH is a more specific technique but high cost of fluorescent-labeled probes is not affordable in countries with limited resources. Detection of translocations using
RT-PCR is a highly specific and sensitive technique for routine diagnostic use. Moreover, Real-time PCR is a highly sensitive and specific technique and amplification is detected during exponential phase of PCR. It requires minimal amount of DNA and also obviates post-PCR analysis thus prevents carry over contamination in clinical setting.

PCR analysis of fusion genes is based on the design of oligonucleotide primers at opposite sides of the breakpoint fusion region, so that the PCR product contains the tumor specific fusion sequence. In most types of chromosomal aberrations with fusion genes, the breakpoints in different patients are spread over >10kb or more, distances which are difficult by DNA-PCR. So the fusion genes are transcribed into fusion mRNA, which serve as the PCR target after reverse transcription (RT) into cDNA (Foroni et al., 1999). The commonly detected fusion genes in ALL are mBCR-ABL, E2A-PBX1, TEL-AML1 and MLL-AF4.

1.7.3. Chromosomal Translocations and Fusion gene transcripts in B-ALL

1.7.3.1. t(9;22)(q34;q11) or mBCR-ABL

The Philadelphia chromosome (Ph) being the hallmark of Chronic Myeloid Leukemia (CML), also occurs in ~5% of childhood ALL and in 20-50% of adult ALL. The translocation of 3' portion of C-ABL proto-oncogene on chromosome 9 to the 5' region of BCR gene on chromosome 22 generates a BCR-ABL chimeric oncogene. The breakpoint in BCR gene is located in 'minor-
breakpoint cluster region’ (m-bcr) between two alternative exons and exon 2: breakpoints in ABL gene is located in the large intron region (~200kb) between ABL exon 1b and exon 2 (also called exon a2). The breakpoints on chromosome 22 differ in their position within the BCR gene, giving rise to e1-a2 (occurs in >95% of BCR-ABL fusion gene) and e1-a3 (rare) types of fusion genes. This results in the production of a BCR-ABL protein of 190 kDa in molecular weight (p190\textsuperscript{BCR-ABL}). It results in aberrant tyrosine kinase activity and occurs in 3-5% of childhood ALL and 25-40% of adult ALL (Rabbits, 1994; Ferrando et al., 2000). The presence of Ph-chromosome and consequently the BCR-ABL translocation is an independent unfavorable prognostic factor, which affects both hematological complete remission rate and probability of DFS (Crist et al., 1990b).

In CML, the breakpoint on chromosome 22 occurs at central region of BCR gene (between exons 13 and 15) called ‘Major Breakpoint Cluster Region’ (M-BCR) and produces fusion transcripts that encode a BCR-ABL protein of 210 kDa called p210\textsuperscript{BCR/ABL} (Melo, 1996).

1.7.3.2. t(1;19)(q23;p13) or E2A-PBX1

It results in fusion of E2A gene on chromosome 19p13 with PBX1 locus on chromosome 1q23. E2A-PBX1 fusion gene encodes a hybrid transcription factor containing the transactivation domain of E2A and the DNA binding homeodomain of PBX1. The genomic organization of E2A is well defined and breakpoints occur almost exclusively in a 3.5kb intron region between exon 13
and 14. The genomic organization of PBX1 gene is not yet fully known and the breakpoints are dispersed over an intronic region of about 50kb between exon 13 and 14. E2A encodes the two transcription factors E12 and E47 via alternative splicing of precursor mRNA. E12 and E47 are ubiquitous helix-loop-helix (HLH) containing proteins that bind to the E-box element in the κ light-chain DNA enhancer region (Mellentin et al., 1990).

PBX1 plays a role in the regulation of cell differentiation (Kamps, 1997). Wild type E2A is essential for normal lymphopoiesis and regulation of B-cell development. E2A −/− mice demonstrate a total block in B-lymphoid development with absence of DH-JH and VH-DH-JH rearrangement (Zhuang et al., 1994; Bain et al., 1994). Loss of E2A also leads to block in T-lymphoid development, with development of thymic lymphomas (Bain et al., 1997). Leukemogenic effect of E2A-PBX1 is mediated by induction of cell differentiation arrest (Hunger, 1996; Kamps, 1997).

This translocation is strongly associated with pre-B-cell ALL and associated with poor prognosis (Crist et al., 1990a). It occurs in 5-6% of pediatric ALL cases and less than 5% in adults (Nichols and Nimer, 1992; Rabbits, 1994; Thomas Look, 1995; Ferrando et al., 2000).

1.7.3.3. t(12;21)(p13;q22) or TEL-AML1

This translocation occurs by the fusion of 5' portion of TEL (ETV6) on chromosome 12 to the entire coding region of AML1 (CBFA2) on chromosome
21 (Golub et al., 1995; Romana et al., 1995a). TEL gene is a member of the ETS family of transcription factor genes and AML1 codes for the α subunit of core binding factor, a master regulator of the formation of definitive hematopoietic stem cells (Loh and Rubnitz, 2002; Speck and Gilliland, 2002). TEL is required for the homing of hematopoietic progenitor cells to the bone marrow (Wang et al., 1998), whereas AML1 is the DNA-binding component of the heterodimeric transcription factor (CBFα plus CBFβ) called core-binding factor (CBF), which has a central role in hematopoiesis (Lorsbach and Downing, 2001; Speck and Gilliland, 2002). The HOX genes probably operate downstream of the transcriptional cascade initiated by the CBF (Buske and Humphries, 2000; Canon and Banerjee, 2000).

The binding of AML1 to a sequence termed the core enhanced sequence recruits transcription factors and co-activators including histone acetylases, which add acetyl groups to DNA-bound histones, thereby causing conformational changes in chromatin that activates the transcription of target genes. Like AML1, the abnormal TEL-AML1 fusion protein can bind to the core enhanced sequence, but instead of activating transcription, it recruits histone deacetylases which induce closure of the chromatin structure and hence inhibition of transcription (Hiebert et al., 1996). These changes in the normal AML1-mediated transcriptional cascade into repression alter both the self-renewal and differentiation capacity of hematopoietic stem cells (Speck and Gilliland, 2002; Downing, 2003; Zelent et al., 2004; Pui et al., 2004).
The breakpoint in TEL gene occurs in a 15kb region between exon 5 and 6 and breakpoint in AML1 gene occurs most frequently in large intron 1 or in intron 2. TEL gene has been completely sequenced but the genomic organization of AML1 gene has not yet been completely unraveled (Van Dongen et al., 1999). This translocation is found in 16-32% of pediatric B-cell ALL and 3-4% of adult ALL and has a good prognosis (Borkhardt et al., 1997; Ozbek et al., 2003). This translocation is detected in 1% of newborns show that the translocation is present in blood cells at birth, up to 5 to 10 yrs of age before the development of leukemia. This shows that TEL-AML1 translocation is the initiating event in this leukemia (Armstrong, 2005). Fortunately, only about one in every 100 of these newborns develops leukemia later during childhood.

1.7.3.4. t(4;11)(q21;q23) or MLL-AF4

Mixed Lineage Leukemia (MLL) gene encodes on chromosome 11q23. The N-terminal portion of MLL is fused to the C-terminal portion of more than 20 fusion partners. MLL normally regulates the expression of HOX genes; altered HOX gene expression result in leukemogenesis (Domer et al., 1993). It is present in 60% of infant ALL cases, 2% of childhood ALL and 3-6% of adult ALL. This fusion gene is associated with pro-B ALL phenotype, cy CD79a+, CD19+, CD10 +ve and co-expression of myeloid differentiation antigens (CD15/CD33/CD65) (Pui et al., 1991). Many studies have identified MLL-AF4 as an adverse prognostic factor in infant leukemia (Van Dongen et al., 1999b; Felix and Lange, 1999).
Translocations affecting 11q23 region occur in up to 80% of secondary leukemias arising after chemotherapy with topoisomerase II inhibitors (Biondi et al., 2000; Pui et al., 2000). A study suggests that transplacental fetal exposure to substances that inhibits topoisomerase II might be a critical event in the generation of leukemias. Flavanoids (in trans), Quinolone antibiotics, Benzene metabolites, Catechins and Estrogens can inhibit topoisomerase II both in vivo and in vitro and may cause mutations that lead to acute leukemias with MLL rearrangements (Biondi et al., 2000; Greaves and Wiemels, 2003). A case-control study disclosed a significant association between in utero exposure to DNA damaging agents, a nonsteroidal anti-inflammatory drug (Dipyrone), herbal medicines and a mosquitocidal agent (Baygon) in the development of infant leukemia with MLL rearrangements (Alexander et al., 2001).

1.7.4. Tal-1 Deletions in T-ALL

Tal1-deletion is the most common genetic abnormality in T-ALL. The TAL-1 gene (T-cell acute leukemia) also known as SCL or TCL5, located on 1p32, encodes for the basic helix loop helix protein Tal-1 (Begley et al., 1989; Aplan et al., 1990). It is a transcriptionally complex locus, extending over 16 kb containing eight exons which can give rise to different mRNA transcripts by alternative splicing (Aplan et al., 1990; Bernard et al., 1991). TAL-1 gene is essential in the earliest stages of hematopoietic stem cell development and differentiation (Begley et al. 1989; Aplan et al., 1990; Porcher et al., 1996).
Translocation of TAL1 (1p32) next to the TCRδ loci in the t(1;14) (p32;q11), occurs in 3% of T-cell ALL and results in aberrant TAL1 expression (Carroll et al., 1990). A much more frequent mode of TAL1 deregulation is a site-specific deletion (tal-1 deletion) of ~90kb that occurs with a frequency of 6% to 26% of T-ALL cases (Brown et al., 1990). As a result of this 90kb deletion, the coding exons of the tal-1 gene is juxtaposed to the first non-coding exon of the Sil gene, which is almost completely deleted. The expressed SIL-TAL1 fusion transcript produces a normal TAL-1 protein, but it is transcriptionally controlled by the Sil gene promoter (Bernard et al., 1991; Aplan et al., 1992).

Tal-1 deletions are assumed to be caused by the V(D)J recombinase system, because the breakpoints cluster at heptamer and nonamer recombination signal sequences. The fusion regions of the tal-1 deletion breakpoint shows random addition and deletion of germline nucleotides and P-region nucleotides all of which are hallmarks of V(D)J recombination (Blackwell et al., 1989; Caroll et al., 1990).

Tal1-deletions are not detected by classical cytogenetics. Sil contain three donor deletion sites (Sildb1 to Sildb3) of which Sildb1 is most commonly used (98% of cases). TAL1 contains seven acceptor deletion sites (taldb1 to taldb7), with two being involved in almost all cases (taldb1 and taldb2). The vast majority (~90%) of Tal1-deletions are located between Sildb1 and taldb1 and are known as Tal-1 del type I. The tal-1 deletion type 2 occurs between Sildb1 and taldb2 (Breit et al., 1993c).
1.8. QUANTITATION OF MRD IN ALL

The detection of MRD and monitoring the leukemia load during and after therapy of ALL predicts outcome and is important for treatment stratification, antedate relapse and individualization of therapy. In near future, detection of MRD will be an essential part of routine diagnostics in childhood ALL. Several MRD studies in childhood ALL have shown that it is important to determine precisely the level of MRD at early remission time points for discrimination between low-and high-risk patients.

A method that allows sensitive MRD detection in the majority of ALL patients is PCR analysis of clonal rearrangements of T-cell receptor genes and Immunoglobulin. The junctional region of TCR and Ig gene rearrangements are regarded as leukemia-specific DNA fingerprints, owing to the deletion and random insertion of nucleotides during the recombination of various V, D and J gene segments (Macintyre et al., 1990; Van Dongen and Wolvers Tettero, 1991a; Breit and Van Dongen, 1994b). PCR based assessment of MRD allows the detection of leukemic cells with 100 to 100,000-fold higher sensitivity than cytomorphology (Campana and Pui, 1995; Van Dongen et al., 1996).

1.8.1. PCR based methods for MRD detection

During the last decade, most PCR based studies for the detection of MRD using clone specific TCR and Ig gene rearrangements have used semi-quantitative methods such as Dot-blot hybridization or liquid hybridization. For
Dot-blot technique, after PCR amplification for TCR/Ig gene rearrangement, the
PCR product is hybridized with the radioactively labeled patient-specific probe
(Seriu et al., 1995). For liquid hybridization, the PCR product is not fixed to a filter
but hybridized as free DNA (Steenbergen et al., 1995b). Although these MRD-
PCR techniques are highly sensitive, they provide only semi-quantitative data
owing to post-PCR 'end-point' analysis. The PCR technique has the ability to
amplify target DNA up to a plateau but it is impossible to define precisely the
initial amount of target DNA. The methods used to overcome the limitations in
quantitating the DNA are competitive PCR and limiting dilution technique.
Quantitation by competitive PCR is performed by comparing the PCR signal of
the specific target DNA with that of known concentrations of an internal standard,
the competitor (Cross et al., 1993). Quantitation with limiting dilution assay is
done by serial dilution of the target DNA in multiple replicates. The dilution end-
point defines the amount of initial target DNA via Poisson's law. Both approaches
are laborious, require multiple PCR analyses per sample and are therefore
difficult to perform routinely (Ouspenskaia et al., 1995).

1.8.2. RQ-PCR based on TaqMan technology with Hydrolysis Probe

RQ-PCR exploits the 5'-3' exonuclease activity of the Taq Polymerase to
detect and quantify specific PCR products as the reaction proceeds. A major
advantage of using RQ-PCR is that the stage of exponential amplification is
easily identified, thus allowing an accurate quantification. The TaqMan probe is
labeled at the 5' end with a fluorescent reporter dye and at the 3' end with a
quencher dye. When both reporter and quencher are intact, there will not be any fluorescence emission. During strand synthesis, Taq Polymerase cleaves off the 5' end of the reporter and fluorescence is emitted. The amount of fluorescence emitted during each cycle is directly proportional to the amount of target copy numbers. Because of the real-time detection, quantitative data can be obtained in a short period of time, since post-PCR processing is not necessary (Holland et al., 1991; Heid et al., 1996). Several groups have shown that RQ-PCR via TaqMan technology can be used for the quantitative detection of MRD using leukemia-specific chromosome aberrations t(9;22), t(14;18) and t(8;21) as PCR targets (Mensink et al., 1998; Luthra et al., 1998; Marcucci et al., 1998; Pallisgaard et al., 1999; Siraj et al., 2002).

A patient-specific PCR primer or probe is usually designed to junctional region sequence to detect the leukemic cells within the background of normal cells that may have similar gene rearrangements but different junctional regions. During the disease course continuing rearrangements and secondary rearrangements might result in loss of the junctional regions initially identified at diagnosis (Steenbergen et al., 1993; Beishuizen et al., 1994; Pongers-Willemse et al., 1998). Hence, it is important to monitor ALL patients with two or more independent monoclonal Ig/TCR targets to prevent false-negative results during follow-up (Van Dongen et al., 1996).
1.8.3. Detection of MRD for risk stratifying the patients

It is imperative to detect the level of MRD for discrimination between low and high-risk patients (Coustan-Smith et al., 1998; Cave et al., 1998; Van Dongen et al., 1998; Gruhn et al., 1998). This was found especially informative when early remission time-points were analyzed; at later time-points presence of MRD was associated with a high risk of disease recurrence. Low levels or absence of MRD in BM after completion of induction therapy appears to predict good outcome, and the risk of relapse is proportional to detected MRD levels (Brisco et al., 1994; Cave et al., 1998). The results from the large prospective MRD study of the International BFM Study Group indicated that sensitive MRD information about the kinetics of tumor load decrease during the first 3 months of therapy (at the end of induction treatment and before consolidation treatment) by using the kinetics of tumor reduction, it was possible to recognize a low risk group (~43% of patients) with a 3-year relapse rate of only 2%, an intermediate risk group (~43% of patients) with a 3-year relapse of ~25% and a high risk group (~15% of patients) with a 3-year relapse rate of ~75% (Van Dongen, 1998). Multivariate analyses showed that the degree of MRD positivity after induction therapy is the most powerful prognostic factor, independent of other clinically relevant risk factors. Similarly, high levels of MRD positivity (10^{-2} to 10^{-3}) after re-induction treatment in ALL patients are strongly prognostic of relapse regardless of applied treatment (Eckert et al., 2001) and high MRD levels before stem cell transplantation are invariably associated with relapse after transplantation (Knechtli et al., 1998; van der Veldon et al., 2001; Sramkova et al., 2007).
1.8.4. Chromosome aberrations as PCR targets for MRD detection in ALL

Chromosome aberrations can be used as tumor-specific MRD-PCR targets by choosing PCR primers at opposite sides of the breakpoint fusion region. Amplification of such hybrid sequences at the DNA level with PCR is only feasible when the breakpoints cluster in a region of <2kb. Despite the clustering of the breakpoints, the nucleotide sequences of the fusion region of chromosome aberrations differ per patient. Therefore such sequences represent unique patient-specific MRD-PCR targets. Breit et al used tal-1 deletions as PCR target for detection of MRD in T-ALL (Breit et al., 1993c). However, this concerns only 10-25% of T-ALL cases.

In most translocations however, breakpoints occur randomly in a widespread region of >2kb that is difficult to study by DNA-PCR. So, fusion gene is transcribed into fusion gene mRNA that is similar in individual patients despite distinct breakpoints at the DNA level. After reverse transcription into cDNA, these fusion-gene mRNA molecules can be used as appropriate targets for MRD-PCR analysis (Malec et al., 1999; Van Dongen et al., 1999b). Examples include BCR-ABL with t(9;22), TEL-AML1 fusion mRNA with t(12;21) were applied for quantitation studies (Radich et al., 1997; Drunat et al., 2001; Gutierrez et al., 2005).

An advantage of using chromosome aberrations as tumor-specific MRD-PCR targets is their stability during the disease course. Also, because of the cross-contamination of RT-PCR products between patient samples might
severely hamper MRD detection, leading to false positive results. Such cross-contamination is difficult to recognize, since leukemia-specific fusion-gene mRNA products are not patient-specific. An additional (theoretical) disadvantage is that fusion gene transcription might be affected during different phases of chemotherapy, resulting in altered transcript levels (Cazzaniga et al., 2003).

1.8.5. Principles of Real-time Quantification

In RQ-PCR method, a real-time amplification plot is generated. The cycle at which the fluorescence signal exceeds a certain background fluorescence level, referred to threshold cycle (C_T), is directly proportional to the amount of target DNA present in the sample. Based on a two-fold amplification during each PCR cycle, a sample with series of 10-fold dilution should theoretically result in a C_T increase of 3.3 (i.e., 2log10), but in practice the slope of the dilution curve will generally be between 3.2 and 3.9. The amount of residual leukemic cells in follow-up samples obtained during or after treatment can be calculated using the standard curve of the diagnostic sample. Also, a control gene (albumin/β-actin for DNA and Abelson/GAPDH for RNA) should be used to correct for the total amount of DNA and its amplifiability (Pongers-Willemse et al., 1998; Bruggemann et al., 2000; Donovan et al., 2000; Szczepanski et al., 2001b; van der Veldon et al., 2002b; van der Veldon et al., 2003b; Bruggemann et al., 2004).
Fig 1.7. TaqMan Probe Method

For PCR  → Forward primer (FP1) + Reverse Primer (RP1)
For RQ-PCR  → ASO and RP2 + TaqMan Probe
of magnitude (Pongers-Willemse et al., 1998; Verhagen et al., 2000; van der Velden et al., 2002b).

1.8.5.2. Sensitivity of RQ-PCR-based MRD detection

Theoretically the detection limit of the RQ-PCR technique is approximately \(10^{-6}\). Using fusion gene transcripts of chromosome aberrations as MRD-PCR targets, detection limit of \(10^{-4}\) to \(10^{-6}\) can easily be reached in ALL patients. In contrast, the sensitivity of MRD-PCR analysis of junctional regions generally ranges from \(10^{-3}\) to \(10^{-5}\) and is dependent on the type of rearrangement. Junctional regions of complete V-(D)-J rearrangements (e.g. TCR\(\delta\)/IgH) are extensive, whereas junctional regions of V-J rearrangements (e.g. TCR\(\gamma\)) are less extensive.

The sensitivity of MRD-PCR analysis using TCR/Ig junctional region is also influenced by the "background" of normal lymphoid cells showing polyclonal rearrangements. For instance, V\(\delta\)1-J\(\delta\)1 rearrangements frequently occur in T-ALL, but also in a small fraction (0.1-2%) of normal peripheral blood T-cells. V\(\gamma\)1-J\(\gamma\)1.3/2.3 joining comprise 50-60% of TCRG rearrangements in T-ALL, but are also found in a large fraction (70-90%) of normal T lymphocytes. The polyclonal V\(\gamma\)-J\(\gamma\) joinings, particularly in post-induction follow-up samples (Van Wering et al., 2001) is generally less sensitive \((10^{-2}\) to \(10^{-4}\)) than MRD-PCR analysis of long V\(\delta\)1-J\(\delta\)1 junctional regions \((10^{-3}\) to \(10^{-5}\)). A RQ-PCR based quantitation of MRD study by Van der Velden et al using TCRG gene rearrangement reached a
maximal sensitivity of $10^{-4}$ in 50% of patients in precursor-B-ALL and in only two third of patients in T-ALL (van der Veldon et al., 2002b). Similarly, substantial expansions of normal precursor-B-cells with polyclonal IgH rearrangements in regenerating BM after completion of maintenance therapy might affect sensitivity of MRD detection using Ig gene rearrangements as PCR targets. Nevertheless, it should be possible to identify two sufficiently sensitive targets ($\leq 10^{-4}$) for RQ-PCR based MRD detection in at least 80% of patients (Szczepanski et al., 2002a).

1.8.6. Detection of MRD in PB

Two studies found that in most patients with B-lineage ALL, the level of MRD in the PB was at most one tenth that in the BM (Van Rhee et al., 1995; Brisco et al., 1997). However, two other studies found concordant results in BM and PB from T-ALL (Couston-Smith et al., 2002; Van der Veldon et al., 2002a). In a study, MRD was detected in both marrow and blood in 78 pairs of samples, and in the marrow but not in blood in 67 samples; it was undetectable in the remaining 602 pairs out of 747 collected from 231 children (Couston-Smith et al., 2002). An other study showed, concordant results of positivity in 41 BM and PB samples out of 179-paired samples. The results of a recently reported study using immunofluorescence analysis and RQ-PCR are in agreement with these data (van der Veldon et al., 2002a). Moreover, MRD in PB of B-lineage ALL patients at the end of remission induction therapy was associated with a very high risk of disease recurrence (Couston-Smith et al., 2002).
1.9b OBJECTIVES

1. To study the pattern of TCRγ, TCRδ and IgH gene rearrangements and junctional region characteristics in South Indian ALL patients.

2. To quantitate the MRD in ALL using TCRγ, TCRδ and IgH gene rearrangements and patients response to treatment.

3. To identify the prognostic significance of TCRγ and TCRδ gene rearrangements in ALL.

4. To detect the incidence of chromosomal translocations in South Indian Precursor-B-ALL.