INTRODUZIONE

In questi anni di lavoro sperimentale la mia ricerca si e' focalizzata sullo studio del gene *MLL* (*Mixed Lineage Leukemia*) il quale spesso risulta coinvolto in translocazioni cromosomiche ricorrenti che identificano un sottogruppo di leucemie acute con caratteristiche cliniche e biologiche uniche. Questi riarrangiamenti si possono identificare sia nelle leucemie di tipo linfoide che mieloide, e si trovano prevalentemente nelle leucemie infantili e nelle leucemie secondarie dopo trattamento con farmaci inibitori delle topoisomerasiII.

MLL, l'omologo della proteina trithorax (trx) di drosophila, e' una complessa proteina che riveste un ruolo fondamentale nella regolazione della transcrizione genica. In particolare MLL, attraverso la regolazione dei geni omeotici (*Hox*) e di altri geni importanti per la regolazione dello sviluppo, e' essenziale per un corretto sviluppo embrionale e nei processi ematopoietici. Nella leucemia, la normale funzione di MLL viene alterata in seguito alla rottura cromosomica che successivamente a ricombinazione genica forma un prodotto chimerico in cui la parte N-terminale di MLL si fonde con la parte C-terminale di un altro gene partner, di cui se ne conoscono più di 50. Queste proteine chimeriche hanno sempre una funzione anomala spesso in grado di modificare il differenziamento delle cellule ematopoietiche progenitrici dotandole di attivita' leucemogenica attraverso meccanismi ancora largamente sconosciuti.

Il lavoro svolto in questa tesi di dottorato ha indagato se l'osservazione che il riarrangiamento del gene *MLL*, come pure la malattia stessa, possa originare in una cellula indifferenziata staminale ematopoietica; se i punti di rottura a livello genomico del gene *MLL* nella regione conosciuta "breakpoint cluster region (bcr)" in pazienti affetti da leucemia secondaria si distribuiscono in modo causale o in particolari regioni genomiche; se il gene omologo di *MLL* è presente in zebrafish, prerequisito essenziale per utilizzare questo sistema modello dello studio delle funzioni di mll e del processo leucemogenico indotto. In particolare, attraverso la descrizione di un paziente affetto da leucemia pre-pre-B portatore della translocazione t(4;11) è stata investigata l'origine della translocazione. La peculiarita' di questo caso risulta dal fatto che il clone leucemico subiva due

consecutivi cambiamenti immunofenotipici inizialmente LAL dopo il trattamento terapeutico ricadeva come LAM e dopo il trapianto allogenico di midollo osseo ricadeva nuovamente a LAL. La sovraespressione dei geni *HOXA9* e *FTL3* e il mantenimento di un unico pattern genotipico (Ig/TCR, MLL-AF4) durante tutte le fasi della malattia hanno dimostrato il probabile coinvolgimento di una cellula ematopoietica progenitrice di tipo mielo-linfoide all'origine del riarrangiamento. Questa parte di ricerca e' stata condotta mediante analisi morfologica e citofluorimetrica, l'analisi di microarray ha permesso di valutare l'espressione genica durante tutte le fasi della malattia.

In seguito è stata condotta l'analisi della sequenza nucleotidica del gene *MLL* nei punti di rottura e giunzione che ha dimostrato un meccanismo di ricombinazione intercromosomica molto preciso che avviene preferenziamente nella regione 3' terminale dell'introne 8 di *MLL* bcr. Per questa parte di ricerca è stata utilizzata la panhandle PCR, una metodica che permette di clonare entrambe le regioni di giunzione dei punti di rottura del gene *MLL*, indipendentemente dal gene partner di fusione, attraverso la formazione di un loop e successiva amplificazione con primer oligonucleotidi *MLL* specifici. Questa metodica ha consentito inoltre di identificare a livello genomico la fusione tra il gene *ARMC3*, situato sul cromosoma 10p12, e *MLL* in un paziente con leucemia mieloide secondaria. Il gene *ARMC3* codifica per una proteina con funzione sconosciuta caratterizzata dalla presenza di domini proteici denominati "Arm" simili a quelli riscontrati nella famiglia proteica delle catenine, della placoglobina e del soppressore tumorale APC. *ARMC3* e' il primo gene di questo tipo ad essere coinvolto in un riarrangiamento cromosomico con il gene *MLL*. Infine, è stato identificato e clonato il trascritto (cDNA) omologo del gene umano *MLL* di zebrafish dato l'alto grado di conservazione dei processi che governano l'ematopoiesi tra mammiferi e zebrafish e la recente idenfificazione di numerosi importanti fattori transcrizionali omologhi e conservati tra queste due specie. La caratterizzazione di *mll* è stata possibile mediante analisi bioinformatica e una combinazione di analisi molecolari di PCR degenerata, RACE-PCR, long-PCR e PCR convenzionale è stata dimostrata la presenza di un singolo transcritto di 12657 nucleotidi che codifica per una proteina di 4218 residui amminoacidici identica per il 46% con la rispettiva proteina umana di MLL. Il pattern di espressione del supposto trascritto *mll* e' stato esaminato tramite PCR quantitativa e analisi di ibridazione *in situ*.

Questi studi indicano che l'mRNA di *mll* e' inizialmente di origine materna per poi essere espresso durante tutto lo sviluppo embrionale e nell'adulto sottolineandone l'importanza per l'intera vita dell'organismo. Pertanto, questi risultati costituiscono le basi per utilizzare lo zebrafish come sistema modello addizionale per decifrare le funzioni del gene *MLL* nell'ematopoiesi e nella leucemia, e contribuire all'identificazione di nuovi potenziali target terapeutici.

In conclusione, questo dottorato di ricerca ha introdotto nuove conoscenze sul gene *MLL* di elevato significato biologico e quindi potenzialmente clinico. E' stata utilizzata la panhandle PCR che ha permesso la caratterizzazione di nuovi punti di rottura sul gene *MLL* e ha identificato un nuovo gene partner, *ARMC3*. Infine, lo studio ha gettato le basi per l'utilizzo dello zebrafish come modello *in vivo* aprendo una nuova era per la ricerca nella leucemia.

INTRODUCTION

During these years of experimental work, my research focused on the study of *MLL* (*mixed lineage leukemia*) gene, which is one paradigm of leukemogenesis, because it is frequently part of random chromosomal translocations that defines a unique group of acute leukemias in terms of clinical and biological features.

MLL chromosomal translocations are found in leukemias with both lymphoid and myeloid phenotype and are often associated with infant and therapy-related secondary leukemias. In general, leukemias with *MLL* translocations are clinically aggressive and respond poorly to therapy.

MLL, the mammalian homologue of the *Drosophila* protein, trithorax (trx), is a large multi-domain protein that plays a important role in the regulation of transcription. In particular, MLL is essential for both embryonic development and normal hematopoiesis through regulation of the clustered homeobox (*Hox*) genes and other genes important for developmental regulation. In leukemia, MLL function is subverted by rupture, illegitimate ricombination, and the formation of a chimeric fusion in which the N-terminal portion of MLL is fused to the C-terminal portion of over 50 alternative partner genes. The formed MLL chimeric proteins with altered function are able to affect differentation pathways of hematopoietic progenitors and to endow cells leukemogenic activity. However, the mechanism(s) by which MLL fusion proteins induce leukemia is largely unknown.

Research here discussed has been addressed to two different aspects of *MLL* proto-oncogene in acute leukemia, and also the viability of the zebrafish (*Danio rerio*) as a genetic tool to deepen the role of human MLL has been considered.

In particular, this research has regarded: (1) the observation that *MLL* rearrangements and disease itself may initiate within an undifferentiated hematopoietic stem cell; (2) the assessment of *MLL* genomic breakpoints distribution within the breakpoint cluster region (bcr), in particular in secondary leukemias samples, and (3) the identification of orthologue gene coding for *mll* in zebrafish as prerequisite for the study of Mll function, and of MLL-fusion proteins-related leukemia, in this vertebrate model system.

(1) The investigation of the cell in which origin *MLL* translocations, was made by a case of a patient diagnosed with pre-pre-B ALL/t(4;11) leukemia, which during the treatment and after matched bone marrow transplantation, underwent two consecutive switches from lymphoid to myeloid lineage and *vice versa*. The high expression of *HOXA9* and *FLT3* genes remaining genotypically stable in leukemia throughout phenotypic switches, suggests that this leukemia may have originated as a common B/myeloid progenitors. This part of the work has been performed by morphology and flow cytometry analyses combined with microarray analysis, in order to evaluate gene expression during different phases of disease.

(2) The work about the localization of *MLL* genomic breakpoint junction indicates that translocations in treatment related leukemias occur mainly near precise or precise interchromosomal DNA recombination at the sequence level, and confirms a translocation breakpoint hotspot at 3' region in the bcr intron 8. This part of work has been performed by panhandle PCR analysis, a technique that allows the amplification of the *MLL* genomic breakpoint junction from a stem-loop template using primers all from *MLL*. The panhandle PCR also identified a breakpoint junction of the uncovered *ARMC3* from band 10p12 and *MLL* intron 9 in a case of treatment-related myeloid leukemia. ARMC3 protein contains Arm repeats similar to catenin proteins (e.g. b-catenin), plakophilins and the tumor suppressor APC. *ARMC3* is the first gene of this type disrupted by *MLL* translocation.

(3) The overall high-grade conservation of the molecular pathways governing hematopoiesis between mammals and zebrafish, as well as the identification of several well conserved zebrafish transcription factors mammalian orthologs, permitted the cloning of a zebrafish *mll* cDNA. This part of the work has been performed by bioinformatic analysis and a combination of degenerate-PCR, RACE PCR, long-range and conventional PCR strategies. The consensus open reading frame (12657 bp) encodes a polypeptide of 4218 amino acids which shares 46% identity with the human MLL. The temporal pattern of *mll* cDNA expression was examined using quantitative RT-PCR analysis and whole-mount *in situ* hybridization analysis. These findings indicate that there is a single zebrafish *mll* gene with highly conserved functional similarity to human MLL. The temporal pattern of expression indicates that *mll* is important from early embryogenesis through the entire lifespan of the fish. The high evolutionary conservation of critical domains are the starting point to use zebrafish for dissecting the function of

MLL in hematopoiesis and leukemia, and thus, may further contribute to the identification of potential therapeutic targets.

In conclusion, this PhD doctoral thesis brings new insight into the leukemogenic process by novel biological characterization of *MLL* alterations and open for further clinical opportunities. Moreover, this research has showed the panhandle PCR potential to discover new partner genes of *MLL* as the characterization of the uncovered *ARMC3* gene. Finally, the identification and characterization of the zebrafish *mll* ortholog open a new era in leukemia research by using this *in vivo* model.

CHAPTER 1

MLL AN UNUSUALLY PROMISCUOUS ONCOGENE

Acute leukemia is the most common form of childhood cancer representing a third of malignancies and is the primary cause of cancer-related mortality. Acute lymphoblast leukemia (ALL) is more frequent than acute myeloid leukemia (AML), representing 80-85% of all leukemias. Most of these cases present cytogenetic abnormalities which acquire somatic mutations that have been implicated in the etiology of disease[1,2]. They represent the most important prognostic factors helping leukemia subtype classification and therapeutic protocol assignments.

In this chapter the attention will be focused on the translocations that involve the *mixed lineage leukemia* (*MLL*) gene, which identify a unique subtype of acute leukemias. In the second part of this chapter will be discuss a case report in which *MLL* rearrange in an early stage of the hematopoiesis, giving reason of the crucial role of *MLL* abnormalities.

1.1 Clinical features of leukemias with *MLL* translocations

Translocations of the gene *MLL* located at chromosome band 11q23 define a subset of aggressive acute leukemias with distinctive clinical and biological features[3]. As a consequence, *MLL* undergoes fusion with a wide variety of partner genes (up to 50) in which the chimeric proteins that are originated lead to either myeloid or lymphoid leukemias (hence the name MLL). *MLL* translocation are frequently seen in infant leukemias[4] and secondary leukemias following treatment with topoisomerase II poisons[5]. *MLL* translocations are also found in 5-10% of childhood and adult acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML)[5-7]. Other alteration of MLL that have been reported include partial tandem duplication (PTDs) of exons encoding N-terminla sequence of MLL, which occur in approximately 10% of AML cases[8], and amplification of the apparentely un-rearranged MLL gene in rare cases of myelodysplasia and AML[9].

Biologically, *MLL* translocations determine differentiation, lineage and immunophenotype of leukemia blast cell populations. For example, the blast populations in the ALL cases exhibit an early $CD10^-$ $CD24^-$ pro-B cell immunophenotype and uniquely co-express the myeloid associated antigen CD15[10]. *MLL* translocations are strongly associated with myelomonocytic and monoblastic AML in infants and young children[11] and in the treatment-related cases[12,13]; however, leukemias with *MLL* translocations also can present as other AML morphologic subtypes or myelodysplastic syndrome[14-16]. This morphologic and phenotypic heterogeneity is influenced by the partner genes involved[17-19].

In all of these patient populations, *MLL* translocations are poor prognostic factors with significant adverse effects on response to treatment. For an ultra high-risk population within infant ALL, the constellation of poor prognostic features including age <3 months at diagnosis, WBC count >100,000/μL, early pro-B $CD10^-$ immunophenotype and t(4;11) translocation, is associated with event free survival of ~5%[20,21]. Infant leukemias with *MLL* translocations often are resistant to common chemotherapeutic agents[22,23]. Infants also are more vulnerable to toxicities, and more intensive treatment for infant ALL has increased treatment complications without improving outcome[24]. Event free survival rates in infant AML are ~50% using current intensive treatments[25]. *MLL* translocations strongly predict poor clinical outcome and portend a grave prognosis in secondary leukemia also[4]. Prognosis in the secondary cases is affected further by the limited feasibility of administering additional intensive anti-leukemia therapy after primary cancer treatment[26]. Furthermore, *MLL*-rearranged leukemias display a highly distinct gene expression profile that is consistent with an early hematopoietic progenitor expressing select multilineage markers. Clustering algorithms reveal that lymphoblastic leukemias with *MLL* translocations can clearly be separated from conventional acute lymphoblastic and acute myelogenous leukemias[27]. Therefore, on bases of all clinical and molecular features of *MLL*-associated leukemias, nowadays "MLL leukemia" is considered a distinct disease, and not a subtype of more prevalent ALL or AML.

An example of MLL leukemia is described in the case report further reported (section 1.3) which deals of a patient with a pro-B ALL, *MLL-AF4* rearranged[28]. The peculiarity of this patient is the fact that during early induction therapy of ALL, a distinct monocytic blast population progressed while the lymphoblastic

leukemia quickly regressed. We used morphological and immunophenotypic analysis, as well as total RNA investigations by RT-PCR and gene expression profiling to investigate whether the myelo-monocytic blasts were related to the lymphoblastic clone. It was found that the lympoblasts at diagnosis were similar to relapse by morphology, immunophenotype, *MLL* rearrangement, and gene expression; whereas cells during myelo-monocytic blast progression harbored the same *MLL* rearrangement, but showed a very distinct morphologic, immunological, as well as total gene expression profile. We conclude that the leukemic cells originated in a not fully committed *MLL*-rearranged progenitor cell, compatible with leukemic arrest in a precursor with B lymphoid/myeloid potential.

1.2 MLL, MLL Fusion Proteins and Leukemogenesis

The *MLL* gene encodes a large, complex oncoprotein that regulates transcription[29-34]. *MLL* was also originally named *HRX* and *Htrx1* because its speckled nuclear localization (SNL) domains, plant homeodomains (PHDs) and SET domain have regional amino acid similarity to *Drosophila* trithorax (trx)[3,32,35]. *Drosophila* trx group (trxG) and Polycomb-group (PcG) proteins, respectively, maintain expression or repression of homeotic gene complexes during embryonic development[36,37]. The trxG proteins are not required for transcription initiation but maintain transcription through later stages of development[38]. MLL and BMI-1, mammalian homologues of trxG and PcG proteins, are antagonistic regulators of *HOX* gene expression[38]. MLL maintains *HOX* gene expression during skeletal, craniofacial and neural development and hematopoiesis[36,39,40].

The *MLL* gene spans 90-kb of chromosome 11q23, has 36 exons and codes for a protein of 3696 amino[31,32,41] (fig. 1.1.2). Taspase1 cleaves MLL into an amino terminal fragment (MLL^N, 320-kDa) with transcriptional repression properties and a carboxyl terminal fragment (MLL^C, 180-kDa) with transcriptional activation properties, which associate with one another and other chromatin regulatory proteins in a large multi-protein complex[42,43] (fig. 1.1.2). MLL proteolytic cleavage by taspase1 and association of its N and C terminal fragments is critical for proper nuclear sublocalization and *HOX* gene regulation[43].

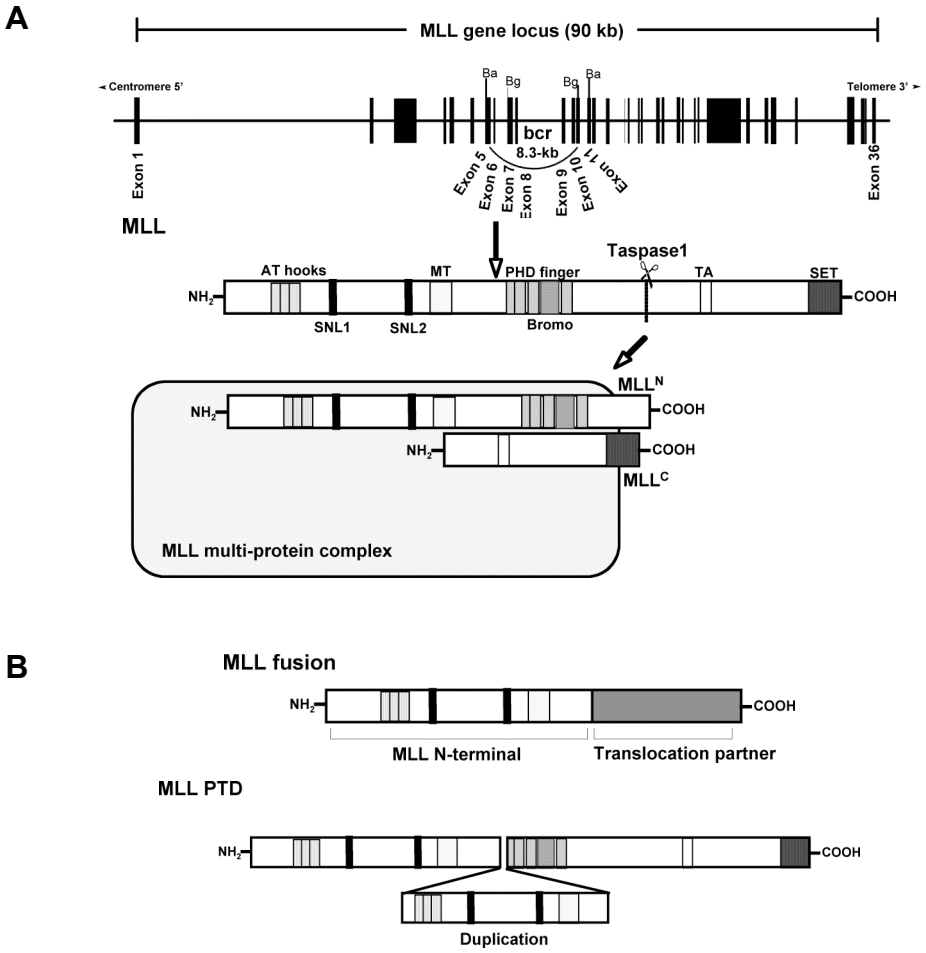


Figure 1.1.2. Schematic representation of the *MLL* gene, MLL protein and MLL fusions. A) The MLL gene is approximately 90-kb long, consists of 36 exons, and encodes a 3,969 amino acid nuclear protein with a complex structure. Narrow lines represent the introns, and vertical narrow boxes/lines represent exons. bcr (breakpoint cluster region) is represented by the semicircular line that encompasses exon 5-11. MLL protein is proteolytically cleaved by Taspase1 into N-terminal (MLL^N, 300 kDa) and C-terminal (MLL^C, 180 kDa) fragments that interact with each other to stabilize the complex. MLL is a component of a multiprotein complex with at least 27 proteins. The vertical arrow indicates the location of the breakpoint that occurs in leukemia wit *MLL* gene translocations. B) Structure of MLL fusion proteins generated by MLL translocations. N-terminal portion of MLL is fused to the C-terminal portion of the fusion partner protein (MLL fusion) or internally duplicated (MLL PTD, partial tandem duplication) in leukemia. Abbreviations: SNL, speckled nuclear localization; MT, DNA methyltransferase homology domain; PHD, plant homeodomain zinc fingers; bromo, bromodomain; TA, transcriptional-activation domain; SET, su(var)3-9, enhancer of zeste, trithorax domain; Ba, *Bam*HI; Bg, *Bgl*II

Although there has been substantial research on the functional motifs of MLL, the roles of MLL are incompletely understood. Constructs comprising MLL AT hook motifs have been shown to promote p21 and p27 upregulation, cell cycle arrest and monocyte differentiation[44]. The amino terminal SNL motifs direct MLL subnuclear localization[35]. The cysteine-rich CXXC region is similar to the

CXXC region in DNA methyltransferase 1 that recognizes CpG di-nucleotides[45]. The MT domain is part of a transcriptional repression region[35,42,44,46]. The PHD mediates MLL homodimerization and protein interactions including binding to a nuclear cyclophilin, which modulates target gene expression[47]. MLL possesses a transcriptional activation domain (TA) between the PHD fingers and C-terminal SET domain. TA domain binds to CBP (CREB-binding protein), a histone acetylase that promotes transcriptional activation by acetylating histones H3 and H4 at arget gene loci[48]. The SET domain interacts with the SWI/SNF chromatin remodeling complex, which activates transcription[49]. Consistent with its role in epigenetic gene regulation, the SET domain has specific histone H3 lysine-4-specific methyltransferase activity that regulates *HOX* promoters[50]. The MLL protein is a component of a large multiprotein supercomplex with at least 27 proteins, including components of the human transcription complexes TFIID, SWI/SNF, NuRD, hSNFsH and Sin3A[51]. The complex acetylates, deacetylates, and methylates nucleosome-attached histones, resulting in chromatin remodeling.

MLL translocations disrupt an 8.3 kb breakpoint cluster region between exons 5-11 and involve >50 partner genes that encode diverse partner proteins[52-55]. Many MLL partner proteins have structural motifs of nuclear transcription factors[31,32,57-65], transcriptional regulatory proteins[18,66-68] or other nuclear proteins[69-71]. Others are cytoplasmic proteins[72-91], cell membrane proteins or proteins in different cellular locations[92-96]. Table 1.1.2 lists several of the cloned MLL fusion partners. *MLL* also undergoes self-fusions and MLL itself is a partner protein[62,97,98]. While some *MLL* partner genes are members of the same gene families[54,55,59,86-89,99,100] or encode proteins with otherwise similar functions[60,65,101,102], there is no unifying functional relationship between the many partner genes. The most frequently found fusion genes are MLL-AF9/t(9;11), MLL-AF4/t(4;11) and MLL-ENL/t(11,19)[103]. In ALL, the partner genes are limited and *AF4* is the most common, whereas in AML the partner genes are much more diverse. The partner genes in *de novo* and treatment-related leukemias are at least partially overlapping. Of interest also is that some of the MLL partner proteins such as AF4 and AF9 interact with one another[104]. MLL fusion proteins, which retain the AT-hook, SNL and MT domains of MLL but replace the MLL PHD, transactivation, and SET domains with the carboxyl partner protein, transform hematopoietic progenitors and cause leukemia in mice[49,105-110]. The MLL fusion protein also lacks the taspase1

proteolytic cleavage site and cannot interact with the MLL C terminus fragment[42,43]. Murine models of MLL fusion oncoproteins have suggested that the function of nuclear partner proteins involves transcriptional activation[35,102,111], whereas cytoplasmic partner proteins result in forced MLL dimerization or oligomerization[112]. Murine models have also demonstrated that MLL fusion proteins constitutively activate *Hoxa9* and that *Hoxa9* activation is essential for leukemogenesis with some MLL fusion proteins (*e.g.* MLL-ENL)[113]. However, altered *Hox* expression influences phenotype, latency and penetrance, but is not essential for leukemogenesis with other MLL fusion proteins (e.g. MLL-AF9, MLL-GAS-7)[114,115].

Table 1.1.2. MLL fusion Partners

Gene	Chromosome Locus	Function	Localization
AF4	4q21	Transcriptional activator	N
AF6	6q27	Maintenance of cell-cell junction and cell polarity	C
AF9	9p22	Transcriptional factor	N
AF10	9p22	Transcriptional factor	N
ELL	19p13.3	RNA polymerase II transcription elongation factor	N
ENL	19p13.3	Transcriptional activator	N
AFX	Xq13	Forkhead transcriptional factor	N
Septin6	Xq22	Septin family	N
AF1p	1p32	Regulation of endocytosis	C, N
AF1q	1q21	?	?
LAF4	2q11	Transcriptional activator	N
AF3p21	3p21	?	?
GMPS	3q25	Guanosine monophosphate synthetase	C
LPP	3q28	Regulation of cell motility and focal adhesion	C, N
AF5q31	5q31	?	?
GRAF	5q31	Negative regulator of RhoA	C
FKHRL1	6q21	Forkhead transcriptional factor	N
DAB2IP	9q34	Ras GTPase-activating protein	C
FBP17	9q34	?	?
ABI1	10p11.2	?	?
LCX	10q22	?	?
CALM	11q14-q21	Regulation of endocytosis	C, N
LARG	11q23.3	Activator of Rho GTPases	C
CBL	11q23.3	Negative regulator of receptor tyrosine kinases	C
GPHN	14q24	Gly and GABA receptor assembly	C
AF15q14	15q14	?	?
MPFYVE	15q14	?	?
CBP	16p13	Transcriptional coactivator	N
GAS7	17p13	Actin assembly	C
AF17	17q21	Transcriptional factor	N
LASP1	17q21	?	C
MSP	17q25	?	C
EEN	19p13.3	?	C
hCDCrel	22q11	?	C
p300	22q13	Transcriptional activator	N

Abbreviations: N, nuclear localization; C, cytoplasmic localization; (?), not known.

In infant leukemias the *MLL* translocation is an acquired, *in utero* alteration and there is a short latency to the diagnosis of leukemia during the first year of life[86,116,117]. In treatment-related leukemias with *MLL* translocations the typical latency is about two years after the chemotherapy exposure[14]. Latency to leukemia in patients and in mice has suggested that secondary alterations may be important in addition to the translocations for leukemia to occur[35,54].

Although the oncogenesis of *MLL*-associated leukemias is very complicated, several studies suggest that MLL fusion proteins transform by a gain-of-function rather than a loss-of-function mechanisms, in which the replacement of MLL sequence with domains of the fusion partner is a crucial step to cause leukemia. In addition, the translocation product likely affects normal cellular mechanisms, by direct action on *HOX* gene expression occurring at the level of chromatin-mediated activation. However, more efforts may be done to complete the molecular picture of MLL-mediated leukemia, for example the discovery of the function of many *MLL* partner genes that is uncompleted, as well as the characterization of genes and pathways that are under the control of MLL fusion proteins. Expectations that this work will facilitate the understanding of *MLL*-associate leukemias are supported by this research, with the aim to develop successful and novel MLL-directed therapies.

1.3 Two Consecutive Immunophenotypic Switches in a Child with MLL-rearranged Acute Lymphoblastic Leukemia

(Germano *et al.* Haematologica 2006)

Abstract

An 18-month-old girl was diagnosed with pre-pre-B ALL/t(4;11) leukemia, which during the treatment and after matched bone marrow transplantation (BMT) underwent two consecutive switches from lymphoid to myeloid lineage and *vice versa*. The high expression of *HOXA9* and *FLT3* genes in a genotypically stable leukemia throughout the phenotypic switches, suggest that this leukemia may have originate in a common B/myeloid progenitors.

Case report

An 18-month-old girl was admitted to the San Giovanni Rotondo " Casa Sollievo della Sofferenza" IRCCS Hospital, who presented with a WBC of 201,000/mm^3 (80% blasts), hemoglobin of 7.1 g/dL, and platelet count of 42,000/mm^3. After obtaining informed consent from the patient's parents, a bone marrow aspirate was performed. The sample was submitted to our Onco-Hematology laboratory (the reference laboratory of all cases of acute leukemia enrolled in the Associazione Italiana Ematologia Oncologia Pediatrica [AIEOP] protocols) for morphologic, cytochemical, cytogenetic, molecular and flow cytometric evaluations. The bone marrow (BM) aspirate contained 95% of tumor cells having the morphology of lymphoblasts (French-American-British [FAB] L1) (fig. 1.1.3A). Flow cytometric immunophenotypic analysis of blast cells showed: CD19+, CD34+, CD45+, CD133+, CD135+, NG2+, TdT+, HLA-DR (bright) (table 1.1.3). Conventional cytogenetic analysis revealed 46, XX, t(4;11)(q21;q23) in all 20 metaphases analyzed. RT-PCR analysis showed the presence of two amplified *MLL-AF4* fusion transcripts (fig. 2.1.3A). The diagnosis of pre-pre-B ALL was made.

The insufficient clearance of circulating blasts (36,000/mm^3) on day 8 after a steroid prophase (prednisone poor response; PPR) and despite an initial relevant cytoreduction (4,000/mm^3 on day 5), the girl was assigned to the high-risk group in accordance with the AIEOP-BFM-ALL-2000 protocol. Thirteen days later the WBC count increased a second time to 96,600/mm^3, the hemoglobin level was 9.8 g/dL, platelets 56,000/mm^3. Peripheral blood and BM examination revealed respectively 98% and 85% blasts characteristic of monoblastic features and consistent with M5 morphology. Cytochemistry was strongly positive for myeloperoxidase (fig. 1.1.3B,C). Flow cytometric analysis of BM showed: CD15 (bright), CD33 (bright), CD45+, CD64+, CD135+, HLA-DR (dim) but CD19-, CD34-, TdT- (table 1.1.3). The results were consistent with a typical AML (FAB M5) diagnosis. The translocation 46, XX, t(4;11)(q21;q23) was still present (20 of 20 analyzed metaphases). Consequently, her treatment was changed to the AIEOP-AML-2002 protocol. Five months later, after complete remission was achieved, she received allogeneic BMT from an HLA-matched normal donor. Two months after the BMT the patient again relapsed. The leukemic cells displayed morphology (fig. 1.1.3D) and immunophenotype (table 1.1.3) of the initial

diagnosis of ALL. Cytogenetic analysis again revealed a 46, XX, t(4;11)(q21;q23) in all 20 metaphases analyzed. She died a few weeks later without responding to any further treatment forms.

DNA and RNA from BM samples were collected throughout the disease course (diagnosis, myeloid phase and relapse) of this patient. Samples were studied for Ig/TCR clonality evaluation by PCR heteroduplex analysis, and for gene expression profiling. A heteroduplex and homoduplex bands corresponding respectively to VH-JH and Vg-Jg gene rearrangements were identified from the diagnosis, myeloid phase and relapse samples (fig. 2.1.3B). The sequence analysis revealed an identical VH4-JH6, VH4-JH4 and Vg11-Jg1.3/2.3 junctions region between the specimens analyzed[1]. Additional RT-PCR analyses showed the presence of two *MLL-AF4* amplified products also during the myeloid phase and relapse (fig. 2.1.3A). The sequencing of these PCR products showed two alternatively spliced *MLL-AF4* transcripts joining *MLL* exon 10 to either *AF4* exon 4 or exon 5. Gene expression analysis confirmed the MLL signature for each disease phase analyzed (ALL and AML). Furthermore, a closer examination of these genes showed a significant over-expression of *HOXA9* and *FLT3* gene targets during all phases of the acute leukemia (data not shown).

Several hypotheses have been suggested to explain lineage conversion in acute leukemia, but its precise mechanism remains unclear[2,3]. We have described a case of an acute leukemia in which the blast cells rapidly changes lineage from pre-pre-B ALL to AML after 13 days of high-risk AIEOP-BFM chemotherapy. The patient achieved a complete remission by conventional AML-type treatment which included BMT. Eight months later she relapsed again and the blast cells showed a clear return to lymphoid B-cell phenotype. Analyzing their characteristic, as shown in table 1.1.3, the Ig/TCR gene rearrangements and the cytogenetic t(4;11) abnormalities demonstrated that the leukemic cells switched throughout each disease phase maintaining the same clonal relationship. Re-examining the blast cells (85%) at 13 days showed that $<1x10^{-3}$ were CD34+/CD19+, indicating that it was unlikely that this cells contaminated the myeloid phase. Therefore, the CD34+/CD19+ lymphoid-restricted cells at diagnosis were not fully B-lymphoid committed but also able to differentiate into monocytoid blasts retaining the Ig/TCR and *MLL-AF4* leukemic-specific rearrangements. In addition, the expression of the progenitor/stem cell-related markers, such as CD133 and CD34,

only during the lymphoid phase could indicate that an immature lymphoid progenitor develops the potentiality to address different lineages. Gene expression studies of *MLL*-rearranged ALL demonstrate that these leukemias represent a unique disease when compared to other ALLs[4,5]. Moreover, the differences in gene expression support the hypothesis that the cell of origin of MLL is an early hematopoietic progenitor with both myeloid and B-lymphoid potential[4,6]. In our patient this is further supported by the high level expression of *HOXA9* and *FLT3* genes in all phases of acute leukemia. HOXA9 and FLT3 are expressed in early hematopoietic progenitors and both are necessary for the appropriate expansion of the hematopoietic stem cells[7-9]. Taken this hypothesis together with our finding of related Ig/TCR and t(4;11)/*MLL-AF4* gene rearrangements during the lineage switches, this indicates that the leukemia population in our patient could have originated from a common B/myeloid progenitors with the capacity to differentiate into committed cells of either lymphoid or myeloid lineage. In contrast, the translocation t(4;11) that leads to a *MLL-AF4* fusion gene has been preferentially associated with B-cell phenotype lineage[10], but in this case the leukemic clone retain the possibility to induce both myeloid or lymphoid gene expression, suggesting that an immature progenitor/stem cell may be the target of the chromosomal translocation. Therefore, an interpretation of this case is that the *MLL*-rearranged leukemic clone is able to differentiate as lymphoid and myeloid under therapeutic effects by amplifying or suppressing the normal differentiation programs for their survival/expansion.

In conclusion this report confirms that some forms of acute leukemia may arise from very immature cells belonging to a common myeloid/lymphoid progenitor. Moreover, it provides further information into the mechanism of leukemic lineage switches and underlines that it could be useful to test new therapeutic protocols fit to these particular severe leukemias.

References

1. Germano G, del Giudice L, Palatron S, Giarin E, Cazzaniga G, Biondi A, Basso G. *Leukemia* 2003;17:1573-82.
2. Ridge SA, Cabrera ME, Ford AM, Tapia S, Risueno C, Labra S, et al. *Leukemia* 1995; 9: 2023-26
3. Bierings M, Szczepanski T, van Wering ER, Willemse MJ, Langerak AW, Revesz T, van Dongen JJ. *Br J Haematol*. 2001;113:757-62.

4. Armstrong SA, Staunton JE, Silverman LB, Pieters R, den Boer ML, Minden MD, et al. *Nat Genet*. 2002;30:41-7.
5. Kohlmann A, Schoch C, Dugas M, Schnittger S, Hiddemann W, Kern W, Haferlach T. *Leukemia*. 2005;19:953-64.
6. Armstrong SA, Korsmeyer SJ. *Cell Cycle*. 2003;2:408-9.
7. Lawrence HJ, Christensen J, Fong S, Hu YL, Weissman I, Sauvageau G et al. *Blood*. 2005;106:3988-94.
8. Gilliland DG, Griffin JD. *Blood*. 2002 100:1532-42.
9. Ayton PM, Cleary ML. *Genes Dev*. 2003;17:2298-307.
10. Johansson B, Moorman AV, Haas OA, Watmore AE, Cheung KL, Swanton S, Secker-Walker LM. *Leukemia*. 1998;12:779-87.

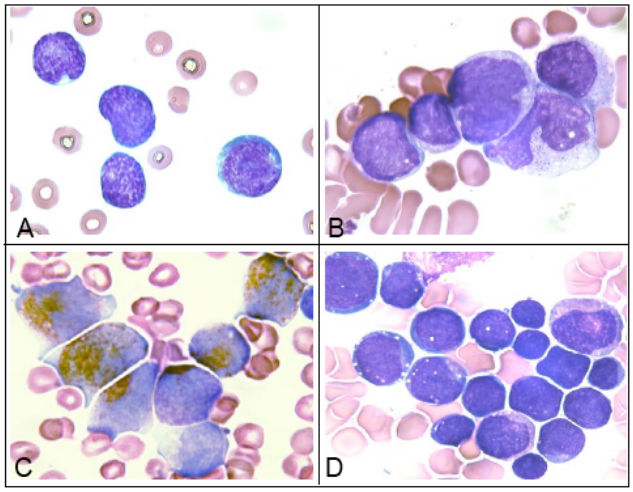


Figure 1.1.3. Morphologic analysis of the leukemia cells in bone marrow smears. (A) Blast cells at diagnosis. (B) Blast cells at day 13 when ALL switched to AML and (C) cytochemical myeloperoxidase staining. (D) Blast cells at relapse (Wright-Giemsa (1000).

Table 1.1.3. Summary of cytomorphological, immunophenotypic and immunogenotypic features during the three phases of acute leukemia.

	At diagnosis (ALL)	**At myeloid phase (day 13) (AML)**	**At relapse (ALL)**
Cytomorphology			
FAB classification	L1	M5	L1
% BM blast	95	85	60
Myeloperoxidase	Negative	Positive	Negative
Esterase	Negative	Positive	Negative
Immunophenotype			
CD19	+	-	+
CD10	-	-	-
CD15	-	bright	dim
CD33	-	bright	-
CD45	+	+	+
CD64	-	+	-
CD133	+	-	+
CD34	+	-	+
CD135	+	+	+
TdT	+	-	+
NG2	+	-	+
HLA-DR	bright	dim	bright
Immunogenotype			
Ig/TCR gene rearrangements	VH4-JH6, VH4-JH4 (he); Vγ11-Jγ1.3/2.3 (ho)	VH4-JH6, VH4-JH4 (he); Vγ11-Jγ1.3/2.3 (ho)	VH4-JH6, VH4-JH4 (he); Vγ11-Jγ1.3/2.3 (ho)
Karyotype	46, XX, t(4;11)(q21;q23)	46, XX, t(4;11)(q21;q23)	46, XX, t(4;11)(q21;q23)
MLL gene rearrangements	*MLL*(ex10)-*AF4* (ex4) *MLL* (ex10)-*AF4* (ex5)	*MLL*(ex10)-*AF4* (ex4) *MLL* (ex10)-*AF4* (ex5)	*MLL*(ex10)-*AF4* (ex4) *MLL* (ex10)-*AF4* (ex5)

Abbreviations: FAB, French-American-British classification; (he), heteroduplex rearrangement; (ho), homoduplex rearrangement; ex, exon. (*MLL* gene rearrangement was alternatively spliced).

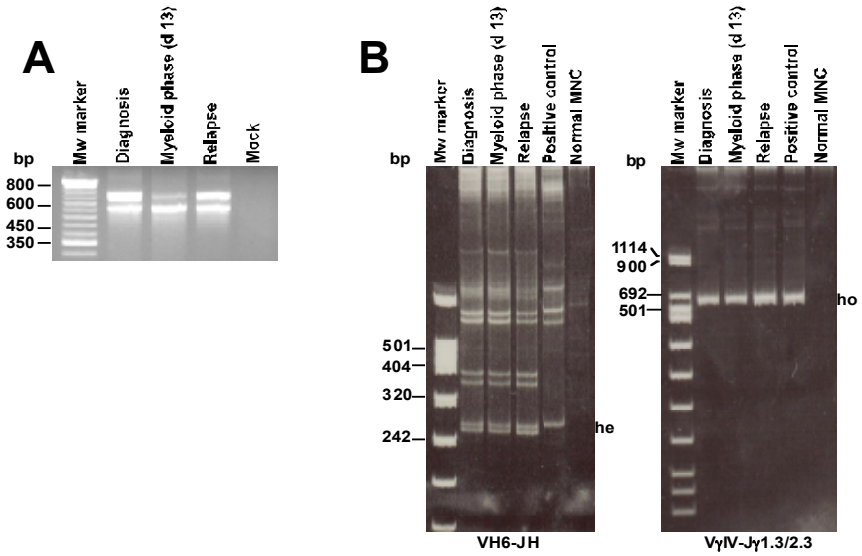


Figure 2.1.3. (A) RT-PCR analysis of the *MLL-AF4* fusion gene showing expression of two different products in diagnosis, myeloid phase (day 13) and relapse RNA leukemic cells. (B) Heteroduplex PCR analysis of *IGH* and *TCRG* gene rearrangements. Blast cells at diagnosis as well as myeloid phase and relapse contained a biallelic heteroduplex (he) VH6-JH and a monoclonal homoduplex (ho) VgIV-Jg1.2/2.3 gene rearrangement. Mw marker, molecular weight marker; Mock, negative control; Normal MNC, normal mononuclear cells.

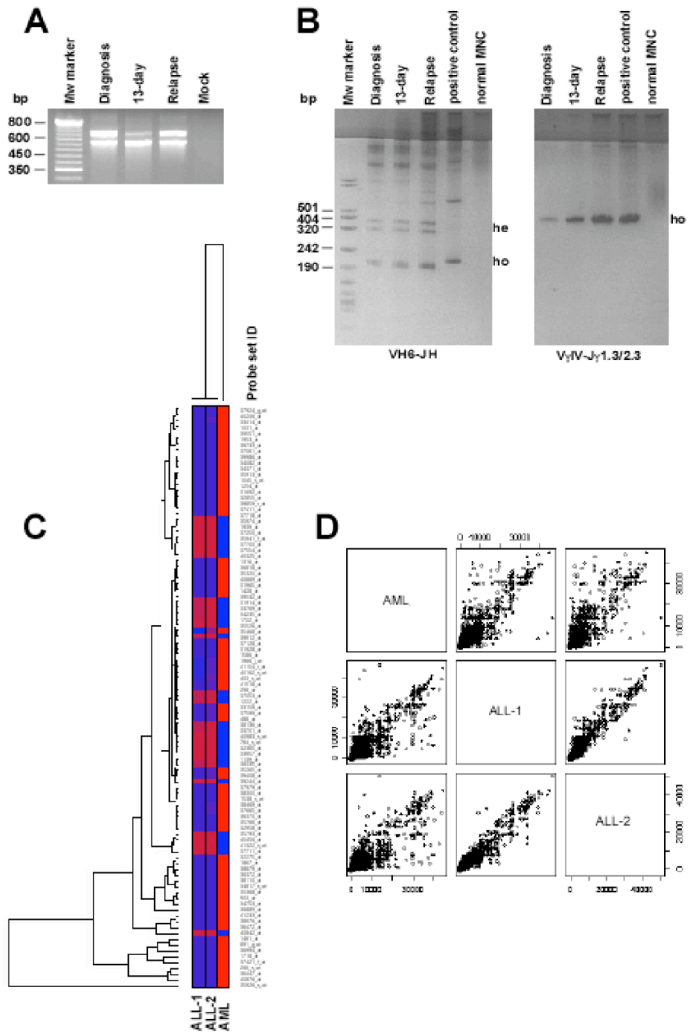


Figure 2.1.3. (A) RT-PCR analysis of the *MLL-AF4* fusion gene showing expression of two different products in diagnosis, 13-day and relapse RNA leukemic cells. Total RNA was extracted from BM cells of diagnosis, 13-day and relapse using TRIzol((Life Technologies, NY, USA). Integrity of the RNA was assessed by capillary electrophoresis (Agilent Tecnologies, Palo Alto, CA, USA). (B) Heteroduplex PCR analysis of *IGH* and *TCRG* gene rearrangements. Blast cells at diagnosis as well as 13-day and relapse contained a biallelic heteroduplex (he) V_H6-J_H and a monoclonal homoduplex (ho) V(IV-J(1.2/2.3 gene rearrangement. PCR products were detected by silver staining on a 12% non-denaturing polyacrylamide minigel. Mononuclear cells were separated by Ficoll-Paque centrifugation and DNA was extracted and purified using Gentra kit (Gentra System, Minneapolis, MN, USA). Mw marker, molecular weight marker; Mock, negative control; normal MNC, normal mononuclear cells. (C) Hierarchical clustering based on HG-U95Av2 expression data of diagnosis ALL (ALL-1), 13-day (AML) and relapse (ALL-2). The genes used in this analysis are the top 100 genes chosen by t-test statistic that are the most differentiated among the tumor samples. The normalized expression value for each gene is indicated by color, with red representing high expression and blue representing low expression. cRNA was prepared according to the standard Affymetrix protocol (BioArray(High Yield RNA Transcript Labeling Kit; Enzo Diagnostics, Farmingdale, NY, USA). Expression values were determined using Affymetrix MAS 5.0 software. (D) Scatter plots indicating statistical correlations among the samples on the basis of their gene expression values. Tighter the dot cloud around the diagonal stronger the correlation among samples. The correlation between ALL-1 and ALL-2 is stronger than between ALL and AML samples.

General references

1 Rabbitts TH. Chromosomal translocations in human cancer. Nature 1994;372(6502):143-9.372.

2 Look AT. Oncogenic transcription factors in the human acute leukemias. Science 1997;278(5340):1059-64.

3 Gilliland DG, Jordan CT, Felix CA. The molecular basis of leukemia. Hematology (Am Soc Hematol Educ Program) 2004:80-97.

4 Rowley JD, Olney HJ. International workshop on the relationship of prior therapy to balanced chromosome aberrations in therapy-related myelodysplastic syndromes and acute leukemia: overview report. Genes Chromosomes Cancer 2002;33(4):331-45.

5 Pui CH, Chessells JM, Camitta B, et al. Clinical heterogeneity in childhood acute lymphoblastic leukemia with 11q23 rearrangements. Leukemia 2003;17(4):700-6.

6 Bacher U, Kern W, Schnittger S, Hiddemann W, Haferlach T, Schoch C. Population-based age-specific incidences of cytogenetic subgroups of acute myeloid leukemia. Haematologica 2005;90(11):1502-10.

7 Mancini M, Scappaticci D, Cimino G, et al. A comprehensive genetic classification of adult acute lymphoblastic leukemia (ALL): analysis of the GIMEMA 0496 protocol. Blood 2005;105(9):3434-41.

8 S.A Schichman, et al., *Proc. Natl. Acad. Sci USA*. **91**, (1994), 6236.

9 B. Poppe, et al., Blood. 103, (2004), 229.

10 Borkhardt A, Wuchter C, Viehmann S, et al. Infant acute lymphoblastic leukemia - combined cytogenetic, immunophenotypical and molecular analysis of 77 cases. Leukemia 2002;16(9):1685-90.

11 Pui CH, Kane JR, Crist WM. Biology and treatment of infant leukemias. Leukemia 1995;9(5):762-9.

12 Ratain MJ, Kaminer LS, Bitran JD, et al. Acute nonlymphocytic leukemia following etoposide and cisplatin combination chemotherapy for advanced non-small-cell carcinoma of the lung. Blood 1987;70(5):1412-7.

13 Pui CH, Kalwinsky DK, Schell MJ, Mason CA, Mirro J, Jr., Dahl GV. Acute nonlymphoblastic leukemia in infants: clinical presentation and outcome. J Clin Oncol 1988;6(6):1008-13.

14 Smith MA, Rubenstein L, Ungerleider RS. Therapy-related acute myeloid leukemia following treatment with epipodophyllotoxins: estimating the risks. Med Pediatr Oncol 1994;23:86-98.

15 Felix CA, Hosler MR, Winick NJ, Masterson M, Wilson AE, Lange BJ. *ALL-1* gene rearrangements in DNA topoisomerase II inhibitor-related leukemia in children. Blood 1995;85:3250-6.

16 Winick N, McKenna RW, Shuster JJ, et al. Secondary acute myeloid leukemia in children with acute lymphoblastic leukemia treated with etoposide. J Clin Oncol 1993;11:209-17.

17 Hunger SP, Sklar J, Link MP. Acute lymphoblastic leukemia occurring as a second malignant neoplasm in childhood: report of three cases and review of the literature. J Clin Oncol 1992;10:156-63.

18 Sobulo OM, Borrow J, Tomek R, et al. MLL is fused to CBP, a histone acetyltransferase, in therapy related acute myeloid leukemia with a t(11;16)(q23;p13.3). Proc Natl Acad Sci, USA 1997;94:8732-7

19 Rowley JD, Reshmi S, Sobulo O, et al. All patients with t(11;16)(q23;p13.3) that involve *MLL* and *CBP* have treatment-related hematologic disorders. Blood 1997;90:535-41.
20 Reaman G, Sposto R, Sensel M, et al. Treatment outcome and prognostic factors for infants with acute lymphoblastic leukemia treated on two consecutive trials of the Children's Cancer Group. J Clin Oncol 1999;17(2):445-55.
21 Reaman G, Zeltzer P, Bleyer WA, et al. Acute lymphoblastic leukemia in infants less than one year of age: A cumulative experience of the Children's Cancer Study Group. J Clin Oncol 1985;3:1513-21.
22 Pieters R, den Boer ML, Durian M, et al. Relation between age, immunophenotype and *in vitro* drug resistance in 395 children with acute lymphoblastic leukemia-implications for treatment of infants. Leukemia 1998;12:1344-8.
23 Pui CH, Gaynon PS, Boyett JM, et al. Outcome of treatment in childhood acute lymphoblastic leukaemia with rearrangements of the 11q23 chromosomal region. Lancet 2002;359(9321):1909-15.
24 Hilden JM, Dinndorf PA, Meerbaum SO, et al. Analysis of prognostic factors of acute lymphoblastic leukemia in infants: report on CCG 1953 from the Children's Oncology Group. Blood 2006;108(2):441-51.
25 Woods WG, Neudorf S, Gold S, et al. A comparison of allogeneic bone marrow transplantation, autologous bone marrow transplantation, and aggressive chemotherapy in children with acute myeloid leukemia in remission. Blood 2001;97(1):56-62.
26 Barnard DR, Lange B, Alonzo TA, et al. Acute myeloid leukemia and myelodysplastic syndrome in children treated for cancer: comparison with primary presentation. Blood 2002;100(2):427-34
27 S.A. Armstrong, T.R. Colub, S.J. Korsmeyer, *Seminars in Hematology*. 40,
28 G. Germano, M. Pigazzi, L. del Giudice, et al., *Haematologica*. 91, (2006), ECR09.
29 Rasio D, Schichman SA, Negrini M, Canaani E, Croce CM. Complete exon structure of the *ALL1* gene. Cancer Research 1996;56:1766-9.
30 Djabali M, Selleri L, Parry P, Bower M, Young BD, Evans GA. A *Trithorax*-like gene is interrupted by chromosome 11q23 translocations in acute leukemias. Nature Genetics 1992;2:113-8.
31 Gu Y, Nakamura T, Alder H, et al. The t(4;11) chromosome translocation of human acute leukemias fuses the *ALL-1* gene, related to Drosophila *Trithorax*, to the *AF-4* gene. Cell 1992;71:701-8.
32 Tkachuk DC, Kohler S, Cleary ML. Involvement of a homolog of Drosophila *Trithorax* by 11q23 chromosomal translocations in acute leukemias. Cell 1992;71:691-700.
33 Ma Q, Alder H, Nelson KK, et al. Analysis of the murine All-1 gene reveals conserved domains with human ALL-1 and identifies a motif shared with DNA methyltransferases. Proc Natl Acad Sci USA 1993;90:6350-4.
34 Domer PH, Fakharzadeh SS, Chen C-S, et al. Acute mixed-lineage leukemia t(4;11)(q21;q23) generates an *MLL-AF4* fusion product. Proc Natl Acad Sci USA 1993;90:7884-8.
35 Ayton PM, Cleary ML. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. Oncogene 2001;20(40):5695-707.

36 Yu BD, Hanson RD, Hess JL, Horning SE, Korsmeyer SJ. MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. Proc Natl Acad Sci U S A 1998;95(18):10632-6.
37 Mahmoudi T, Verrijzer CP. Chromatin silencing and activation by Polycomb and trithorax group proteins. Oncogene 2001;20(24):3055-66.
38 Hanson RD, Hess JL, Yu BD, et al. Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. Proc Natl Acad Sci U S A 1999;96(25):14372-7.
39 Yu BD, Hess JL, Horning SE, Brown GAJ, Korsmeyer SJ. Altered *Hox* expression and segmental identity in *Mll*-mutant mice. Nature 1995;378:505-8.
40 Hess JL, Yu BD, Li B, Hanson R, Korsmeyer SJ. Defects in yolk sac hematopoiesis in Mll-null embryos. Blood 1997;90(5):1799-806.
41 K. Nilson, et al., Br. J. Haematol. 93, (1996), 966.
42 Yokoyama A, Kitabayashi I, Ayton PM, Cleary ML, Ohki M. Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties. Blood 2002;100(10):3710-8.
43 Hsieh JJ, Cheng EH, Korsmeyer SJ. Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. Cell 2003;115(3):293-303.
44 Caslini C, Shilatifard A, Yang L, Hess JL. The amino terminus of the mixed lineage leukemia protein *(MLL*) promotes cell cycle arrest and monocytic differentiation. PNAS 2000;97(6):2797-802.
45 Lee JH, Voo KS, Skalnik DG. Identification and characterization of the DNA binding domain of CpG-binding protein. J Biol Chem 2001;276(48):44669-76.
46 Xia ZB, Anderson M, Diaz MO, Zeleznik-Le NJ. MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein. Proc Natl Acad Sci U S A 2003;100(14):8342-7.
47 Fair K, Anderson M, Bulanova E, Mi H, Tropschug M, Diaz MO. Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. Mol Cell Biol 2001;21(10):3589-97.
48 Ernst P, Wang J, Huang M, Goodman RH, Korsmeyer SJ. MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. Mol Cell Biol. 2001;21(7):2249-58.
49 Rozenblatt-Rosen O, Rozovskaia T, Burakov D, et al. The C-terminal SET domains of ALL-1 and TRITHORAX interact with the INI1 and SNR1 proteins, components of the SWI/SNF complex. Proceedings of the National Academy of Sciences of the United States of America 1998;95(8):4152-7.
50 Milne TA, Briggs SD, Brock HW, et al. MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell 2002;10(5):1107-17.
51 Nakamura T, Mori T, Tada S, Krajewski W, et al. ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol Cell. 2002;10(5):1119-28.
52 Rowley JD. The critical role of chromosome translocations in human leukemias. Annu Rev Genet 1998;32:495-519.
53 Felix CA. Acute Lymphoblastic Leukemia in Infants. In: Pediatric Acute Lymphoblastic Leukemia: Challenges and Controversies in 2000.

Hematology 2000: Education Program of the American Society of Hematology 2000:294-8.
54 Ayton PM, Cleary ML. MLL in Normal and Malignant Hematopoiesis. In: Ravid K, Licht JD, editors. Transcription Factors: Normal and Malignant Development of Blood Cells. New York: Wiley-Liss, Inc.; 2001.
55 Huret JL. 11q23 rearrangements in leukaemia. In: Atlas Genet Cytogenet Oncol Haematol; 2001.
56 Harrison CJ, Cuneo A, Clark R, et al. Ten novel 11q23 chromosomal partner sites. Leukemia 1998;12:811-22.
57 Morrissey J, Tkachuk DC, Milatovich A, Francke U, Link M, Cleary ML. A serine/proline-rich protein is fused to HRX in t(4;11) acute leukemias. Blood 1993;81:1124-31.
58 Taki T, Hayashi Y, Taniwaki M, et al. Fusion of the *MLL* gene with two different genes, AF-6 and AF-5alpha, by a complex translocation involving chromosomes 5, 6, 8 and 11 in infant leukemia. Oncogene 1996;13(10):2121-30.
59 Taki T, Kano H, Taniwaki M, Sako M, Yanagisawa M, Hayashi Y. *AF5q31*, a newly identified *AF4*-related gene, is fused to *MLL* in infant acute lymphoblastic leukemia with ins(5;11)(q31;q13q23). Proc Natl Acad Sci U S A 1999;96:14535-40.
60 Hillion J, Le Coniat M, Jonveaux P, Berger R, Bernard OA. AF6q21, a novel partner of the MLL gene in t(6;11)(q21;q23), defines a Forkhead transcriptional factor subfamily. Blood 1997;9:3714-9.
61 Chaplin T, Bernard O, Beverloo HB, et al. The t(10;11) translocation in acute myeloid leukemia (M5) consistently fuses the leucine zipper motif of *AF10* onto the *HRX* gene. Blood 1995;86:2073-6.
62 Schichman SA, Caligiuri MA, Gu Y, et al. *ALL-1* partial duplication in acute leukemia. Proc Natl Acad Sci 1994;91:6236-9.
63 Prasad R, Leshkowitz D, Gu Y, et al. Leucine-zipper dimerization motif encoded by the *AF17* gene fused to *ALL-1 (MLL)* in acute leukemia. Proc Natl Acad Sci 1994;91:8107-11.
64 Nakamura T, Alder H, Gu Y, et al. Genes on chromosomes 4, 9, and 19 involved in 11q23 abnormalities in acute leukemia share sequence homology and/or common motifs. Proc Natl Acad Sci USA 1993;90:4631-5.
65 Borkhardt A, Repp R, Haas O, et al. Cloning and characterization of AFX, the gene that fuses to MLL in acute leukemias with a t(X;11)(q13;q23). Oncogene 1997;14:195-202.
66 Taki T, Sako M, Tsuchida M, Hayashi Y. The t(11;16)(q23;p13) translocation in myelodysplastic syndrome fuses the *MLL* gene to the *CBP* gene. Blood 1997;89:3945-50.
67 Thirman MJ, Levitan DA, Kobayashi H, Simon MC, Rowley JD. Cloning of *ELL*, a gene that fuses to *MLL* in a t(11;19)(q23;p13.1) in acute myeloid leukemia. Proc Natl Acad Sci 1994;91:12110-4.
68 Ida K, Kitabayashi I, Taki T, et al. Adenoviral E1A-associated protein p300 is involved in acute myeloid leukemia with t(11;22)(q23;q13). Blood 1997;90:4699-704.
69 Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H, Hayashi Y. LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). Cancer Res 2002;62(14):4075-80.

70 Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR. TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). Leukemia 2003;17(3):637-41.
71 Hayette S, Tigaud I, Vanier A, et al. AF15q14, a novel partner gene fused to the MLL gene in an acute myeloid leukaemia with a t(11;15)(q23;q14). Oncogene 2000;19(38):4446-50.
72 Bernard O, Mauchauffe M, Mecucci C, Van Den Berghe H, Berger R. A novel gene, *AF-1p*, fused to *HRX* in t(1;11)(p32;q23), is not related to *AF-4*, *AF-9* nor *ENL*. Oncogene 1994;9:1039-45.
73 Tse W, Zhu W, Chen HS, Cohen A. A novel gene, AF1q, fused to MLL in t(1;11) (q21;q23), is specifically expressed in leukemic and immature hematopoietic cells. Blood 1995;85(3):650-6.
74 Sano K, Hayakawa A, Piao J-H, Kosaka Y, Nakamura H. Novel SH3 protein encoded by the AF3p21 gene is fused to the mixed lineage leukemia protein in a therapy-related leukemia with t(3;11)(p21;q23). Blood 2000;95:1066-8.
75 Pegram LD, Megonigal MD, Lange BJ, et al. t(3;11) translocation in treatment-related acute myeloid leukemia fuses MLL with the GMPS (GUANOSINE 5' MONOPHOSPHATE SYNTHETASE) gene. Blood 2000;96(13):4360-2.
76 Daheron L, Veinstein A, Brizard F, et al. Human LLP gene is fused to MLL in a secondary acute leukemia with a t(3;11)(q28;q23). Genes Chromosomes & Cancer 2001;31:382-9.
77 Borkhardt A, Bojesen S, Haas OA, et al. The human GRAF gene is fused to MLL in a unique t(5;11)(q31;q23) and both alleles are disrupted in three cases of myelodysplastic syndrome/acute myeloid leukemia with a deletion 5q. Proc Natl Acad Sci USA 2000;97(16):9168-73.
78 Raffini LJ, Slater DJ, Rappaport EF, et al. Panhandle and reverse-panhandle PCR enable cloning of der(11) and der(other) genomic breakpoint junctions of MLL translocations and identify complex translocation of MLL, AF-4, and CDK6. Proc Natl Acad Sci USA 2002;99(7):4568-73.
79 Fuchs U, Rehkamp G, Haas OA, et al. The human formin-binding protein 17 (FBP17) interacts with sorting nexin, SNX2, and is an MLL-fusion partner in acute myelogeneous leukemia. Proc Natl Acad Sci U S A 2001;98(15):8756-61.
80 Taki T, Shibuya N, Taniwaki M, et al. *ABI-1*, a human homolog to mouse Abl-1 interactor 1, fuses the *MLL* gene in acute myeloid leukemia with t(10;11)(p11.2;q23). Blood 1998;92:1125-30.
81 Fu J, Hsu J, Tang T, Shih L. Identification of CBL,a Proto-oncogene at 11q23.3, as a Novel MLL Fusion Partner in a Patient With de Novo Acute Myeloid Leukemia. Genes, Chromosomes & Cancer 2003;37:214-19.
82 Chinwalla V, Chien A, Odero M, Neilly MB, Zeleznik-Le NJ, Rowley JD. A t(11;15) fuses MLL to two different genes, AF15q14 and a novel gene MPFYVE on chromosome 15. Oncogene 2003;22(9):1400-10.
83 Megonigal MD, Cheung NK, Rappaport EF, et al. Detection of leukemia-associated MLL-GAS7 translocation early during chemotherapy with DNA topoisomerase II inhibitors. Proc Natl Acad Sci U S A 2000;97(6):2814-9.
84 Strehl S, Borkhardt A, Slany R, Fuchs UE, Konig M, Haas OA. The human LASP1 gene is fused to MLL in an acute myeloid leukemia with t(11;17)(q23;q21). Oncogene 2003;22(1):157-60.

85 So C, Caldas C, Liu M-M, et al. *EEN* encodes for a member of a new family of proteins containing a Src homology 3 domain and is the third gene located on chromosome 19p13 that fuses to *MLL* in human leukemia. Proc Natl Acad Sci USA 1997;99:2563-8.
86 Megonigal MD, Rappaport EF, Jones DH, et al. t(11;22)(q23;q11.2) In acute myeloid leukemia of infant twins fuses MLL with hCDCrel, a cell division cycle gene in the genomic region of deletion in DiGeorge and velocardiofacial syndromes. Proc Natl Acad Sci U S A 1998;95(11):6413-8.
87 Osaka M, Rowley JD, Zeleznik-Le NJ. MSF (MLL septin-like fusion), a fusion partner gene of MLL, in a therapy-related acute myeloid leukemia with a t(11;17)(q23;q25). Proc Natl Acad Sci USA 1999;96:6428-33.
88 Taki T, Ohnishi H, Shinohara K, et al. AF17q25, a putative septin family gene, fuses the MLL gene in acute myeloid leukemia with t(11;17)(q23;q25). Cancer Res 1999;59(17):4261-5.
89 Borkhardt A, Teigler-Schlegel A, Fuchs U, et al. An ins(X;11)(q24;q23) fuses the MLL and the Septin 6/KIAA0128 gene an infant with AML-M2. Genes Chromosomes & Cancer 2001;32:82-8.
90 Ono R, Taki T, Taketani T, et al. *SEPTIN6*, a human homologue to mouse *Septin6*, is fused to *MLL* in infant acute myeloid leukemia with complex chromosomal abnormalities involving 11q23 and Xq24. Cancer Res 2002;62:333-7.
91 Slater DJ, Hilgenfeld E, Rappaport EF, et al. MLL-SEPTIN6 fusion recurs in novel translocation of chromosomes 3, X, and 11 in infant acute myelomonocytic leukaemia and in t(X;11) in infant acute myeloid leukaemia, and MLL genomic breakpoint in complex MLL-SEPTIN6 rearrangement is a DNA topoisomerase II cleavage site. Oncogene 2002;21(30):4706-14.
92 Eguchi M, Eguchi-Ishimae M, Seto M, et al. GPHN, a novel partner gene fused to MLL in a leukemia with t(11;14)(q23;q24). Genes Chromosomes Cancer 2001;32(3):212-21.
93 Wechsler DS, Engstrom LD, Alexander BM, Motto DG, Roulston D. A novel chromosomal inversion at 11q23 in infant acute myeloid leukemia fuses MLL to CALM, a gene that encodes a clathrin assembly protein. Genes, Chromosomes & Cancer 2003;36(1):26-36.
94 Kourlas PJ, Strout MP, Becknell B, et al. Identification of a gene at 11q23 encoding a guanine nucleotide exchange factor: evidence for its fusion with MLL in acute myeloid leukemia. Proc Natl Acad Sci USA 2000;97(5):2145-50.
95 Prasad R, Gu Y, Alder H, et al. Cloning of the *ALL-1* fusion partner, the *AF-6* gene, involved in acute myeloid leukemias with the t(6;11) chromosome translocation. Cancer Res 1993;53:5624-8.
96 LoNigro L, Slater DJ, Rappaport EF, et al. Two new partner genes of MLL and additional heterogeneity in t(11;19)(q23;p13) translocations. Blood 2002;100 (Suppl 1):531a.
97 Caligiuri MA, Strout MP, Schichman SA, et al. Partial tandem duplication of *ALL1* as a recurrent molecular defect in acute myeloid leukemia with trisomy 11. Cancer Res 1996;56:1418-25.
98 Megonigal MD, Rappaport EF, Jones DH, et al. Panhandle PCR strategy to amplify MLL genomic breakpoints in treatment-related leukemias. Proc Natl Acad Sci USA 1997;94(21):11583-8.

99 Nilson I, Reichel M, Ennas MG, et al. Exon/intron structure of the *AF-4* gene, a member of the *AF-4/LAF-4/FMR-2* gene family coding for a nuclear protein with structural alterations in acute leukaemia. Br J Haematol 1997;98:157-69.
100 Tatsumi K, Taki T, Taniwaki M, et al. The *CDCREL1* gene is fused to *MLL* in de novo acute myeloid leukemia with t(11;22)(q23;q11.2) and its frequent expression in myeloid leukemia cell lines. Genes Chromosomes & Cancer 2001;30:230-5.
101 So CW, Cleary ML. MLL-AFX requires the transcriptional effector domains of AFX to transform myeloid progenitors and transdominantly interfere with forkhead protein function. Mol Cell Biol 2002;22(18):6542-52.
102 So CW, Cleary ML. Common mechanism for oncogenic activation of MLL by forkhead family proteins. Blood 2003;101(2):633-9.
103 Secker-Walker LM. General Report on the European Union Concerted Action Workshop on 11q23, London, UK, May 1997. Leukemia 1998;12(5):776-8.
104 Erfurth F, Hemenway CS, de Erkenez AC, Domer PH. MLL fusion partners AF4 and AF9 interact at subnuclear foci. Leukemia 2004;18:92-102.
105 Corral J, Lavenir I, Impey H, et al. An *Mll-AF9* fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes. Cell 1996;85:853-61.
106 Lavau C, Szilvassy SJ, Slany R, Cleary ML. Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. Embo J 1997;16(14):4226-37.
107 Lavau C, Luo RT, Du C, Thirman MJ. Retrovirus-mediated gene transfer of MLL-ELL transforms primary myeloid progenitors and causes acute myeloid leukemias in mice [In Process Citation]. Proc Natl Acad Sci U S A 2000;97(20):10984-9.
108 Lavau C, Du C, Thirman M, Zeleznik-Le N. Chromatin-related properties of CBP fused to MLL generate a myelodysplastic-like syndrome that evolves into myeloid leukemia. Embo J 2000;19(17):4655-64.
109 So CW, Karsunky H, Passegue E, Cozzio A, Weissman IL, Cleary ML. MLL-GAS7 transforms multipotent hematopoietic progenitors and induces mixed lineage leukemias in mice. Cancer Cell 2003;3(2):161-71.
110 Liedman D, Zeleznik-Le N. Retroviral transduction model of mixed lineage leukemia fused to CREB binding protein. Curr Opin Hematol 2001;8(4):218-23.
111 Zeisig BB, Schreiner S, Garcia-Cuellar MP, Slany RK. Transcriptional activation is a key function encoded by MLL fusion partners. Leukemia 2003;17(2):359-65.
112 So CW, Lin M, Ayton PM, Chen EH, Cleary ML. Dimerization contributes to oncogenic activation of MLL chimeras in acute leukemias. Cancer Cell 2003;4(2):99-110.
113 Ayton PM, Cleary ML. Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. Genes Dev 2003;17(18):2298-307.
114 Kumar AR, Hudson WA, Chen W, Nishiuchi R, Yao Q, Kersey JH. Hoxa9 influences the phenotype but not the incidence of Mll-AF9 fusion gene leukemia. Blood 2004;103(5):1823-8.

115 So CW, Karsunky H, Wong P, Weissman IL, Cleary ML. Leukemic transformation of hematopoietic progenitors by MLL-GAS7 in the absence of Hoxa7 or Hoxa9. Blood 2004;103(8):3192-9.
116 Gale K, Ford A, Repp R, et al. Backtracking leukemia to birth: Identification of clonotypic gene fusion sequences in neonatal bloodspots. Proc Natl Acad Sci USA 1997;94:13950-4.
117 Ford AM, Ridge SA, Cabrera ME, et al. *In utero* rearrangements in the trithorax-related oncogene in infant leukaemias. Nature 1993;363:358-60.

CHAPTER 2

PANHANDLE PCR APPROACHES AND THE *ARMC3* GENE

The human *MLL* gene is one of the few genes involved with multiple translocation partners in leukemia. So far, approximately 88 different recombination products have been described able to disrupt *MLL* and cause aggressive leukemias[1,2]. Thus, rapid and accurate methods to identify and characterize genomic breakpoint junctions and fusion transcripts are essential for risk group stratification and treatment protocol assignments.

In this chapter the attention will be focused on the panhandle PCR method that uses a stem-loop template and two sides PCR to identify *MLL* breakpoint regardless of the partner gene. In the second part, *MLL* genomic breakpoint junctions will be characterized using the panhandle PCR in a cohort of children with therapy-related leukemia cytogenetically 11q23 rearranged. Cohort patients made up of therapy-related leukemias was chosen because it's known that chemotherapy agents promote double-strands breaks (DSBs) by targeting the essential cellular enzyme topoisomerase II (topoII), and are associated with characterized balanced chromosomal translocations. There is a clear association between exposure to etoposide and chromosomal rearrangements involving *MLL*. Finally, this study identified a new breakpoint junction of *ARMC3* gene from band 10p12 and *MLL* intron 9.

2.1 Panhandle PCR approaches

Panhandle polymerase chain reaction (PCR) method has been developed in order to identify and characterize genomic breakpoint junction sequences and fusion transcripts of *MLL* translocations, when only the end of *MLL* sequence is known[3-5]. Panhandle PCR methodology takes advantage of that nearly all chromosomal translocations occur within the specific region of *MLL* located between exons 5-11[6]. This region is known as breakpoint cluster region (bcr) and is specifically bounded by two *Bam*HI and three *Bg*lII restriction sites,

respectively[4-6] (fig.1.2.1). During *MLL* rearrangement, the gene is disrupted and an 8.3-kb *Bam*HI fragment or 6- and 2.9-kb *Bgl*II DNA fragments are able to fuse with one genomic DNA fragment that by chance exhibits *Bam*HI or *Bgl*II restriction sites. Thus, the recombination process generates a chimeric DNA that contains a part of known sequence from *MLL* juxtaposed, and an unknown sequence from one partner gene. These new chimeric DNAs are the primary targets for the panhandle PCR approach that involves attaching sequence that is complementary to *MLL* sequence to a *Bam*HI or *Bgl*II restriction site in the partner gene, which leads to the annealing of *MLL* and its complement in the handle and formation of a single-stranded intrastrand loop template that is made double stranded and amplified in two sequential PCRs with primers all from *MLL*[4,7]. The principle of panhandle PCR was invented by Jones and co-workers in 1992. The panhandle PCR was developed to amplify directly human genomic DNA target sequences of 2-4.4-kb from *β*-globin and cystic fibrosis transmenbrane conductance regulator genes utilizing primers that flanked only one side of the region amplified. This method included the generation of the genomic DNA template by restriction enzyme digestion followed by ligation to a single-stranded oligonucleotide. The ligated-oligonucleotide has free 3' end that is complementary to the known region of DNA. Thus, the single-strands of DNA, which contain the complement of the ligated-oligonucleotide, undergo self-annealing forming a stem-loop structure. The ligated oligonucleotide can then prime template-directed DNA polymerization so that the PCR amplification of unknown DNA can be carried out because sequence now flanks both ends of the unknown DNA[8].

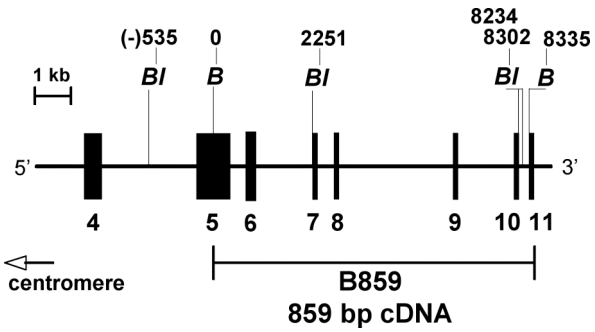


Figure 1.2.1. Partial restriction map of the *MLL* gene breakpoint cluster region (bcr). The B859 probe[4,5] is a cDNA fragment containing *MLL* exons 5-11 sequences. Narrow bars represent the introns and black boxes represent the exons *B* indicates *BamHI*; *Bl*, *Bgl*II.

Different panhandle PCR approaches have been developed to identify and characterize *MLL* breakpoint junctions at both genomic and transcriptome levels. The genomic included the originals *Bam*HI and *Bgl*II based panhandle PCRs for cloning the derivative chromosome 11 [der(11)] breakpoint junctions and the reverse *Bam*HI and *Bgl*II based panhandle PCRs for breakpoint junctions on the other derivative chromosomes [der(other)] of *MLL* translocations[4,7,9]. While, cDNA panhandle PCR has been developed to identify der(11) fusion transcripts[10]. The general strategy of the genomic panhandle PCR includes restriction enzyme cleavage of the genomic leukemic samples to create a 5' overhang. This latter permits the ligation of a phosphorylated oligonucleotide that is complementary to known sequence of *MLL* bcr. Following is promoted the intrastrand annealing of the ligated oligonucleotide to the its complementary *MLL* bcr sequence (handle) that ultimately leads to the formation of a single-stranded intrastrand loop (pan) that contains the breakpoint junction sequence for amplification with primers all from *MLL*. Therefore, before performing panhandle PCR is convenient to analyze the *Bam*HI or *Bgl*II digested DNA samples by Southern blot analysis using a specific *MLL* bcr probe[4,5] (fig. 1.2.1). The Southern blot permits to identify if *MLL* is rearranged and leads to the choice of the most amenable panhandle PCR approach by the rearrangements sizes that approximate the target for panhandle PCR. The cDNA panhandle PCR approach is easier because it does not require restriction enzyme cleavage, nor ligation reaction. Moreover, in cDNA panhandle PCR the smaller amplicons contains the informative exonic sequences and reveal the partner genes more readily. The cDNA panhandle involves reverse transcription of first-strand cDNA from total RNA using *MLL*-random hexamer oligonucleotides. This produces first-strand cDNAs of different sizes. Those derived from the normal *MLL* allele and from der(11) transcripts have a known *MLL* sequence and its inverse complement at the 5' and 3' ends respectively. This permits the generation of second-strand cDNAs by *MLL* specific primer extension, formation of stem-loop templates by intrastrand annealing with the fusion point of the chimeric transcript in the loop. This generates the double-stranded template with *MLL* sequence at both ends that can be amplified with primers all derived from *MLL*.

As example, here is reported in detail the steps of the original *Bam*HI based panhandle PCR developed to amplify the genomic breakpoint from the der(11)

chromosome (fig. 2.2.1) [5]. The first step consists of genomic digestion with *Bam*HI to create a 5' overhang followed by treatment with calf intestinal alkaline phosphatase to prevent religation. The purpose of steps 2 and 3 is to form the handle. Formation of the handle attaches known *MLL* DNA to 3' of the unknown partner DNA and brings the translocation breakpoint and unknown partner DNA within an intrastrand loop. Step 2 involves ligation of a single-stranded 5' phosphorylated oligonucleotide (5'-GAT CGA AGC TGG AGT GGT GGC CTG TTT GGA TTC AGG-3') to the 3' ends of the digested DNA.

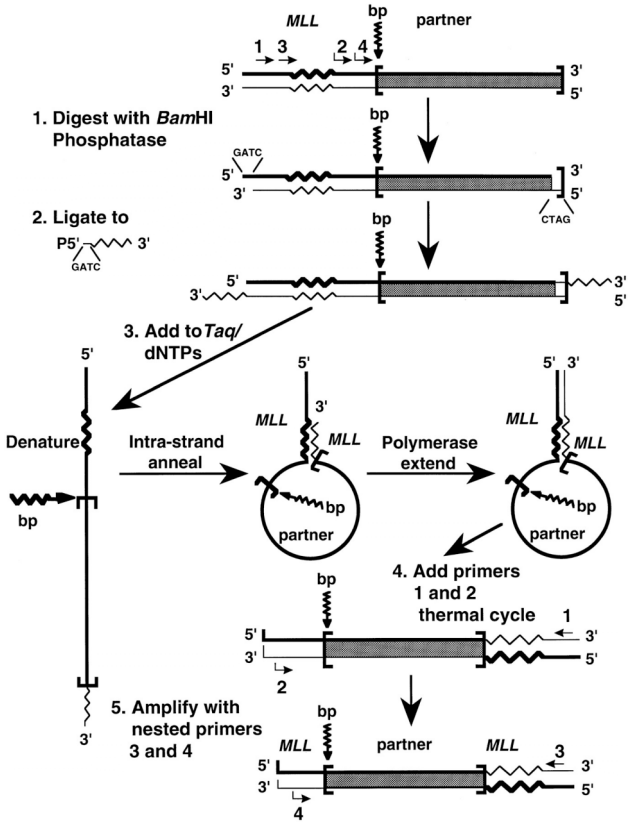


Figure 2.2.1. Schematic of panhandle PCR method to amplify *MLL* genomic breakpoint on der(11) chromosome [5]. Bold lines indicate sense DNA strand containing *MLL* and partner gene sequence from which the template is created, and sense orientations of primers. The 5' phosphorylated oligonucleotide contains a *Bam*HI overhang at the 5' end and the complement of sense sequence in *MLL* exon 5 at 3' end. The *Bam*HI overhang at the 5' end of the phosphorylated oligonucleotide promotes attachment of the complement of sequence in *MLL* exon 5 at the *Bam*HI site in the sense strand of the partner gene to create the template (step 2). Formation of the handle is completed in step 3 by intrastrand annealing of the ligated oligonucleotide to the complementary sequence in MLL and polymerase extension of the recessed 3' end. With MLL sequences at both ends of the template, primers 1 and 2 from exon 5 are used to amplify the breakpoint junction (step 4). In step 5, nested PCR with primers 3 and 4 enhances yield.

The 4-base 5' end of the oligonucleotide is complementary to the 5' overhang of

*Bam*HI digested DNA; its 3' end is complementary to nucleotides in *MLL* exon 5, which is in the 5' bcr. The sense strand becomes the template strand in Step 3. Formation of the handle is completed in Step 3 by intrastrand annealing of the ligated oligonucleotide to the complementary sequence in *MLL* and polymerase extension of the recessed 3' end. An aliquot of the *Bam*HI digested, ligated DNA is added to a reaction mixture of DNA polymerase, dNTPs, and PCR reaction buffer. After addition of the DNA, the reaction mixtures is heated at 94°C for 1 minute to make the template single-stranded. Intrastrand annealing of the ligated oligonucleotide to the complementary sequence in *MLL* and polymerase extension of the recessed 3' end completed formation of the handle during a 2 minute ramp to 72°C and incubation at 72°C for 30 seconds. Intrastrand annealing contains the translocation breakpoint and unknown partner DNA within an intrastand loop or pan-like structure. With *MLL* sequences at both ends of the template, the next step is to add *MLL* primers that are all sense with respect to exon 5 to amplify the breakpoint junction. The position and orientations of the primers with respect to the ligated oligonucleotide are shown in fig. 2.2.1. *MLL* primer 1 (5'-TCC TCC ACG AAA GCC CGT CGA G-3') is homologous to exon 5 upstream to the *MLL* sequence that is complementary to the ligated oligonucleotide. *MLL* primer 2 (5'-TCA AGC AGG TCT CCC AGC CAG CAC-3') is homologous to exon 5 between the MLL sequence that is complementary to the ligated oligonucleotite and the breakpoint junction. The step 5 consists of a nested PCR reaction with primer 3 and 4, also from *MLL* exon 5, enhances the yield of products from panhandle PCR. The sequences of nested primers are 5'-AGC TGG ATC CGG AAA AGA GTG AAG AAG GGA ATG TCT CGG-3' and 5'-AGC TGG ATC CGT GGT CAT CCC GCC TCA GCC AC-3'.

In summary, the prognostic significance of *MLL* rearrangements demands precise identification of its partner genes. Panhandle PCRs approaches is the unique method that allow the amplification of known sequence flanked by unknown sequence and it have been carried out with success in *MLL*-rearranged leukemias to identify *MLL* genomic breakpoint junctions and a part of the numerous *MLL* fusion partner genes. Overall, this method can contribute for new partner genes discovery and then it help to define the complete "*MLL* recombinome network" as well as to clarify the impact of the *MLL* translocations with sporadic partner genes on leukemia prognosis. Moreover, the achievement of

der(11) and der(other) genomic breakpoint junction sequences by panhandle PCR may provide important clues about the mechanism(s) that cause the translocations in terms of DNA sequence damages and of DNA processing repair.

Here, 8 cases of secondary leukemias that were previously screened by Southern blot for *MLL* translocations were revisited by Panhandle PCR. For all cases *MLL* fusion gene was identified. In seven cases *MLL* was fused with a known partner gene, whereas one patient carried a *MLL-AF10* and a novel fusion between *ARMC3* (*ARMADILLO REPEAT CONTAINING 3*) and *MLL*, in a 3-way complex translocation.

2.2 Characterization of *MLL* translocation breakpoint junctions in secondary leukemias by panhandle PCR approaches

Chemotherapy agents that target the cellular enzyme DNA topoisomerase II (topoII) are known to promote double-strand breaks (DSBs) and are associated with secondary leukemias (also termed therapy-related leukemias) characterized by balanced chromosomal translocations[11,12]. TopoII resolves knots and tangles of DNA by inducing transient DSBs[13,14]. Under normal conditions, the cleaved DNA ends are immediately relegated by topoII. However, in the presence of topoII-inhibiting agents (topoII poisons), like etoposide (representative of the chemotherapeutic epipodophyllotoxin), the topoII-DNA cleavage complex is stabilized and increases the risk of chromosomal translocations[15]. The induction of DSBs and *MLL* translocations, which are often reported in secondary leukemias, are assumed to occur through topoII-induced strand cleavage. Indeed, some recent studies have showed that etoposide induced DSBs which lead to *MLL* translocations in cultured human hematopoietic progenitor cells[16-18]. Moreover, evidence of a direct role of topoII poisons in mediating *MLL* gene translocations came from published case reports about therapy-related acute myeloid leukemia (t-AML) with *MLL* translocations in patients who received single agent therapy with etoposide [19].

The objective of this study is the understanding the nature of the damage leading to *MLL* translocations in secondary leukemias induced by the chemotherapeutic topoII poisons. The hypothesis is that etoposide treatment could stimulate a spectrum of different DNA lesions, which being repaired, are able to

form heterogeneous breakpoint junctions on *MLL*, that represent the heterogeneity of the DNA damage.

In this work, I characterized the genomic breakpoint junction sequences in 8 cases of secondary leukemias using various panhandle PCR approaches. Some of these cases (t-2, t-3, t-4, t-5, t-6, t-9 and t-36) were previously analyzed for the presence of *MLL* bcr rearrangements by Southern blot analysis of *Bam*HI digested DNA (Fig. 1.2.2)[20,21]. However, some of these patients (t-2, t-3 and t-4) presented large *MLL* bcr rearrangements (>8.3-kb *Bam*HI germline fragment) to be considered as target of panhandle PCR amplification. For this reason, these cases were further analyzed by hybridization of *Bgl*II digested DNA with the B859 cDNA probe (Fig. 1.2.2).

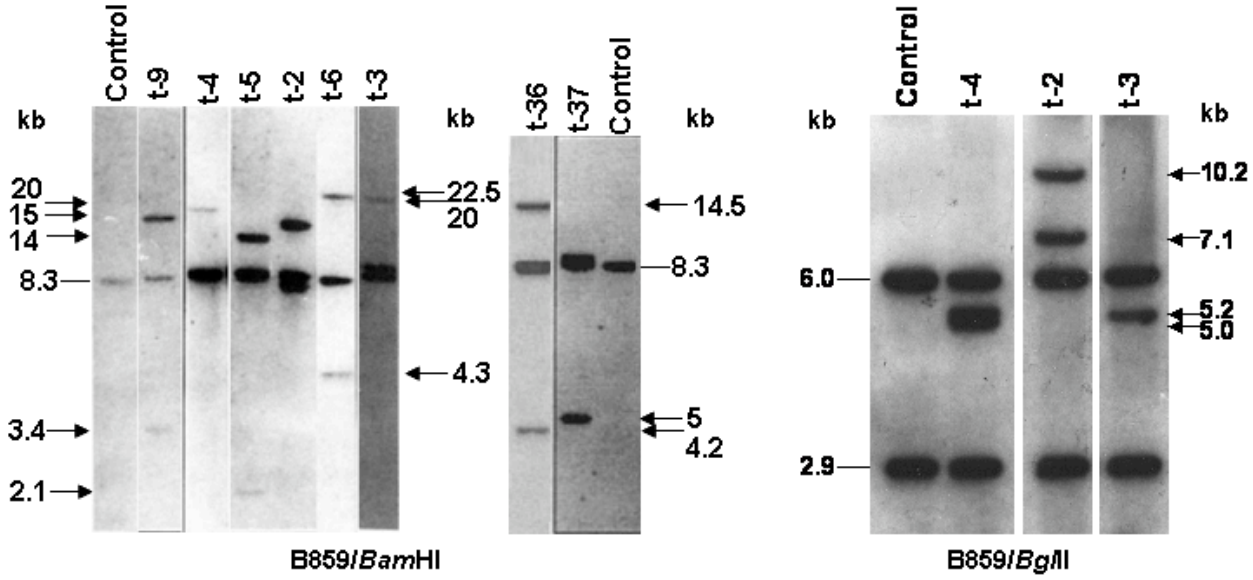


Figure 1.2.2. Characterization of *MLL* bcr rearrangements in secondary leukemias by Southern blot analysis[10,20,21]. Southern blot analysis of *Bam*HI-digested DNA of cases 2, 3, 4, 5, 6, 9, and 36 were reported previously; additional molecular analysis of the *MLL* bcr in cases 2, 3, 4 and 37 were performed herein. Southern blot analysis was done with 10 ug of genomic DNA extracted from leukemic cells and digested with *Bam*HI or *Bgl*II. The membrane was assayed for hybridization to B859 cDNA probe. Peripheral blood lymphocyte DNA from normal subjected was used as the control. Dash shows germ-line band; arrows show rearrangements.

In table 1.2.2 the status of *MLL* bcr rearrangements are summarized and the respective fragment sizes of eight cases of secondary leukemias obtained by Southern blot analyses. All eight leukemic samples from children previously treated for a primary cancer and diagnosed with secondary leukemia were

collected at the Children's Hospital of Philadelphia that has approved this research. The clinical and cytogenetics features of these patients are summarized in table 2.2.2 and 3.2.2. Each patient had received a DNA topoII inhibitor or developed secondary leukemia typical of that associated therapy. Primary cancer includes five solid tumors and three leukemias. Seven leukemias were monoblastic variants of AML, five was FAB M5, one was FAB M1, one was FAB M5, and one presented as MDS (table 1.2.2). The cytogenetic analyses reveal that seven cases where positive for the band 11q23 translocated with bands 19p13, or 9p22, or 3q25, or 10q22; while the karyotype was normal for one case (table 2.2.2).

Table 1.2.2. Southern blot analyses and *MLL* bcr status in secondary leukemias

Pt	***Bam*HI/B859 *MLL* gene rearrangements/size (kb)**	***Bgl*I/B859 *MLL* gene rearrangements/size (kb)**
2	R2, 15.5, 7.8	R2, 10.1, 7.5
3	R2, 20	R1, 5.2
4	R1, 20	R2 5.2, 5.0
5	R2, 14, 2.1	N.D.
6	R2, 22.5 4.3	N.D.
9	R2, 15, 3.4	N.D.
36	R2, 14.5 4.2	N.D.
37	R2, 10.6. 5.0	N.D.

Abbreviations: Pt, patient; R1, one rearrangement; R2, two rearrangements; N.D., not done.

Table 2.2.2. Clinical features of patients with secondary leukemia

			PRIMARY CANCER																				
			Therapy																				
Pt	**Dx**	**Age (yrs)**	A-ase	ADR	AMD	Ara-C	BCNU	BMT*	CBDCA	CPPD	Cort**	CPM	DNM	IFOS	MTX	PCZ	6MP	6TG	VCR	VLB	VM26	VP16	XRT
2	ALL	2.7	●			●					●		●		●		●		●			●	
3	ALL	1.5	●			●					●	●	●		●		●		●			●	●
4	OS	7.0		●						●					●				●			●	
5	HD	15.7					●				●	●				●				●			●
6	NBL	3.5		●				o		●		●							●		●		●
9	PNET	13.8		●								●		●					●			●	
36	ARMS	1.4												●					●			●	
37	ALL	1.8	●	●		●					●	●	●		●		●	●	●				

Table 2.2.2. Clinical features of patients with secondary leukemia (continued)

SECONDARY LEUKEMIA																			
			Therapy																
Pt	dx/ FAB	Mos From Dx of primary cancer	A-ase	ADR	Ara-C	BMT*	Cort**	CPPD	DNM	EPO	5-AZA	HU	MIT	MTX	6MP	6TG	VCR	VP16	Status (Mos)
2	AML/M4	50.0		●	●						●					●		●	NED 19
3	AML/M4	42.0					●					●		●	●		●		DOD 1.5
4	AML/M1	60.0	●		●														DOD 0.5
5	AML/M4	16.0			●		●		●								●		DOC 4
6	AML/M4	18.0				●													DOD 2
9	AML/M5	11.0																	DOD 0.2
36	AML/M4	60.0		●													●	●	DOD 6
37	MDS/RAEB-t	62.0	●				●							●		●		●	DOD 7

Abbreviations: Pt, patient; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; MDS, myelodysplastic syndrome; Dx, diagnosis; OS, osterorsarcoma; HD, Hodgkin's disease; NBL, neuroblastoma; ARMS, alveolar rhabdomyosarcoma; A-ase, L-asparginase; ADR, Adriamycin (doxorubicin), AMD, dactinomycin; Ara-C, cytosine arabinoside; BCNU, carmustine; BMT*, bone marrow transplant; CBDCA, carboplatin; Cort**, cortcosteroid; CPPD, cispaltin; CMP, cyclophosphamide; DNM, daunorubicin; EPO, erythropoietin; 5-aza, 5-azacytine; HU, hydroxyurea; IFOS, ifosfamide; MIT, mitoxantrone; MTX, methotrexate; PCZ, procarbazine; 6MP, 6-mercaptopurine; 6TG, 6-thioguanine; VCR, vincristine; VLB, vinblastine; VM26 teniposide; VP16 etoposide; XRT, radation therapy; FAB, French-American-Brithish; DOD, dead of disease; DOC, dead of complications; NED, no evidence of disease.

Table 3.2.2. Cytogenetics characteristic of patients with secondary leukemias

Pt	dx/FAB	Karyotype
2	AML/M4	46,XX,t(11;19)(q23;p13.3) [10 cells]/46,XX[10cells]
3	AML/M4	47,XX,+8,t(11;19)(q23;p13)[19 cells]/48,XX,+8,+8,t(11;19)(q23;p13)[6cells]
4	AML/M1	46,XX,t(9;11)(p22;q23)[27cells]
5	AML/M4	46,XY,t(11q-;19p+)[25cells]
6	AML/M4	46,XY,t(1;7)(q32;q32-34),inv(2)(p21;q37),t(3;11)(q25;q23),t(7;?)(q22;?)[29cells]
9	AML/M5	46,XY,t(9;11)(p22;q23)[20cells]
36	AML/M4	46,XY
37	MDS/RAEB-t	46,XY,t(10;11)(q22;q23),del(10)(p11.2)[20cells]

Abbreviations: Pt, patient; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; MDS, myelodysplastic syndrome; FAB, French-American-British.

High molecular weight DNA was isolated from cryopreserved leukemic marrow cells using 4M GITC-5.7 M CsCl gradients as described[4]. Panhandle PCR amplification of *MLL* genomic breakpoints was done as previously reported[4,7,9]. The nested panhandle PCR products were purified using Geneclean III kit (Qbiogene) and subcloned by recombination PCR[10,22]. Alternatively the nested panhandle PCR products were subcloned into pCR 2.1 TOPO vector using a TOPO TA cloning Kit (Invitrogen). Recombinant plasmids containing inserts of the desired size were identified by a PCR screen with the primers used in the nested panhandle PCR, and the corresponding subclones were prepared for sequencing[10,22]. The der(11) and der(other) breakpoint junctions were confirmed

by amplification of genomic DNAs with *MLL*- and partner gene-specific primers and direct sequencing.

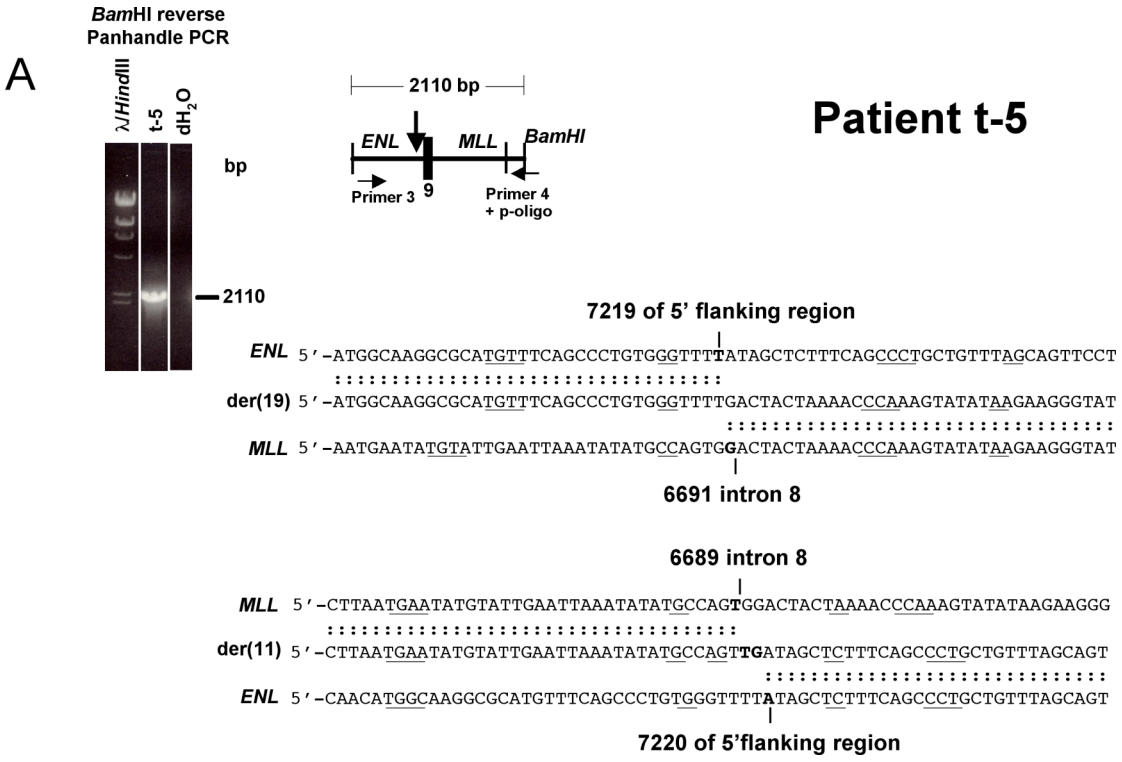
In cases t-5, t-6, t-9 and t-36 the genomic breakpoint junction sequences were obtained by *Bam*HI reverse panhandle PCR because the *Bam*HI fragment containing the der(11) breakpoint junctions were too large to be amplified (fig. 1.2.2). Thus, the der(11) breakpoint junctions were obtained by PCR with gene specific primers.

In case t-5 a 2,110 bp reverse panhandle PCR product indicated that 2.1 kb *MLL* bcr rearrangement found by Southern blot was from the der(19) chromosome (figs. 1.2.2 and 2.2.2A). The *ENL* or *MLLT1* der(19) breakpoint was positioned at 7219nt upstream of exon 1 (GenBank accession no. NM_005934). The *MLL* der(19) breakpoint was positioned at 6991nt in intron 8 (GenBank accession no. U04737) (Figure 2b). PCR with *MLL*- and *ENL*-specific primers amplified a 476-bp product containing the der(11) genomic breakpoint junction predicted by the der(19) sequence. The *MLL* breakpoint on the der(11) chromosome was at 3' bcr position 6689 nt in intron 8; the *ENL* breakpoint on the der(11) chromosome was positioned at 7220nt upstream of exon 1 (fig. 2.2.2A). The 5'-TG-3' sequence was included in the breakpoint of der(11).

In case t-6, the karyotype was t(3;11)(q25;q23), the 4.3-kb *Bam*HI rearrangements (fig. 1.2.2) contained the der(3) breakpoint junction sequence. The *GMPS* (*GUANOSINE 5'-MONOPHOSPAHTE SYNTHETASE*) der(3) breakpoint corresponds to nucleotide 406, 407 or 408 upstream of exon 1 (GenBank accession no. NM_003875); the *MLL* der(3) breakpoint was positioned at 5244, 5244 or 5246nt in intron 8 (fig. 2.2.2B). The presence of homologous 5'-AG-3' sequences in *GMPS* and *MLL* genes, precluded a more precise breakpoint assignment. The amplification of a 457 bp product with *MLL* intron 8- and *GMPS*-specific primers and the sequence, revealed that the der(11) genomic breakpoint sequence was positioned at 5238-5242nt in intron 8 of *MLL* and at 410-414nt upstream of the exon 1 of *GMSP*. A 5'-CTCT-3' sequence in *MLL* and *GMSP* precluded more precise breakpoint assignment (fig. 2.2.2B).

Southern blot analysis of patient t-9 revealed 15- and 3.4-kb *MLL* bcr rearrangements (fig. 1.2.2). The reverse *Bam*HI panhandle PCR identified an *AF9*-intron 5 and *MLL*-intron 8 genomic fusion breakpoint of the der(9) chromosome (fig. 2.2.2C). The *AF9* breakpoint was at 28337nt in intron 5 (GenBank accession

no. NM_004529); the *MLL* breakpoint was at 6588nt in intron 8. An insertion of 10 bp sequence (5'-ATTTCTATTT-3') was present between the two genes. *MLL* and *AF9* primers were designed to amplify the der(11) genomic breakpoint junction predicted by der(9) sequence. The expected product size was 432 bp that revealed *MLL* der(11) breakpoint at position 6569nt in intron 8, and the der(11) breakpoint in the partner gene was positioned at 28357nt in *AF9* intron 5. A "T" residue was inserted between the *MLL* and *AF9* breakpoint junction.



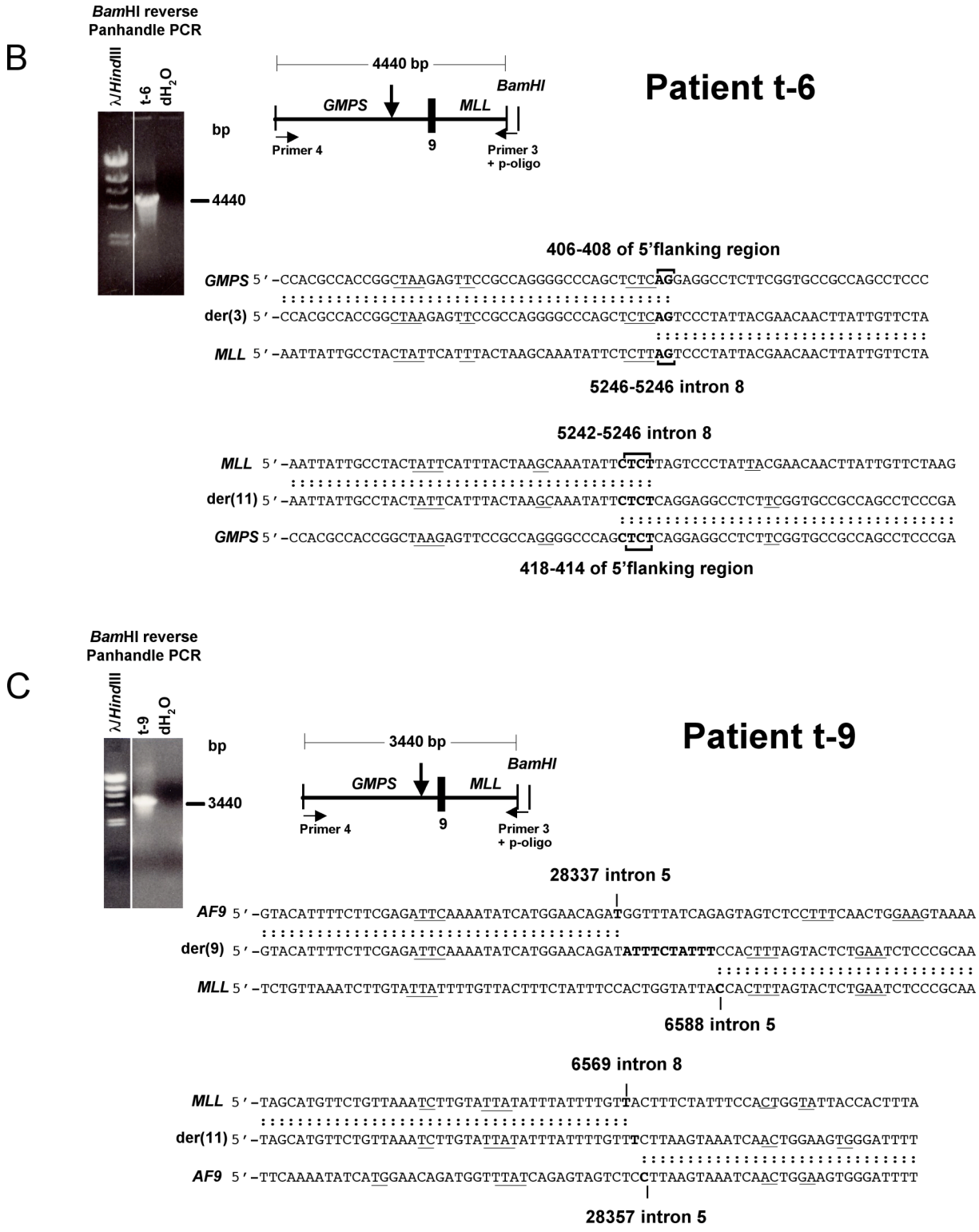


Figure 2.2.2. Reverse *Bam*HI panhandle PCR products from leukemias of patients t-5 (A, top left), t-6 (B, top left) and t-9 (C top left) and corresponding der(other) genomic breakpoint junction sequences from subclones generated by TOPO TA cloning of products of *Bam*HI panhandle PCR. Sequences of the der(11) genomic breakpoint junctions in all three cases were characterized by PCR with gene-specific primers (A, B, C, bottom).

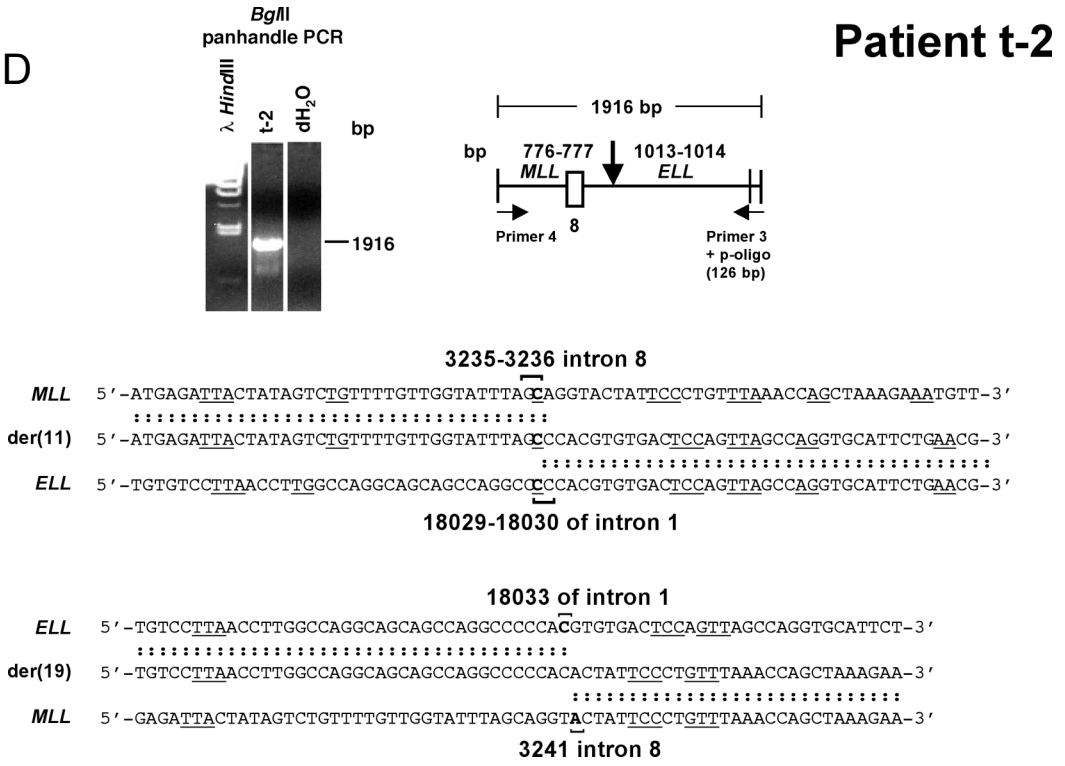
For the t-AML of patient t-36, who will be extensively discussed in the next section, *Bam*HI reverse panhandle PCR identified a new *ARMC3-MLL* fusion in a cryptic *MLL-AF10* three-way complex translocation where the karyotype was normal. *ARMC3* (*ARMADILLO REPEAT CONTAINING 3*), which encodes Arm repeats similar to catenin family proteins, is the first gene to be involved in an *MLL* translocation.

Original *Bam*HI panhandle PCR was utilized to characterize the genomic breakpoint sequences of patient t-37 diagnosed as secondary MDS with a t(10;11) translocation. Panhandle PCR amplified the der(11) genomic breakpoint. The der(11) breakpoint in *MLL* was positioned at 993-995nt of the bcr, in intron 6. The partner DNA at the breakpoint junction was *CXXC6* (*CXXC FINGER PROTEIN 6*). The *CXXC6* der(11) breakpoint was at 878-876nt in intron 10 (GenBank accession no. NM_030625). The presence of two "GG" residues at the breakpoint junction in both genes precluded precise breakpoint assignment (data not shown).

Finally, for the cases (t-2, t-3 and t-4) that showed a large *Bam*HI *MLL* bcr rearrangements (fig. 1.2.2), the der(11) breakpoint junctions have been obtained by original *Bgl*II panhandle PCR approach, after Southern blot analysis of *Bgl*II digested DNA, using the B859 cDNA probe (fig. 1.2.2).

Although, case t-2 showed a large *Bgl*II *MLL* bcr rearrangements (fig. 1.2.2), the original *Bg*lII panhandle PCR permitted to clone the der(11) genomic breakpoint sequence with a 1923bp product rather than a 10.7- or 7.5-kb as predicted on Southern blot (figs. 1.2.2 and 2.2.2D). The presence of a cryptic polymorphism on *MLL* genomic sequence, which is recognized by the phosphorylated oligonucleotide (p-oligo) able to promote the intrastrand annealing and panhandle formation, could explain the amplification of an unexpected shorter PCR product. The MLL der(11) breakpoint was at 5' in the bcr 3235nt or 3236nt in intron 8. The der(11) breakpoint in the partner gene was positioned at 18029nt or 18030nt in *ELL* intron 1 (GenBank accession no. NM_006532). A "C" residue at the breakpoints of both genes, precluded more precise assignment. *ELL* and *MLL* primers were designed to amplify the der(19) genomic breakpoint junction predicted by the der(11) sequence. The der(19) breakpoint in *ELL* was positioned at 18033nt in intron 1 and der(19) breakpoint in *MLL* was positioned at 3241nt in intron 8 (fig. 2.22D).

In case t-3 the Southern blot analysis revealed a 5.1-kb *Bgl*II rearrangement containing the der(11) brekpoint junction as proved by *Bgl*II panhandle PCR (fig. 1.2.2). In fact, the 4851 bp PCR product revealed the *MLL* der(11) breakpoint at position 6424nt, 6425nt or 6426nt in intron 8. The der(11) breakpoint in the partner gene was positioned at 3422nt, 3423nt or 3424nt in *ENL* intron 6 (fig. 2.2.2E). The der(11) breakpoint in *MLL* and *ENL* involved an overlapping of the 5'-AC-3' sequence that precluded precise assignment. PCR with *ENL*- and *MLL*-specific primers amplified a 419 bp product containing the der(19) genomic breakpoint predicted by der(11) sequence. The *MLL* on the der(19) chromosome was positioned at 6421nt, 6422nt or 6423nt in intron 8; the *ENL* breakpoint on der(19) chromosome was at 3419nt, 3420nt or 3421nt in intron 6 (fig. 2.2.2E). In case t-4, der(11) breakpoint fused *MLL* intron 8 (position 6133nt) to *AF9* intron 5 (position 15639nt) (fig. 2.2.2F). The der(11) breakpoint junction informed gene specific primers to amplify the der(9) breakpoint junction. The der(9) breakpoint in *AF9* was positioned at 15690-15692nt in intron 5 and MLL der(9) breakpoint was at 6138-6140nt in intron 8 (fig. 2.2.2F).



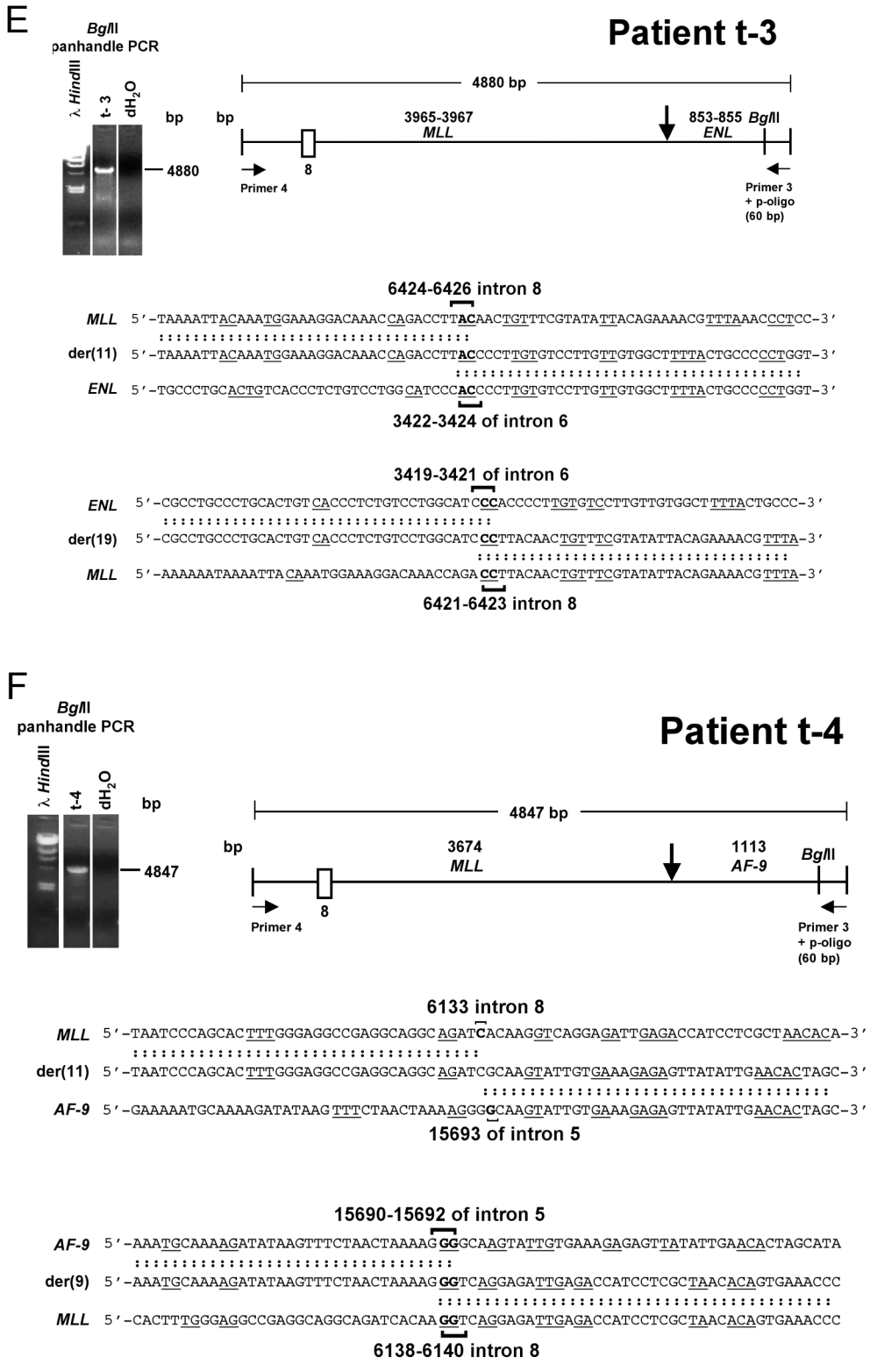


Figure 2.2.2. *Bgl*II panhandle PCR products from leukemias of patients t-2 (D, top left), t-3 (E, top left) and t-4 (F top left) and corresponding der(11) genomic breakpoint junction sequences from subclones generated by TOPO TA cloning of products of *Bgl*II panhandle PCR. Sequences of the der(other) genomic breakpoint junctions in all three cases were characterized by PCR with gene-specific primers (D, E, F, bottom).

The *MLL* genomic breakpoint junctions on the der(11) and der(other) chromosomes of *MLL* translocations in leukemias following chemotherapeutic topoisomerase II poisons are summarized in fig. 3.2.2. The sequence findings in these cases reinforce the theme that *MLL* translocations in treatment related leukemias are precise or near-precise recombinations with gains or losses of either no or, more often, a few bases in the formation of both breakpoint junctions, suggesting a minimal processing. Furthermore, these precise or near-precise recombinations with few bases gained of lost relative to the normal sequences of *MLL* and its partner genes indicate that the translocation breakpoints are in close proximity in the regions where damage occurred. In addition, the finding of short sequence homologies in *MLL* and its partner genes at the translocation breakpoint junctions (fig. 2.2.2) suggest DNA damage resolution by the repair mechanism of nonhomologous end-joining (NHEJ), which typically introduces small deletions and insertions to create homologous overhangs or, less often, accurately joins the ends[4,10,22-26]. Results revealed additional heterogeneity in *MLL* translocation breakpoint distribution within the bcr, and confirmed a translocation breakpoint hotspot at 3' in the bcr intron 8 (fig. 3.2.2).

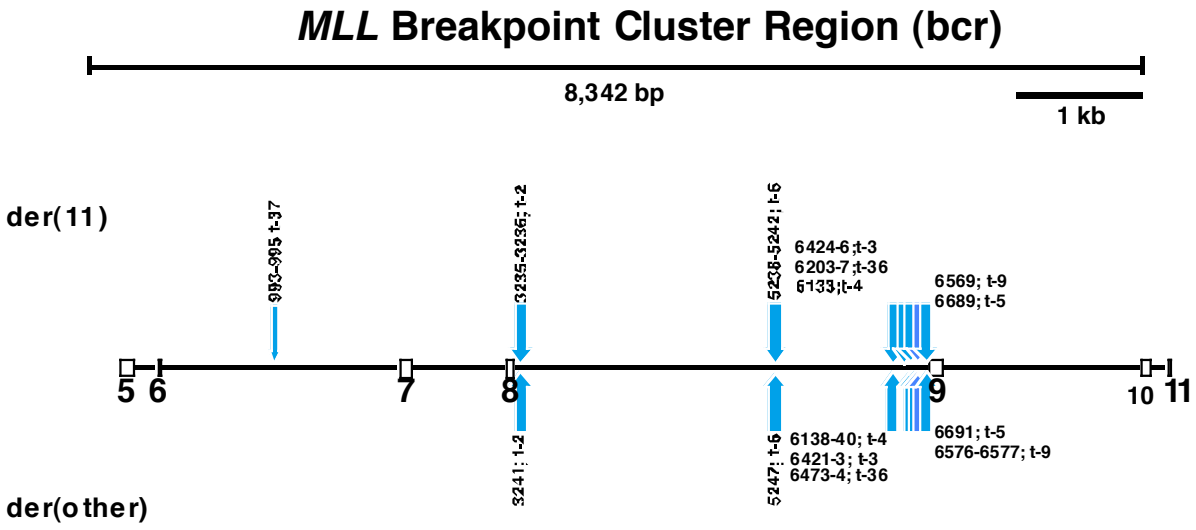


Figure 3.2.2. Heterogeneity in *MLL* translocation breakpoint distribution with hotspot 3' in *MLL* intron 8, and pattern of near precise and precise interchromosomal DNA recombinations in leukemias related to topoisomerase II poisons. *MLL* translocation breakpoint sites where only the der(11) (fine arrow) or both genomic breakpoint junctions (bold arrows, top and bottom of bcr) have been cloned.

Cloning of additional genomic breakpoint junctions on both derivative chromosomes is required to better understand the spectrum of the damage and its resolution that results in translocations and answer whether, how often and to what degree precise recombinations, deletions, insertions, duplications and NHEJ have occurred in creation of the breakpoint junctions.

Secondary leukemias with *MLL* translocations have a *de novo* counterpart in acute leukemia in infants. *MLL* translocation breakpoints in infant are distributed heterogeneously in the bcr with a 3' bias[9,10,24,27,28]. However, the breakpoint junction sequence in infant leukemias often show regions up to several hundred bases from *MLL* and/or from its partner gene on both derivative chromosome suggesting duplications or, less often, deletions of several hundred bases[4, 9, 24, 29].

In conclusion, this study confirms the workability of panhandle PCR method for cloning both genomic breakpoint junctions, which is the prerequisite to localize the genomic regions where DNA damage has occurred and has been resolved to translocations. In addition, these findings support the model for which topoisomerase II inhibitors contribute to *MLL* chromosomal translocations through an NHEJ mechanism. Finally, a novel *ARMC3-MLL* rearrangement was discovered in a t-AML case.

2.3 Cryptic, complex *MLL*, *AF10*, *ARMC3* rearrangement in secondary AML disrupts mammalian armadillo homologue

Giuseppe Germano[1,2], Blaine W. Robinson[1], Carolyn A. Felix[1,3]

(in process)

[1]Division of Oncology, the Children's hospital of Philadelphia; [2]laboratory of haematology and oncology, University of Padua, Padua, it; [3]Department of Pediatrics, University of Pennsylvania school of medicine, Philadelphia, pa 19104

Abstract

5'-*MLL-AF10*-3' translocations form by complex rearrangements because the telomere to centromere orientation of *AF10* (*MLLT10*) at band 10p12 is opposite to the orientation of *MLL* at band 11q23. Previously we discovered a cryptic *MLL* translocation and 5'-*MLL-AF10*-3' fusion transcripts in a case of secondary AML with a normal karyotype. Here we demonstrate that this translocation was a 3-way

rearrangement disrupting *ARMC3 (ARMADILLO REPEAT CONTAINING 3)*. *Bam*HI reverse panhandle PCR identified the unknown genomic 5'-*partner gene-MLL*-3' breakpoint junction fusing intron 17 of *ARMC3* from band 10p12 and *MLL* intron 8 5' to the secondary leukemia *MLL* translocation breakpoint hotspot. The out-of-frame transcript joining *ARMC3* exon 17 and *MLL* exon 9 contained a premature termination codon and would alter and truncate the carboxyl terminus of ARMC3. PCR with gene-specific primers identified the 5'-*MLL-AF10*-3' genomic breakpoint junction at identical 5'-ATTAG-3' sequences in *MLL* intron 8 and *AF10* intron 9. Since *ARMC3* is in the same orientation and adjacent and centromeric to *AF10*, the complex 3-way rearrangement occurred by splitting of band 11q23 and insertion of band 10p12 material containing the 3' portion of *AF10* through the 5' portion of *ARMC3* into the *MLL* bcr. 265-271 bases from *MLL* were deleted during the translocation. No 5'-*ARMC3-AF10*-3' genomic breakpoint junction or fusion transcript was detected, indicating deletion of additional material on the der(10) chromosome. *ARMC3*, which encodes Arm repeats similar to catenin family proteins, is the first gene of this type involved in an *MLL* translocation.

(Proc 47th Annual Meeting American Society of Hematology

Introduction

Translocations disrupting an 8.3 kb breakpoint cluster region (bcr) between *MLL* exons 5-11 at chromosome band 11q23 are common molecular aberrations in infant acute leukemias and leukemias related to chemotherapeutic topoisomerase II poisons and less common alterations in adult acute leukemias and myelodysplastic syndrome (Gilliland, et al. 2004; Rowley and Olney 2002; Rowley, et al. 1996). *MLL* translocations are clinically important because of their unfavorable prognosis (Pui, et al. 2002). *MLL* translocations involve many partner genes that encode diverse partner proteins, ~50 of which have been identified through molecular cloning (Ayton and Cleary 2003). The MLL oncoprotein undergoes proteolytic cleavage into amino and carboxyl terminal segments that reassociate together in a multiprotein complex (Hsieh, et al. 2003a; Hsieh, et al. 2003b; Yokoyama, et al. 2002), which regulates maintenance of *HOX* gene expression and epigenetic modifications of nucleosomes and histones (Milne, et al. 2002). *MLL* translocations generate chimeric oncoproteins from the 5'-*MLL-partner gene*-3'

rearrangements lacking in the proteolytic cleavage site, which play key roles in leukemogenesis by altering the transcriptional regulation of MLL target genes (Ayton and Cleary 2003). However, the contribution of 5'-*partner gene-MLL*-3' rearrangements are not understood (Raffini, et al. 2002). Furthermore, although many *MLL* partner genes themselves encode important proteins in transcriptional regulation or cell signaling pathways in different cellular compartments (Ayton and Cleary 2001b; Ayton and Cleary 2003) the significance of their disruption in *MLL* leukemogenesis is unknown.

Most *MLL* rearrangements are simple reciprocal translocations; however, a notable exception is the t(10;11)(p12;q23), which fuses *MLL* with *AF10* and occurs primarily in AML (Beverloo, et al. 1995; Chaplin, et al. 1995a; Chaplin, et al. 1995b; Chaplin, et al. 2001). Due to the 3' to 5' orientation of *AF10* relative to the centromere and the 5' to 3' orientation of *MLL,* 5'-*MLL-AF10-3'* rearrangements occur as complex insertions and inversions disrupting multiple genes or chromosomes, the 5'-*partner gene-MLL*-3' junction invariably involves a different partner gene than *AF10*, and reciprocal 5'-*AF10-MLL-3'* fusion transcripts are not formed (Chaplin, et al. 2001). Here we characterized a novel three-way translocation of *MLL* with *AF10.* The partner gene in the 5'-*partner gene-MLL*-3' junction was the mammalian *armadillo* homologue *ARMC3*, which encodes armadillo repeat motifs similar to those in other members of the catenin protein family, the cardinal member of which b-catenin has been broadly implicated in cancer (Brembeck, et al. 2006), hematopoiesis and leukemogenesis (Jamieson, et al. 2004). The translocation resulted in a 5'-*ARMC3-MLL*-3' transcript that would alter and truncate ARMC3 in addition to 5'-*MLL-AF10*-3' fusion transcripts.

Materials and methods

Patients and samples

The IRB at the Children's Hospital of Philadelphia approved this research. Patient t-36 had received vincristine, etoposide, and ifosphamide without irradiation for alveolar rhabdomyosarcoma diagnosed at 16 months of age. Five years after the primary cancer diagnosis, the diagnosis of secondary FAB M4 AML with a normal karyotype was made. The clinical history, Southern blot analysis of the *MLL* bcr in *Bam*HI digested DNA from the bone marrow at the time of secondary leukemia

diagnosis, and cDNA panhandle PCR characterization of 5'-*MLL-AF10-3'* fusion transcripts (GenBank Nos. AF272383, AF272384, AF272385) were previously described (Megonigal, et al. 2000b).

Reverse panhandle PCR analysis of unknown 5'-*partner gene-MLL*-3' genomic breakpoint junction

The 5'-*partner gene-MLL*-3' genomic breakpoint junction was cloned from *Bam*HI digested DNA from the bone marrow at the time of leukemia diagnosis by reverse panhandle PCR (Raffini, et al. 2002) as modified by LoNigro *et al.* to use two primers instead of one in the first and nested PCRs (LoNigro, in preparation). The nested reverse panhandle PCR products were purified using a Geneclean III kit (Qbiogene; Carlsbad, CA) and subcloned by the TOPO TA Cloning method (Invitrogen, Carlsbad, CA). Recombinant plasmids were screened by PCR (Megonigal, et al. 2000a; Megonigal, et al. 2000b), and subclones of interest were sequenced using automated methods.

The 5'-*partner gene-MLL*-3' genomic breakpoint junction was confirmed by PCR with the sense primer 5'-TTC TGT GTG TTT GTC TTG TCT CTG AG-3' from positions 1625 to 1650 relative to the start of *ARMC3* intron 17 (GenBank No. NM_173081) and the antisense primer 5'-TGC GGG AGA TTC AGA GTA CTA AAG T-3' from *MLL* bcr positions 6619 to 6595 in intron 8 (GenBank No. U04737).

Reverse Transcriptase (RT) PCR analysis

The 5'-*ARMC3-MLL*-3' fusion transcript was identified by RT-PCR. Random hexamer primed first strand cDNA was generated using the High-Capacity cDNA Archive Kit (Applied Biosystems; Foster City, CA) and amplified with the sense primer 5'-CAA AGA ATG GTG TCC TCC CTC T-3' from *ARMC3* exon 17 (cDNA positions 2213-2234; GenBank No. NM_173081) and the antisense primer 5'-CAA AAC TTG TGG AAG GGC TCA C-3' from *MLL* exon 9 (cDNA positions 4376-4355; GenBank No. L04284). A TOPO TA Cloning Kit was used to subclone the RT-PCR products into the pCR® 2.1-TOPO® vector (Invitrogen, Carlsbad, CA) for automated sequencing.

Characterization of 5'-MLL-AF10-3' genomic breakpoint junction

The *5'-MLL-AF10-3'* genomic breakpoint junction was amplified with gene specific primers. The sense and antisense primers were 5'-TAG GTT TGA CCA ATT GTC CCA ATA-3' and 5'-CAC CTT TAA GTC TTG CCC ATG AAC-3', from *MLL* bcr positions 5948 to 5971 in intron 8 (GenBank No. U04737), and positions 2875 to 2852 relative to the start of *AF10* intron 9 (GenBank No. U13948), respectively.

Results

BamHI reverse panhandle PCR uncovers 5'-ARMC3-MLL-3' breakpoint junction

Although the karyotype of the secondary AML of patient t-36 was normal, the prior etoposide exposure and FAB M4 morphology led to Southern blot evaluation for a cryptic *MLL* translocation, which detected ~14.5 and 4.2 kb *MLL* bcr rearrangements in *Bam*HI digested DNA (Megonigal, et al. 2000b). As previously reported, cDNA panhandle PCR revealed three in-frame transcripts joining either *MLL* exon 7 or 8 to *AF10* exon 10, the latter of which also was alternatively spliced in *AF10* (Megonigal, et al. 2000b), but no product was obtained by *Bam*HI based panhandle PCR as designed for der(11) breakpoint junctions (data not shown).

Here, newer *Bam*HI reverse panhandle PCR enabled cloning the unknown 5'-*partner gene*-*MLL*-3' genomic breakpoint junction contained in the 4.2 kb *Bam*HI rearrangement (fig. 1.2.3A). The sequence of the 4094 bp product included the *ARMC3 (ARMADILLO REPEAT CONTAINING 3)* gene from chromosome band 10p12 (GenBank No. NM_173081) at the 5' side fused to the 3' portion of the *MLL* bcr. The *ARMC3* breakpoint was position 1901 or 1902 relative to the start of intron 17. The *MLL* bcr breakpoint at position 6473 or 6474 in intron 8 was 5' to the secondary leukemia *MLL* translocation breakpoint hotspot (Whitmarsh, et al. 2003). Identical C residues at the *ARMC3* and *MLL* breakpoints precluded more precise breakpoint determinations and suggested NHEJ DNA repair typical of *MLL* translocations (Gillert, et al. 1999; Raffini, et al. 2002; Weinstock, et al. 2006; Whitmarsh, et al. 2003). PCR with *ARMC3* sense and *MLL* antisense primers gave the predicted 424 bp product (data not shown), direct sequencing of which confirmed the 5'-*ARMC3*-*MLL*-3' breakpoint junction.

The novel 5'-*ARMC3-MLL*-3' rearrangement produced an out-of-frame transcript with *ARMC3* exon 17 and *MLL* exon 9 at the point of fusion (fig. 12.3B). The transcript contained a termination codon just 15 codons after the point of fusion and would alter and truncate the carboxyl terminus of ARMC3.

PCR with gene-specific primers identifies 5'-MLL-AF10-3' genomic breakpoint junction

Southern blot analysis, restriction mapping, the 5'-*MLL-AF10*-3' fusion transcripts and the 5'-*ARMC3-MLL*-3' genomic breakpoint junction informed gene-specific primers to clone the 5'-*MLL-AF10*-3' genomic breakpoint junction. Direct sequencing of the 394 bp product in fig. 2.2.3 showed that *MLL* bcr position 6202, 6203, 6204, 6205, 6206 or 6207 in intron 8 was fused to position 2737, 2738, 2739, 2740, 2741 or 2742 relative to the start of *AF10* intron 9. Comparison of the 5'-*MLL-AF10*-3' and 5'-*ARMC3-MLL*-3' genomic breakpoint junctions indicated that 265-271 bases from *MLL* intron 8 were deleted during the complex translocation. The homologous 5'-ATTAG-3' sequences at the breakpoints in *MLL* and *AF10* are consistent with DNA repair by NHEJ (Gillert, et al. 1999; Raffini, et al. 2002; Whitmarsh, et al. 2003), whereas the deleted sequence suggests multiple sites of damage to the *MLL* bcr or extensive processing of a single break site (Super, et al. 1997).

ARMC3 is adjacent and centromeric to *AF10* at chromosome band 10p12 and *ARMC3* and *AF10* have the same 5' to 3' orientation from telomere to centromere (http://www.ncbi.nlm.nih.gov/mapview). The chromosomal locations of *AF10* and *ARMC3* suggest that the complex 3-way rearrangement occurred by splitting and insertion of chromosome 10p12 material into the *MLL* bcr at band 11q23 (fig. 3.2.3).

Multiple additional PCRs with gene-specific primer combinations did not reveal a 5'-*AF10-ARMC*-3' genomic breakpoint junction which was predicted to have formed on the der(10) when the region from the 3' portion of the *AF10* gene through to the 5' portion of *ARMC3* was translocated into *MLL* (fig. 3.2.3). RT-PCR with gene specific primers also did not show any evidence of a 5'-*AF10-ARMC*-3' fusion transcript. These results are consistent with loss of additional material from the der(10) chromosome during the complex rearrangement.

Discussion

We uncovered a 5'-*ARMC3*-*MLL*-3' rearrangement in a cryptic complex three way *MLL* translocation involving *AF10,* indicating that the case of secondary AML was marked by chromosomal translocation of a member of the gene family encoding characteristic armadillo (Arm) repeats, which has broad importance in cancer, hematopoiesis and leukemogenesis. *ARMC3* (*ARMADILLO REPEAT CONTAINING 3*) on chromosome 10p12.31 is a new partner gene in 5'- *partner gene*-*MLL*-3' rearrangements. The 2831 bp *ARMC3* transcript with 19 exons predicting a 872 amino acid soluble cytoskeleton protein was recently found to exhibit organ specific expression in skeletal muscle, lung, prostate and testis (Li, et al. 2006). Detection of *ARMC3* mRNA in bronchial epithelium as well as in ependymomas, which arise in a ciliated region of the brain, has suggested a potential role of ARMC3 in the genesis of cilia (Lonergan, et al. 2006).

The Arm repeat is an ~40 amino acid long tandemly repeated sequence motif first identified in the *Drosophila* segment polarity gene armadillo (Riggleman, et al. 1989). The new MLL partner protein ARMC3 contains 10 Arm repeats. Characteristic Arm repeats occur in the mammalian catenin protein family, which includes a-catenin, b-catenin, g-catenin/plakoglobin, plakophilin and $p120^{ctn}$, the latter two of which are members of the d subfamily of catenin family proteins (Anastasiadis and Reynolds 2000; Hatzfeld 2007; McCrea and Park 2007). The catenin protein family, which has human disease and cancer relevance, is defined by the presence of a repeat domain consisting of at least six individual Arm motifs (Hatzfeld 1999). The adenomatous polyposis coli (APC) gene also is a member of this family; however, it has unique features separate from the classical catenins (Hatzfeld 1999). Relationships of ARMC3 to other catenin family proteins are shown in FIG 4.

b-catenin, the human homologue of *Drosophila* armadillo and the founding member of the catenin protein family has dual cell signaling functions involving cell adhesion in cadherin complexes and transcriptional regulation (Harris and Peifer 2005). b-catenin has a central place in the Wnt signaling pathway. Wnt signaling inhibits post-translational b-catenin degradation, enabling b-catenin-TCF/LEF (T-cell transcription factor/lymphoid enhancer binding factor) interactions (Chung, et al. 2002). Furthermore, b-catenin is also regulated by Notch signaling, another critical determinant of hematopoietic cell fate and

leukemogenesis (Hayward, et al. 2005). Observations of abundant b-catenin expression in primary AML and a case of T-cell ALL and heterogeneous b-catenin expression in hematopoietic cell lines have suggested a potential role of b-catenin in leukemia cell proliferation, adhesion and survival (Chung, et al. 2002). The activated nuclear form of b-catenin regulates the self-renewal and proliferation of chronic myelogenous leukemia (CML) granulocyte-macrophage progenitors, the leukemia stem cells in blast crisis CML (Jamieson, et al. 2004). The potential cancer relevance of $p120^{ctn}$ interactions with cadherin complexes, Rho family GTPases and the transcription factor Kaiso are under study also (McCrea and Park 2007; Xiao, et al. 2007). The discovery of the 5'-*ARMC3*-*MLL*-3' junction in a complex 3-way translocation in the case of secondary AML raises questions as to whether ARMC3 is one of a growing number of Arm-related proteins with significance in cancer because ARMC3 shares the conserved structural feature of Arm-repeat domains, the hallmark of catenin proteins (fig. 5.2.3). Discovery that the new 5'-*ARMC3*-*MLL*-3' rearrangement creates a fusion transcript that would alter and truncate the ARMC3 carboxyl terminus provides the foundation to elucidate the place of ARMC3 in MLL leukemogenesis and, possibly, more generally in malignancies associated with abnormal catenin signaling pathways.

5'-*MLL*-*AF10*-3' rearrangements occur as complex translocations with inversions or insertions because of the telomere to centromere orientation of *AF10* on chromosome 10p12 (Chaplin, et al. 2001; Klaus, et al. 2003). Recent interest in MLL-AF10 leukemogenesis involves the discovery that the octapeptide motif and leucine zipper regions of AF10 retained in MLL-AF10 fusion proteins interact with the histone H3K79 methyltransferase hDOT1L (Okada, et al. 2005). Even though MLL-AF10 is associated with aggressive features and short latency in murine retroviral transplantation models of MLL leukemogenesis (So and Cleary 2003), the nature of the novel partner gene in the 5'-*ARMC3*-*MLL*-3' rearrangement raises questions about whether disruption of the partner protein by the 5'-*partner gene*-*MLL*-3' rearrangement provides a second 'critical hit' (He, et al. 2000) in leukemias with *MLL* translocations. There is precedent in APL that reciprocal RARA-partner gene and partner gene-RARA rearrangements both contribute to the leukemogenic phenotype (He, et al. 2000). Similar studies are warranted in *MLL* leukemias where there are myriad partner proteins with

important cellular functions and 5'-*partner gene-MLL*-3' transcripts are often present also (Raffini, et al. 2002).

References

-Anastasiadis PZ, Reynolds AB. 2000. The p120 catenin family: complex roles in adhesion, signaling and cancer. J Cell Sci 113 (Pt 8):1319-34.Ayton PM, Cleary ML. 2001a. MLL in Normal and Malignant Hematopoiesis. In: Ravid K, Licht JD, editors. Transcription Factors: Normal and Malignant Development of Blood Cells. New York: Wiley-Liss, Inc.

-Ayton PM, Cleary ML. 2001b. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. Oncogene 20(40):5695-707.

Ayton PM, Cleary ML. 2003. Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. Genes Dev 17(18):2298-307.

-Beverloo HB, Coniat ML, Wijsman J, Lillington DM, Bernard O, de Klein A, van Wering E, Welborn J, Young BD, Hagemeijer A and others. 1995. Breakpoint heterogeneity in t(10;11) translocation in AML-M4/M5 resulting in fusion of *AF10* and *MLL* is resolved by fluorescent *in situ* hybridization analysis. Cancer Res 55:4220-4224.

-Brembeck FH, Rosario M, Birchmeier W. 2006. Balancing cell adhesion and Wnt signaling, the key role of beta-catenin. Curr Opin Genet Dev 16(1):51-9.

-Chaplin T, Ayton P, Bernard OA, Saha V, Valle VD, Hillion J, Gregorini A, Lillington D, Berger R, Young BD. 1995a. A novel class of zinc finger/leucine zipper genes identified from the molecular cloning of the t(10;11) translocation in acute leukemia. Blood 85:1435-1441.

-Chaplin T, Bernard O, Beverloo HB, Saha V, Hagemeijer A, Berger R, Young BD. 1995b. The t(10;11) translocation in acute myeloid leukemia (M5) consistently fuses the leucine zipper motif of *AF10* onto the *HRX* gene. Blood 86:2073-2076.

-Chaplin T, Jones L, Debernardi S, Hill AS, Lillington DM, Young BD. 2001. Molecular analysis of the genomic inversion and insertion of AF10 into MLL suggests a single-step event. Genes Chromosomes Cancer 30(2):175-80.

-Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 31(13):3497-500.

-Chung EJ, Hwang SG, Nguyen P, Lee S, Kim JS, Kim JW, Henkart PA, Bottaro DP, Soon L, Bonvini P and others. 2002. Regulation of leukemic cell adhesion, proliferation, and survival by beta-catenin. Blood 100(3):982-90.

-Daser A, Rabbitts TH. 2005. The versatile mixed lineage leukaemia gene MLL and its many associations in leukaemogenesis. Semin Cancer Biol 15(3):175-88.

-DiMartino JF, Ayton PM, Chen EH, Naftzger CC, Young BD, Cleary ML. 2002. The AF10 leucine zipper is required for leukemic transformation of myeloid progenitors by MLL-AF10. Blood 99(10):3780-5.

-Gillert E, Leis T, Repp R, Reichel M, Hosch A, Breitenlohner I, Angermuller S, Borkhardt A, Harbott J, Lampert F and others. 1999. A DNA damage repair mechanism is involved in the origin of chromosomal translocations t(4;11) in primary leukemic cells. Oncogene 18:4663-4671.

-Gilliland DG, Jordan CT, Felix CA. 2004. The molecular basis of leukemia. Hematology (Am Soc Hematol Educ Program):80-97.

-Harris TJ, Peifer M. 2005. Decisions, decisions: beta-catenin chooses between adhesion and transcription. Trends Cell Biol 15(5):234-7.

-Hatzfeld M. 1999. The armadillo family of structural proteins. Int Rev Cytol 186:179-224.

-Hatzfeld M. 2007. Plakophilins: Multifunctional proteins or just regulators of desmosomal adhesion? Biochim Biophys Acta 1773(1):69-77.

-Hayward P, Brennan K, Sanders P, Balayo T, DasGupta R, Perrimon N, Martinez Arias A. 2005. Notch modulates Wnt signalling by associating with Armadillo/beta-catenin and regulating its transcriptional activity. Development 132(8):1819-30.
-He LZ, Bhaumik M, Tribioli C, Rego EM, Ivins S, Zelent A, Pandolfi PP. 2000. Two critical hits for promyelocytic leukemia. Mol Cell 6(5):1131-41.
-Hess JL. 2004. MLL: a histone methyltransferase disrupted in leukemia. Trends Mol Med 10(10):500-7.
-Hsieh JJ, Cheng EH, Korsmeyer SJ. 2003a. Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. Cell 115(3):293-303.
-Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P, Korsmeyer SJ. 2003b. Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. Mol Cell Biol 23(1):186-94.
-Jamieson CH, Ailles LE, Dylla SJ, Muijtjens M, Jones C, Zehnder JL, Gotlib J, Li K, Manz MG, Keating A and others. 2004. Granulocyte-macrophage progenitors as candidate leukemic stem cells in blast-crisis CML. N Engl J Med 351(7):657-67.
-Klaus M, Schnittger S, Haferlach T, Dreyling M, Hiddemann W, Schoch C. 2003. Cytogenetics, fluorescence in situ hybridization, and reverse transcriptase polymerase chain reaction are necessary to clarify the various mechanisms leading to an MLL-AF10 fusion in acute myelocytic leukemia with 10;11 rearrangement. Cancer Genet Cytogenet 144(1):36-43.
-Li X, Liu B, Ji CN, Kang Y, Mao Y. 2006. Cloning and expression of ARMC3_v2, a novel splicing variant of the human ARMC3 gene. Genetika 42(7):999-1003.
-Linder B, Newman R, Jones LK, Debernardi S, Young BD, Freemont P, Verrijzer CP, Saha V. 2000. Biochemical analyses of the AF10 protein: the extended LAP/PHD-finger mediates oligomerisation. J Mol Biol 299(2):369-78.
-Lonergan K, Chari R, deLeeuw R, Shadeo A, Chi B, Tsao M-S, Jones S, Marra M, Ling V, Ng R and others. 2006. Identification of novel lung genes in bronchial epithelium by serial analysis of gene expression. Am J Respir Cell Mol Biol 35(6):651-661.
-McCrea PD, Park JI. 2007. Developmental functions of the P120-catenin sub-family. Biochim Biophys Acta 1773(1):17-33.
-Megonigal MD, Cheung NK, Rappaport EF, Nowell PC, Wilson RB, Jones DH, Addya K, Leonard DG, Kushner BH, Williams TM and others. 2000a. Detection of leukemia-associated MLL-GAS7 translocation early during chemotherapy with DNA topoisomerase II inhibitors. Proc Natl Acad Sci U S A 97(6):2814-9.
Megonigal MD, Rappaport EF, Wilson RB, Jones DH, Whitlock JA, Ortega JA, Slater DJ, Nowell PC, Felix CA. 2000b. Panhandle PCR for cDNA: A rapid method for isolation of -
-MLL fusion transcripts involving unknown partner genes. Proc Natl Acad Sci U S A 97(17):9597-9602.
-Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, Hess JL. 2002. MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell 10(5):1107-17.
-Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, Su L, Xu G, Zhang Y. 2005. hDOT1L links histone methylation to leukemogenesis. Cell 121(2):167-78.
-Pui CH, Gaynon PS, Boyett JM, Chessells JM, Baruchel A, Kamps W, Silverman LB, Biondi A, Harms DO, Vilmer E and others. 2002. Outcome of treatment in childhood acute lymphoblastic leukaemia with rearrangements of the 11q23 chromosomal region. Lancet 359(9321):1909-15.
-Raffini LJ, Slater DJ, Rappaport EF, Lo Nigro L, Cheung N-KV, Biegel JA, Nowell PC, Lange BJ, Felix CA. 2002. Panhandle and reverse panhandle PCR enable cloning of der(11) and der(other) genomic breakpoint junctions of *MLL* translocations and identify complex translocation of *MLL, AF-4* and *CDK6*. Proc Natl Acad Sci U S A 99:4568-4573.
-Riggleman B, Wieschaus E, Schedl P. 1989. Molecular analysis of the armadillo locus: uniformly distributed transcripts and a protein with novel internal repeats are associated with a Drosophila segment polarity gene. Genes Dev 3(1):96-113.

-Robinson BW, Slater DJ, Felix CA. 2006. BglII-based panhandle and reverse panhandle PCR approaches increase capability for cloning der(II) and der(other) genomic breakpoint junctions of MLL translocations. Genes Chromosomes Cancer 45(8):740-53.
-Rowley JD, Olney HJ. 2002. International workshop on the relationship of prior therapy to balanced chromosome aberrations in therapy-related myelodysplastic syndromes and acute leukemia: overview report. Genes Chromosomes Cancer 33(4):331-45.
-Rowley JD, Vignon C, Gollin SM, Rosenberg CL, Wyandt HE, Milunsky A. 1996. Rearrangements involving chromosome band 11q23 in acute leukemia. N Engl J Med 334:601-603.
-So CW, Cleary ML. 2003. Common mechanism for oncogenic activation of MLL by forkhead family proteins. Blood 101(2):633-9.
-Super HG, Strissel PL, Sobulo OM, Burian D, Reshmi SC, Roe B, Zeleznik-Le NJ, Diaz MO, Rowley JD. 1997. Identification of complex genomic breakpoint junctions in the t(9;11) MLL-AF9 fusion gene in acute leukemia. Genes Chromosomes Cancer 20(2):185-95.
-Weinstock DM, Richardson CA, Elliott B, Jasin M. 2006. Modeling oncogenic translocations: distinct roles for double-strand break repair pathways in translocation formation in mammalian cells. DNA Repair (Amst) 5(9-10):1065-74.
-Whitmarsh R, Saginario C, Zhuo Y, Hilgenfeld E, Rappaport EF, Megonigal MD, Carroll M, Liu M, Osheroff N, Cheung N-KV and others. 2003. Reciprocal DNA topoisomerase II cleavage events at 5'-TATTA-3' sequence in *MLL* and *AF-9* create homologous single-stranded overhangs that anneal to form der(11) and der(9) genomic breakpoint junctions in treatment-related AML without further processing. Oncogene 22:8448-8459.
-Xiao K, Oas RG, Chiasson CM, Kowalczyk AP. 2007. Role of p120-catenin in cadherin trafficking. Biochim Biophys Acta 1773(1):8-16.
-Yokoyama A, Kitabayashi I, Ayton PM, Cleary ML, Ohki M. 2002. Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties. Blood 100(10):3710-8.

Figure 1

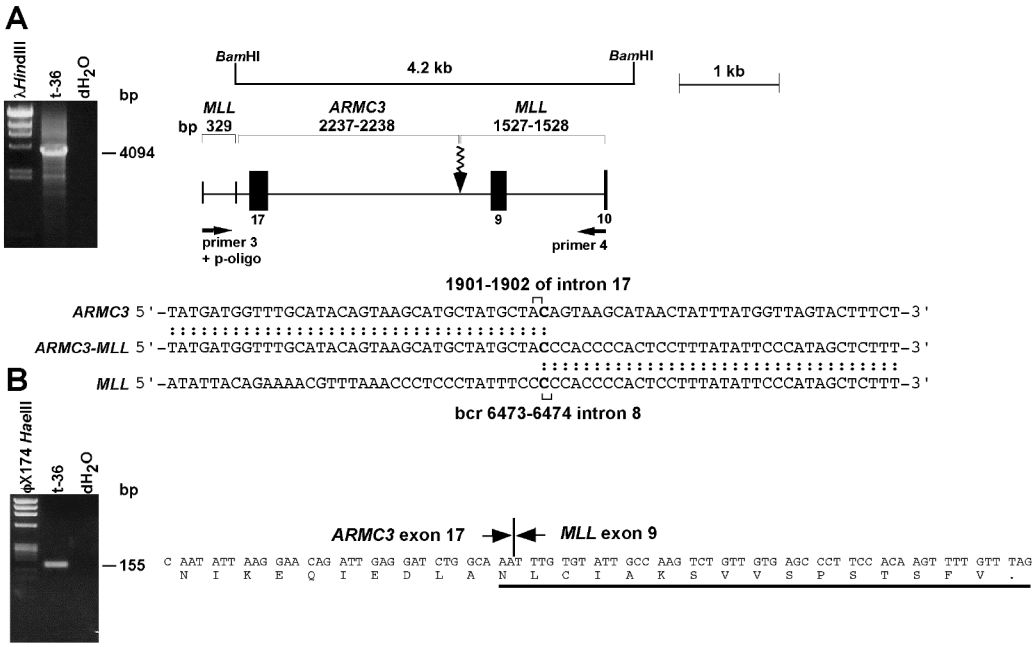


Figure 1.2.3. Molecular identification of *5'-ARMC3-MLL-3'* rearrangement in secondary AML. (A) Reverse panhandle PCR product (left), the size of which indicated that 4.2 kb *Bam*HI rearrangement on prior Southern blot (Megonigal, et al. 2000b) contained the 5'-*partner gene-MLL*-3' genomic breakpoint junction. Subclone sequencing (right) identified 5'-*ARMC3-MLL*-3' genomic breakpoint junction. Relationship of breakpoint junction to proximal intron-exon structure in both genes is shown. 329 bp of 5' sequence are *MLL* primer 3 through phosphorylated (P)-oligonucleotide (LoNigro, In preparation), which was ligated to the unknown partner DNA to form the panhandle template. The adjacent 2237-2238 bp include the 5' *Bam*HI site in *ARMC3* intron 16 through the *ARMC3* translocation breakpoint at position 1901 or 1902 relative to start of intron 17 (GenBank No. NM_173081). The 3' 1527-1528 bp include *MLL* bcr sequence beginning at the translocation breakpoint at position 6473 or 6474 in intron 8, through *MLL* primer 4 (bcr positions 7971-8000), which has been used before (Robinson, et al. 2006). C at breakpoints in both genes (bold) suggests NHEJ repair (Weinstock, et al. 2006). (B) Fusion transcript from 5'-*ARMC3-MLL*-3' rearrangement identified by reverse-transcriptase PCR. cDNA sequence at right shows out-of-frame fusion of *ARMC3* exon 17 and *MLL* exon 9; predicted amino acid sequence of ARMC3 with altered, foreshortened carboxyl terminus (thick underline) is at bottom right.

Figure 2

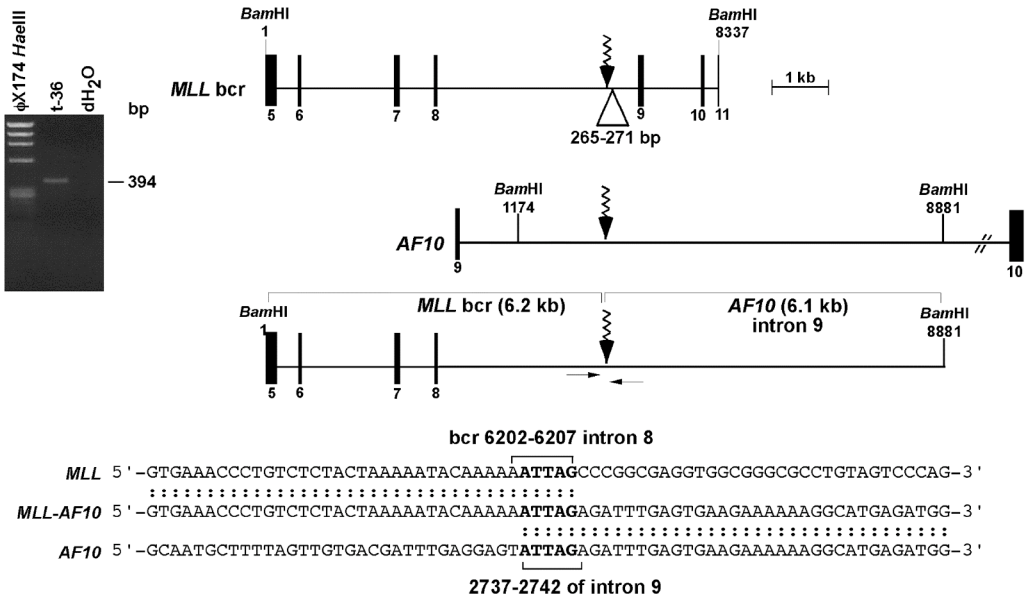


Figure 2.2.3. Characterization of 5'-*MLL-AF10*-3' genomic breakpoint junction. Known 5'-*MLL-AF10*-3' transcripts fusing *MLL* exon 8 and *AF10* exon 10 (Megonigal, et al. 2000b) and position of *MLL* genomic breakpoint in 5'-*ARMC3-MLL*-3' rearrangement informed gene-specific primers (arrows) to amplify 5'-*MLL-AF10*-3' genomic breakpoint junction contained in estimated ~14.5 kb *Bam*HI fragment on prior Southern blot (Megonigal, et al. 2000b). *Bam*HI restriction maps of *MLL* bcr (GenBank No. U04737), involved genomic region of *AF10* (GenBank No. U13948) and 5'-*MLL-AF10*-3' rearrangement are at top. Jagged arrows indicate translocation breakpoints. Coordinates of *Bam*HI sites are relative to start of *MLL* bcr or start of *AF10* intron 9. Breakpoint junction sequence at bottom indicates that 14.5 kb (Megonigal, et al. 2000b) was an overestimate of the rearrangement size. Breakpoint positions (brackets) could not be assigned exactly due to identical 5'-ATTAG-3' in *MLL* and *AF10*. Comparison with 5'-*ARMC3-MLL*-3' breakpoint junction (FIG 1) indicated deletion of bases shown by triangle from positions 6203, 6204, 6205, 6206, 6207 or 6208 to position 6472 or 6473 of *MLL* bcr during the translocation.

Figure 3

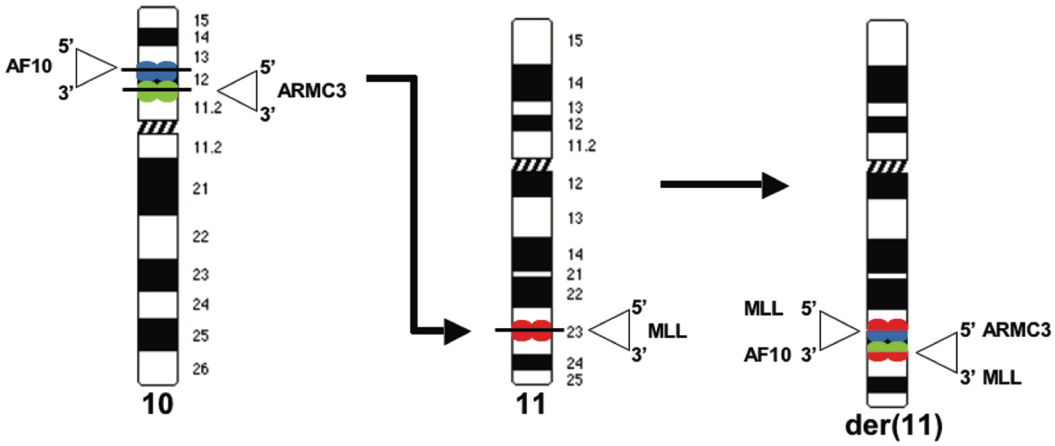


Figure 3.2.3. Schematic of formation of complex rearrangement of *MLL*, *AF10*, and *ARMC3*. *ARMC3* and *AF10* are adjacent genes with 5' to 3' telomere to centromere orientations at chromosome band 10p12, and *ARMC3* is more centromeric. 5'-*MLL*-*AF10*-3' and 5'-*ARMC3*-*MLL*-3' breakpoint junctions are both formed on der(11) when *MLL* bcr is split by insertion of chromosome 10p material containing 3' portion of *AF10* through 5' portion of *ARMC3*. No 5'-*AF10*-*ARMC*3-3' rearrangement was detected on putative der(10) chromosome.

Figure 4

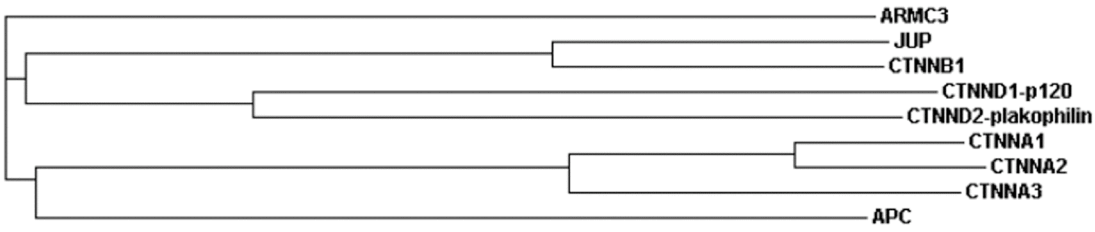


Figure 4.2.3. Dendrogram depicting relationship of ARMC3 to other catenin family proteins. Amino acid sequences of ARMC3 (armadillo repeat containing 3, NM_173081), CTNNA1 (catenin alpha 1, NM_001903), CTNNA2 (catenin alpha 2, NM_004389), CTNNA3 (catenin alpha 3, NM_013266), CTNNB1 (catenin beta 1, NM_001904), JUP (junction plakoglobin, NM_002230), CTNND1-p120 (catenin delta 1, NM_001331), CTNND2-plakophilin (catenin delta 2, NM_001332) and APC (adenomatous polyposis coli, NM_000038), all of which contain Arm domains, were aligned and compared using the ClustalW program (http://www.ebi.ac.uk/clustalw/) to produce a consensus phylogram tree (Chenna, et al. 2003). The analysis suggests that ARMC3 is most closely related to APC of the catenin family proteins, though clearly ARMC3 is in a class by itself, unlike the alpha and delta subfamilies.

Figure 5

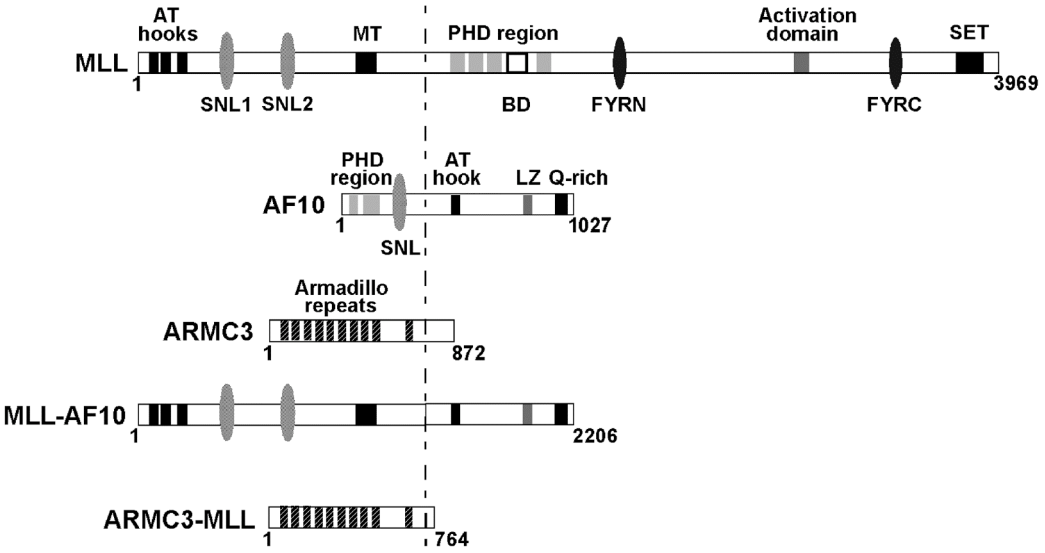


Figure 5. Schematics of native MLL, AF10 and ARMC3 proteins and predicted fusion proteins from complex rearrangement. Native MLL and AF10 domain structures are as described in (Ayton and Cleary 2001a; Ayton and Cleary 2001b; Daser and Rabbitts 2005; DiMartino, et al. 2002; Hess 2004; Linder, et al. 2000). Similar to previously reported MLL-AF10 fusion proteins (DiMartino, et al. 2002), the MLL-AF10 fusion protein joins the amino terminal AT-hook motifs and methyltransferase (MT) domain of MLL to the carboxyl terminal leucine zipper (LZ) domain of AF10. The amino terminus of the ARMC3-MLL protein retains nearly all of ARMC3 including all 10 Arm repeats but the amino acid sequence of the ARMC3 carboxyl terminus changes at the point of fusion and truncates prematurely and the MLL plant homeodomain (PHD), transactivation domain and carboxyl terminal SET domain are not present because of the out-of-frame sequence of fusion transcript.

General references

1 Daser A, Rabbitts TH The versatile mixed lineage leukaemia gene MLL and its many associations in leukaemogenesis. Semin Cancer Biol. 2005;15(3):175-88..

2 Ayton PM, Cleary ML. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. Oncogene 2001;20(40):5695-707M.D.

3 Megonigal MD, Rappaport EF, Jones DH, et al. Panhandle PCR strategy to amplify MLL genomic breakpoints in treatment-related leukemias. Proc Natl Acad Sci USA 1997;94(21):11583-8.

4 Felix CA, Kim CS, Megonigal MD, Slater DJ, Jones DH, Spinner NB, Stump T, Hosler MR, Nowell PC, Lange BJ, Rappaport EF.Panhandle polymerase chain reaction amplifies MLL genomic translocation breakpoint involving unknown partner gene. Blood. 1997;90(12):4679-86.

5 Gu Y, Alder H, Nakamura T, Schichman SA, Prasad R, Canaani O, Saito H, Croce CM, Canaani E.Sequence analysis of the breakpoint cluster region in the ALL-1 gene involved in acute leukemia.Cancer Res. 1994 May 1;54(9):2326-30.

6 Robinson BW, Slater DJ, Felix CA.BglII-based panhandle and reverse panhandle PCR approaches increase capability for cloning der(II) and der(other) genomic breakpoint junctions of MLL translocations. Genes Chromosomes Cancer. 2006;45(8):740-53.

7 Jones DH, Winistorfer SC. Sequence specific generation of a DNA panhandle permits PCR amplification of unknown flanking DNA. Nucleic Acids Res. 1992;20(3):595-600.

8 Raffini LJ, Slater DJ, Rappaport EF, et al. Panhandle and reverse-panhandle PCR enable cloning of der(11) and der(other) genomic breakpoint junctions of MLL translocations and identify complex translocation of MLL, AF-4, and CDK6. Proc Natl Acad Sci USA 2002;99(7):4568-73.

9 Megonigal MD, Cheung NK, Rappaport EF, et al. Detection of leukemia-associated MLL-GAS7 translocation early during chemotherapy with DNA topoisomerase II inhibitors. Proc Natl Acad Sci U S A 2000;97(6):2814-9.

10 Pui CH, Ribeiro RC, Hancock ML, Rivera GK, Evans WE, Raimondi SC, Head DR, Behm FG, Mahmoud MH, Sandlund JT, et al. Acute myeloid leukemia in children treated with epipodophyllotoxins for acute lymphoblastic leukemia. N Engl J Med. 1991;325(24):1682-7.

11 Kumar L.Epipodophyllotoxins and secondary leukaemia.Lancet. 1993;342(8875):819-20.

12 Fortune JM, Osheroff N. Topoisomerase II as a target for anticancer drugs: when enzymes stop being nice. Prog Nucleic Acid Res Mol Biol. 2000;64:221-53.

13 Berger JM, Gamblin SJ, Harrison SC, Wang JC.Structure and mechanism of DNA topoisomerase II. Nature. 1996;379(6562):225-32.

14 Suzuki H, Ikeda T, Yamagishi T, Nakaike S, Nakane S, Ohsawa M.Efficient induction of chromosome-type aberrations by topoisomerase II inhibitors closely associated with stabilization of the cleavable complex in cultured fibroblastic cells. Mutat Res. 1995;328(2):151-61.

15 Blanco JG, Edick MJ, Relling MV. Etoposide induces chimeric Mll gene fusions. FASEB J. 2004;18(1):173-5.

16 Libura J, Slater DJ, Felix CA, Richardson C.Therapy-related acute myeloid leukemia-like MLL rearrangements are induced by etoposide in primary human CD34+ cells and remain stable after clonal expansion.Blood. 2005;105(5):2124-31.
17 Moneypenny CG, Shao J, Song Y, Gallagher EP. MLL rearrangements are induced by low doses of etoposide in human fetal hematopoietic stem cells. Carcinogenesis. 2006;27(4):874-81.
18 Haupt R, Fears TR, Rosso P, Colella R, Loiacono G, de Terlizzi M, Mancini A, Comelli A, Indolfi P, Donfrancesco A, et al.Increased risk of secondary leukemia after single-agent treatment with etoposide for Langerhans' cell histiocytosis. Pediatr Hematol Oncol. 1994;11(5):499-507.
19 Rubin CM, Arthur DC, Woods WG, Lange BJ, Nowell PC, Rowley JD, Nachman J, Bostrom B, Baum ES, Suarez CR, et al.Therapy-related myelodysplastic syndrome and acute myeloid leukemia in children: correlation between chromosomal abnormalities and prior therapy. Blood. 1991;78(11):2982-8.
20 Felix CA, Hosler MR, Winick NJ, Masterson M, Wilson AE, Lange BJ. *ALL-1* gene rearrangements in DNA topoisomerase II inhibitor-related leukemia in children. Blood 1995;85:3250-6.
21 Megonigal MD, Cheung NK, Rappaport EF, et al. Detection of leukemia-associated MLL-GAS7 translocation early during chemotherapy with DNA topoisomerase II inhibitors. Proc Natl Acad Sci U S A 2000;97(6):2814-9.
22 Domer PH, Head DR, Renganathan N, Raimondi SC, Yang E, Atlas M. Molecular analysis of 13 cases of MLL/11q23 secondary acute leukemia and identification of topoisomerase II consensus-binding sequences near the chromosomal breakpoint of a secondary leukemia with the t(4;11). Leukemia. 1995;9(8):1305-12.
23 Super HG, Strissel PL, Sobulo OM, Burian D, Reshmi SC, Roe B, Zeleznik-Le NJ, Diaz MO, Rowley JD.Identification of complex genomic breakpoint junctions in the t(9;11) MLL-AF9 fusion gene in acute leukemia. Genes Chromosomes Cancer. 1997;20(2):185-95.
24 Gillert E, Leis T, Repp R, Reichel M, Hösch A, Breitenlohner I, Angermüller S, Borkhardt A, Harbott J, Lampert F, Griesinger F, Greil J, Fey GH, Marschalek R.A DNA damage repair mechanism is involved in the origin of chromosomal translocations t(4;11) in primary leukemic cells. Oncogene. 1999;18(33):4663-71.
25 Felix CA, Hosler MR, Slater DJ, Megonigal MD, Lovett BD, Williams TM, Nowell PC, Spinner NB, Owens NL, Hoxie J, Croce CM, Lange BJ, Rappaport EF. Duplicated regions of AF-4 intron 4 at t(4;11) translocation breakpoints. Mol Diagn. 1999;4(4):269-83.
26 Rasio D, Schichman SA, Negrini M, Canaani E, Croce CM. Complete exon structure of the ALL1 gene. Cancer Res. 1996;56(8):1766-9.
27 Cimino G, Rapanotti MC, Biondi A, Elia L, Lo Coco F, Price C, Rossi V, Rivolta A, Canaani E, Croce CM, Mandelli F, Greaves M. Infant acute leukemias show the same biased distribution of ALL1 gene breaks as topoisomerase II related secondary acute leukemias. Cancer Res. 1997;57(14):2879-83.
28 Slater DJ, Hilgenfeld E, Rappaport EF, et al. MLL-SEPTIN6 fusion recurs in novel translocation of chromosomes 3, X, and 11 in infant acute myelomonocytic leukaemia and in t(X;11) in infant acute myeloid leukaemia,

and MLL genomic breakpoint in complex MLL-SEPTIN6 rearrangement is a DNA topoisomerase II cleavage site. Oncogene 2002;21(30):4706-14.

CHAPTER 3

MLL IN HEMATOPOIESIS AND LEUKEMIA IN ZEBRAFISH MODEL

Recently, the zebrafish (*Danio rerio*), initially used as a classical developmental and embryological model, become an attractive complement to mouse model for studying hematopoietic development and disorders such as leukemia. Zebrafish belong to the genus of teleosts or bony fish. The split between fish and mammals during evolution dates back 300 million years, but the genetic programs between these organisms are largely conserved. The advanced genome sequencing, the ease of studying gene expression by in situ hybridization, and the availability of various tools for manipulating gene function make zebrafish a powerful model for dissecting the regulation and function of genes.

In this chapter the attention will be focused firstly on the general aspects of zebrafish hematopoiesis and then, as part of the ongoing research about a zebrafish model for the study of *MLL*-related leukemias, will be devoted to describe cDNA sequences, phylogenetic analysis, temporal gene expression and gene-linkage mapping of zebrafish *mll*.

3.1 Zebrafish as a model system for hematopoiesis and leukemia

Zebrafish offer several distinct advantages for developmental and genetic studies including high fecundity, short generation time and small size at maturation[(1)]. The rapid, easily visualized, external development of transparent embryos enables real-time functional observations of hematopoietic development unlike any other models, and blood circulation in zebrafish becomes visible under the microscope by 24 hours postfertilization (hpf) [(2)].

Large segments of zebrafish chromosomes are syntenic with human and mouse genomes[(3)]. Moreover, many mammalian genes have zebrafish orthologs, suggesting that they have evolved from the same ancestral genes sharing common functions[(3)]. Zebrafish has proven especially suitable for the study of hematopoiesis: erythroid, myeloid, and lymphoid lineages have been characterized and shown to follow a similar developmental program to that seen in humans[(4-7)].

In addition, many zebrafish orthologs of blood-specific genes already have been isolated (e.g. *cmyb*, *gata1*, *gata2*, *globin*, *ikaros*, *lmo2*, *pu.1*, *rag1*, *rag2, runx1*, *cbfb,* and *scl*)[8-12]. Gene expression profiling of kidney marrow cells, the site of definitive hematopoiesis in teleosts, has demonstrated that the genetic programs controlling hematopoiesis, angiogenesis and hematopoietic cell function are highly conserved from zebrafish to humans[13].

Although the anatomic sites of hematopoiesis differ, zebrafish hematopoiesis and blood cell morphology closely parallel those of mammals[14]. In mammals, primitive hematopoiesis is largely erythropoietic and extra-embryonic in blood islands of the yolk sac. Later in embryogenesis, mammalian hematopoiesis moves to the aorta-gonad-mesonephros (AGM) and the fetal liver[15], whereas definitive hematopoiesis occurs in the bone marrow where all blood cell lineages are produced[16]. In contrast, zebrafish lack extra-embryonic yolk sac blood islands and primitive hematopoiesis occurs within the intermediate cell mass (ICM) between notochord and endoderm, anteriorly over the yolk cell in the anterior lateral mesoderm (ALM) and posteriorly in a small ventral cluster of cells called posterior lateral mesoderm (PLM)[17,18]. By 10-12 hpf the PLM expresses *scl*, *gata2* and *lmo2*, indicating the formation of hematopoietic stem cells (HSCs)[19,20]. At 12-20 hpf initiation of erytropoiesis is marked by *gata1* expression in a subset of *scl*+ cells in the PLM, whereas myelopoiesis and granulopoiesis, marked by myeloid-specific gene expression (e.g. *pu.1*, *l-plastin*) begins in the ALM[21]. Thus, the PLM and ALM give rise to erythroid and myeloid cells, respectively. By 24 hpf, proerythroblasts from the ICM expressing *gata1* and embryonic globins begin to enter circulation[2]. By 31 hpf, expression of zebrafish *c-myb* and *runx1* orthologs on HSCs herald definitive hematopoiesis in the kidney, and definitive HSCs subsequently colonize the thymus and pancreas[20]. By >96 hpf myelopoiesis occurs in the kidney and the spleen as indicted by MPO+, PAS-, Acid Phosphatase+ cells and *mpo* and *pu.1* gene expression[22]. At 5 dpf, erythrocytes and granulocytes are produced in the kidney and by 13 dpf onward the kidney marrow is the primary hematopoietic organ[23,24]. However, zebrafish have only two granulocyte lineages, one resembling mammalian neutrophils and the second, produced in the spleen and kidney, with features of both mammalian eosinophils and basophils[21,25,26]. Monocyte/ macrophages expressing *c-myb* and *l-plastin* but not the neutrophil marker *mpo* have been identified in zebrafish embryos by 12-20

hpf and in the kidney and spleen of adult fish[2,25]. There is *rag1* expression and evidence of thymic development by 65-75 hpf, and the thymus is fully mature with medullary and cortical tissues and *tcra* gene expression by 3 weeks of age[27]. There is some evidence that B cells first develop in the zebrafish pancreas as evidenced by *rag1* transcripts as early as 3-4 dpf[28,29].

Importantly, zebrafish orthologs have been identified for several known mammalian proto-oncogenes and tumor-suppressor genes involved in leukemogenesis[30-34]. Gene expression profiling has revealed that several *MLL* partner genes are represented in the zebrafish genome[13]. The high evolutionary conservation reinforces the notion that zebrafish is a worthwhile model for investigating hematopoiesis and leukemia.

The utility of zebrafish to induce and study leukemia has been demonstrated recently using transgenic technology. By transiently expressing the human AML-associated *RUNX1-CBF2T1* fusion oncogene under control of the CMV promoter in zebrafish embryos Kalev-Zylinska *et al.* reproduced the hematopoietic defects seen in *RUNX1-CBF2T1* transgenic mice[30]. A transient *TEL-JAK2* fusion oncoprotein transgenic zebrafish also recently was generated[35]. In addition, Langenau et al. reported the first stable transgenic zebrafish, in which expression of a murine c-*Myc-GFP* under control of the *rag2* promoter induced clonal, transplantable T-cell ALL[36]. Finally, zebrafish of the *TEL-AML1* fusion cDNA led to the development of oligoclonal B-lineage ALL[37]. These studies and the cloning of an *Mll* ortholog in *Fugu rubripes* (pufferfish) with functionally similar domains to its mammalian counterparts[38], support the feasibility of modeling the roles MLL and MLL fusion proteins using zebrafish.

3.2 Cloning and characterization of the zebrafish *mll* ortholog

At the outset of this work, little was known about the existence and relationship, if any, of a zebrafish mll ortholog to human MLL. First I utilized bioinformatics tools to interrogate this question. BLASTP searching on the NCBI database server (http://www.ncbi.nlm.nih.gov/BLAST/) using the full-length human MLL protein (GenBank accession no. NP_005924) as the reference sequence identified two putative “similar to MLL proteins” containing 2251 amino

acids (GenBank no. XP_685032) and 1904 amino acids (GenBank no. XP_685116).

GenBank entries for two predicted transcript sequences, XM_680024 and XM_679940, corresponding to the two "similar to MLL proteins" also were identified using BLAST. The two predicted transcript sequences spanned positions 31,979,440 to 31,961,790 and 31,961,430 to 31,952,674, respectively, in close proximity to each other on zebrafish chromosome 15.

The more 5' 5715nt sequence XM_680024 contained 17 predicted exons, whereas there were 7108 bases and 18 predicted exons in XM_679940. Furthermore, Ensembl (www.ensembl.org) projected that the two sequences comprised a single larger transcript (Entrez Gene 557048). Following, the group of Sun *et al.* deposited in the GenBank database partial transcript sequences at the central portion of this region cloned from zebrafish kidney marrow (DQ355790 and DQ355791). The relationship of the two predicted "similar to MLL" transcript sequences to the predicted single transcript (Entrez Gene 557048) is shown in fig. 1.3.2.

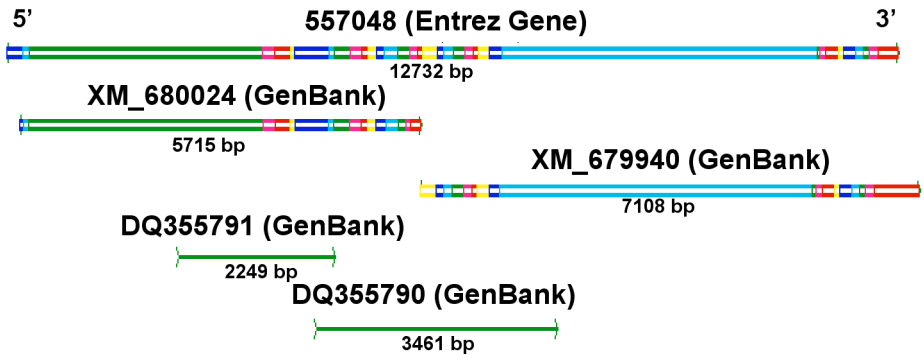


Figure 1.3.2. Predicted and partial cloned zebrafish *mll* cDNA sequences in Ensembl and GenBank databases. There are three predicted MLL transcripts. Entrez Gene (557048) contains complete projected transcript sequence with 35 exons and 12732 bases. Two GenBank entries (XM_680024 and XM_679940) were derived from an annotated genomic sequence (NW_633640) using GNOMON gene prediction method; both contain partial transcript sequences. Entrez Gene and GenBank predictions contain differences usually near exon-exon junctions. GenBank direct submissions (DQ355790 and DQ355791) comprising cloned sequences from zebrafish kidney marrow are from non-contiguous central portions of the full-length transcript.

The GNOMON gene prediction tool, which evaluates transcripts and proteins aligned to a genome (www.ncbi.nlm.nih.gov/genome/), then was used to predict

the genomic structure(s) corresponding to the two zebrafish "similar to MLL" protein sequences (GenBank nos. XP_685032 and XP_685116) in zebrafish genomic DNA. The results of GNOMON analysis also predicted that a single genomic sequence (GenBank accession no. NW_633640) matched both protein sequences.

Next, CDART analysis tools (http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi) were employed in order to compare human MLL and the two zebrafish "similar to MLL" proteins. The CDART algorithm finds protein similarities across significant evolutionary distances using protein domain architecture, i.e. the sequential order of conserved domains in proteins, rather than direct sequence similarity[39]. Interestingly, this analysis suggested that the two predicted proteins' together resembled mammalian MLL in its entirety and that important domains of human and mouse MLL including the CXXC domain, bromodomain, PHD zinc fingers, FYRN, FYRC and SET domain all were present in a hypothetical single zebrafish protein (fig. 2.3.2).

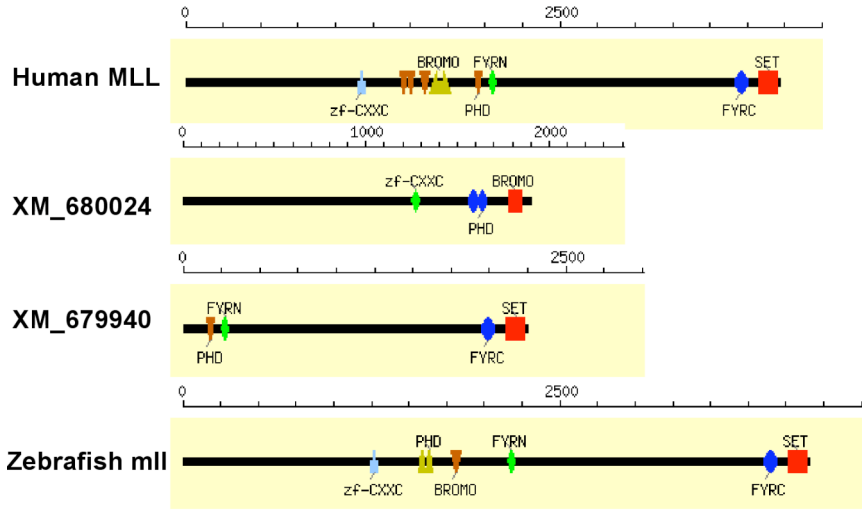
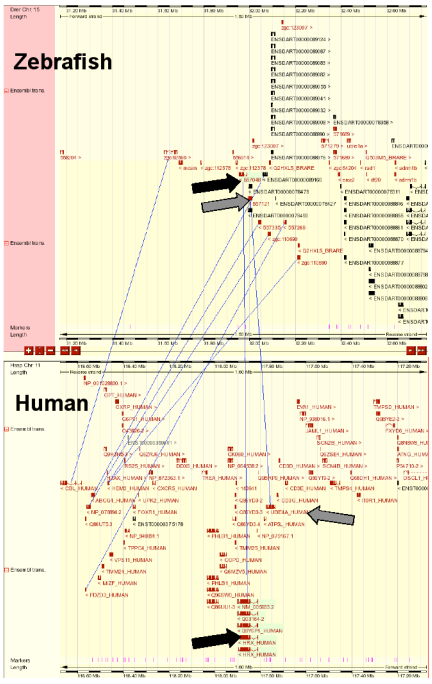


Figure 2.3.2. CDART analysis of human MLL protein and zebrafish "similar to MLL proteins". Conserved domains of human MLL, putative zebrafish "similar to mll" proteins, and hypothetical zebrafish mll obtained by joining of the two zebrafish "similar to MLL" protein sequences.

While there can be one-to-many and many-to-many relationships[40] between human and zebrafish genes due to gene duplication over evolutionary distance, the predictions of a single gene (GenBank accession no. NW_633640) and single larger transcript (Entrez Gene 557048) matching human *MLL* is most consistent

with a one-to-one relationship. Another question in comparing the predicted zebrafish *mll* gene to human *MLL* was whether zebrafish *mll* is an ortholog, i.e. a gene evolved from a common ancestral gene with the same function, or a paralog that arose by duplication with a different function[40]. Because the syntenic relationship[3] between genes is an important predictor of functional similarity, the Ensembl database was utilized to examine synteny between the predicted zebrafish *mll* and human *MLL* genes.

This analysis revealed that there was a conserved block of synteny surrounding zebrafish *mll* and human *MLL* containing several linked genes. In addition, zebrafish *mll* and human *MLL* are in the same map order in similar uninterrupted segments with the gene for ubiquitination factor E4A (fig. 3.3.2).



Genes in Synteny Group	
Human	Zebrafish
UBE4A	557121
NLRX1	557335
PDZD3	557269
HMBS	zgc:110690
CBL	zgc:92560
ABCG4	ENSDART00000089166

Figure 3.3.2. Analysis of synteny of chromosomal regions surrounding human *MLL* and zebrafish *mll*. Ensembl genes were compared within 1.6 Mb regions centered around human *MLL* at chromosome band 11q23 and putative zebrafish *mll* ortholog on chromosome 15. Human *MLL* and zebrafish *mll* and human (UBE4A) and zebrafish (557121) genes for ubiquitination factor E4A are found in same map order in similar uninterrupted segments, and several other genes are within a synteny group.

Thus the existence of a single zebrafish *mll* gene with functional similarity to human *MLL* was supported by several gene and protein prediction methods as well as the syntenic relationship indicated by the respective surrounding genes. This prediction was further strengthened by the prior characterization in pufferfish

(*fugu)*, a teleost more closely related to the zebrafish, of a single *Mll*-like gene with structural similarity and high overall sequence identity to human *MLL*[38]. These initial experiments asked whether cross-species counterparts of amino acid sequences of highly conserved domains of MLL could be used to identify the corresponding orthologous zebrafish transcript sequence. The amino acid sequences from MLL domains determined by ClustalW analysis (http://www.ebi.ac.uk/clustalw/) to be the most highly conserved across species from human through mouse, pufferfish and fly, were used to design degenerate primer mixtures for RT-PCR experiments in which oligo-dT primed cDNA prepared from total RNA from a whole wild-type adult zebrafish was examined.

In the experiment shown in fig. 4.3.2, degenerate RT-PCR with primers derived from the SET domain amino acid sequence produced a 203 bp product, with 99% sequence identity to the corresponding region of the predicted 3' XM_679940 transcript. Similarly, products could be generated in an additional degenerate RT-PCR experiment interrogating the transcript region corresponding to the PHD. These studies using degenerate primers demonstrate that transcript regions encoding specific MLL functional domains are highly conserved throughout evolution. Not only is there high cross-species homology at the amino acid sequence level (fig. 4.3.2), but also the cross-species counterparts of amino acid sequences could be used to generate the predicted transcript, providing the first experimental evidence that the transcript represented the *bona fide* orthologous *mll* gene from the zebrafish species.

Similarly, cross-species Southern blot analysis of zebrafish genomic DNA was performed using the B859 fragment containing exons 5-11 of the human *ALL-1* cDNA[41] to determine if the human probe would detect the predicted zebrafish *mll* gene. First, restriction maps were simulated for the enzymes *Bam*HI, *Bgl*II, *Hind*III, *Nhe*I, *Sac*I and *Xba*I from a projected 36,662 bp genomic sequence corresponding to the predicted single zebrafish *mll* cDNA (Entrez Gene 557048), and the region of highest homology to the human probe was used to project the restriction fragment sizes that would be detected (fig. 5.3.2A).

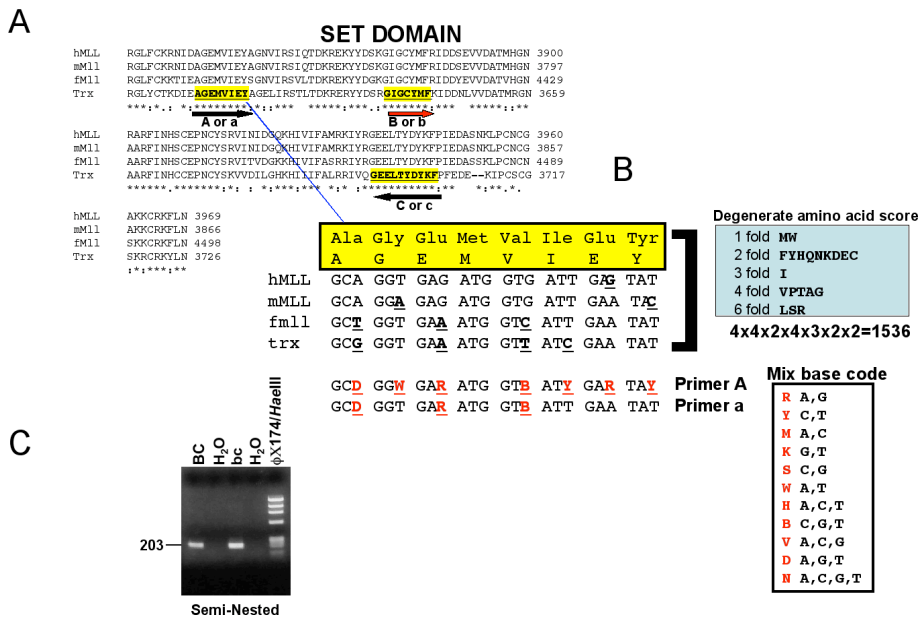


Figure 4.3.2. RT-PCR analysis of predicted zebrafish *mll* encoding SET domain with degenerate primers. A. Highly homologous SET domains from human, mouse, pufferfish (*fugu*) MLL and *Drosophila* trx aligned by ClustalW with identical amino acid sequences shown with asterisks. Degenerate primer mixtures for RT-PCR were designed from highlighted regions (yellow). B. Example of approach to primer design. Amino acid regions (yellow) were selected based on low degeneracy. Fold degeneracy of highlighted amino acid sequences in human, mouse, pufferfish (*fugu*) and *Drosophila* trx was determined from the product of the degenerate amino acid score (http://boneslab.bio.ntnu.no/degpcrshortguide.htm) after examining corresponding transcript regions for codons with a mismatched base. Two primers were designed for each highlighted amino acid sequence using the mixed base code. One (capitalized) counted as degenerate all positions with ≥1 mismatch(es) between the 4 species; the second (lower case) incorporated into the primer sequence any consensus base that matched in 3 or 4 of the species. C. Degenerate primer mixtures A and C or a and c were used in initial PCRs. A 2 ul aliquot of respective initial PCRs was used for semi-nested PCR with degenerate primer mixtures B and C (first lane) or b and c (third lane). Sequencing showed that products of both semi-nested PCRs corresponded to XM_679940 sequence predicted to be zebrafish ortholog of human *MLL*.

The sizes and numbers of hybridizing fragments in zebrafish genomic DNA exactly matched those predicted, except with *Hind*III where a single fragment was expected and two fragments were detected (fig. 5.3.2B). This difference is likely due to generation of the zebrafish *mll* genomic sequence with a gene prediction tool. Therefore, this experiment simulated and detected the putative genomic region corresponding to the human *MLL* bcr in zebrafish *mll*.

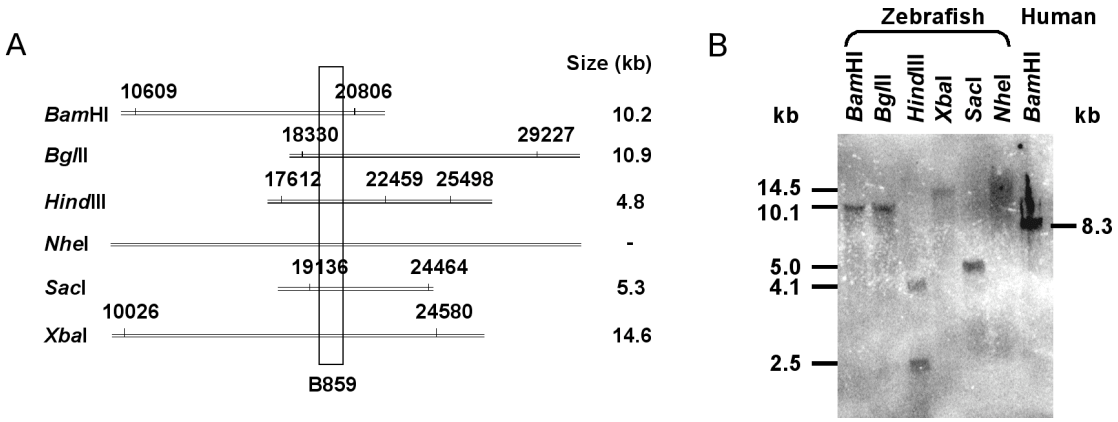


Figure 5.3.2. Cross-species Southern blot analysis of adult zebrafish genomic DNA with B859 fragment of human *ALL-1* cDNA. A. Simulated restriction mapping of predicted zebrafish *mll* genomic sequence. Insets from six restriction enzyme maps of predicted 36,662 bp zebrafish *mll* genomic sequence for Entrez Gene ID 557048 are shown in region most homologous to B859 fragment containing exons 5-11 of human *ALL-1* (*MLL*) cDNA. Approximately 90 bases of the predicted zebrafish *mll* cDNA sequence (Entrez Gene ID 557048) match the probe exactly. B. Autoradiograph of zebrafish genomic DNAs and normal human subject peripheral blood lymphocyte DNA after probing with B859. DNA was extracted from a whole wild type adult zebrafish using DNeasy tissue kit (Qiagen, Valencia, CA). 20 mg of zebrafish DNA was digested to completion with the indicated enzyme. 10 mg of *Bam*HI-digested human DNA was positive control. Conditions for electrophoresis, Southern transfer, nick translation and hybridization were those employed routinely for human DNAs[42].

Additional experiments utilized reverse transcriptase PCR (RT-PCR) analysis of total RNA from a whole wild-type adult zebrafish in order to investigate whether the two predicted "similar to MLL proteins", which, in turn, predicted transcript sequences in close proximity to each other on chromosome 15, were derived from a single gene encoding a putative zebrafish mll with functional domains similar to human MLL. The reaction products generated with a forward primer from the more 5' cDNA (XM_680024) and a reverse primer from the more 3' cDNA (XM_679940) are shown in fig. 6.3.2, indicating that the two cDNA sequences were derived from one gene.

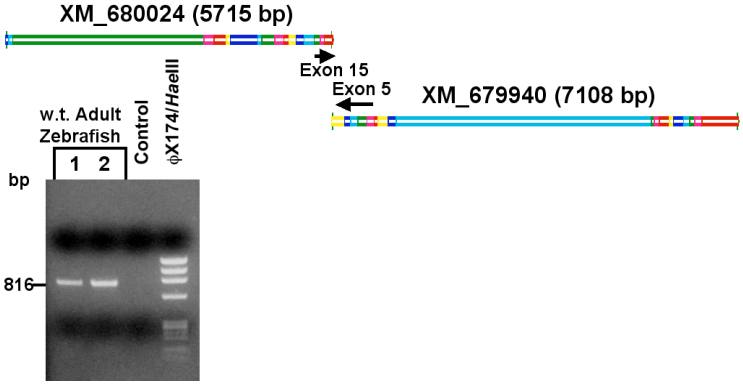


Figure 6.3.2. RT-PCR identification of single zebrafish *mll* transcript spanning predicted cDNAs for both "similar to MLL proteins". Total RNA was extracted from a whole wild-type adult zebrafish using TRIZOL reagent (Invitrogen). Oligo(dT) primed first strand cDNA was synthesized from 5 ug of total RNA using SuperScript II reverse transcriptase (Invitrogen). Sense primer (5'-GAG AGC AGG AAA GCC AAC AG-3') from exon 15 of XM_680024 and antisense primer (5'-TGG TTC AAG TCC ATT AAC AAA TTT TCT-3') from exon 5 of XM_679940 generated a single product that spanned both cDNAs, sequencing of which indicated that the two cDNAs are partial 5' and 3' sequences of a single gene.

Having determined that the zebrafish *mll* ortholog to human *MLL* was a single gene on chromosome 15, I applied the strategies of 5' RACE PCR and long-distance PCR in order to achieve and characterize a near-full length zebrafish *mll* cDNA. As summarized in fig. 1.3.2, the bioinformatics databases contain only partial sequences of predicted *mll* cDNAs and ~5 kb of cloned sequence from the central region of the gene.

Moreover, the cDNAs derived with gene prediction tools are not precise representations of the sequence especially at the exon boundaries. We utilized the 5' rapid amplification of cDNA ends (RACE) procedure[43] and information on the predicted sequence of exon 3 in zebrafish *mll* cDNA to analyze whole wild-type adult zebrafish total RNA and obtained the 561 bp product shown in fig. 7.3.2A containing the unknown 5' UTR.

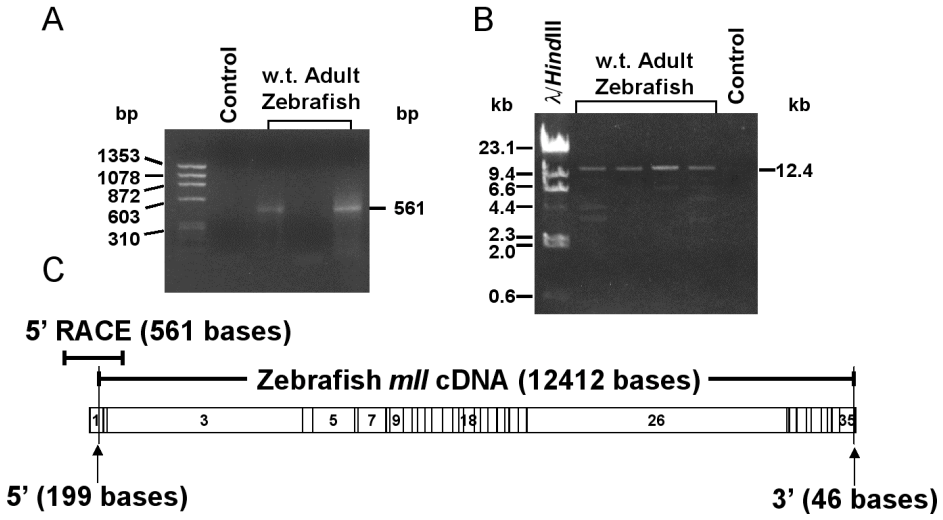


Figure 7.3.2. Cloning and sequencing of near complete zebrafish *mll* cDNA. Protocol was carried out using pooled aliquots of total RNAs from two whole wild-type adult zebrafish extracted with TRIZOL reagent (Invitrogen). A. PCR amplification of 5' UTR of zebrafish *mll* by 5' RACE. The 5' RACE PCR was performed in duplicate. Sequencing demonstrated the 5' UTR of zebrafish *mll*. B. Cloning of 12.4 kb fragment of zebrafish *mll* cDNA. Oligo-dT primed first strand cDNA was synthesized from total RNA from a whole wild-type adult zebrafish using SuperScript™ II First Strand Synthesis reagents (Invitrogen). RT-PCR was performed using Accuprime High Fidelity Taq Polymerase (Invitrogen) with a sense primer from exon 1 (5'-AAT TTC GGG ATG TTT TGG GGG AGT C'-3) and an antisense primer from exon 35 (5'-AGC TTA TTG CCT GGT TCT TCG ATG G-'3) designed from the sequence of Entrez Gene 557048. Five of 7 reactions generated the predicted 12.4 kb product, examples of which are shown. PCR products were gel-purified using a TOPO XL kit (Invitrogen) and subcloned into a TOPO XL vector (Invitrogen). Subclones with the desired insert were identified by PCR screen of bacterial mini-cultures with the exon 1 and exon 35 primers used for the original PCR. A 12412 bp sequence contig was generated from sequencing two subclones and directly sequencing the products of 3 independent PCRs derived from the embryos. C. Summary of 5' UTR and 35-exon overlapping sequence generated in A and B above, which together contain near complete zebrafish *mll* cDNA. The 199 bases at the most 5' end, which are not present in the 12412 base subclone, were obtained by 5' RACE, such that only 46 bases from the 3' end have not been cloned and sequenced.

In addition, using either total RNAs from pooled 24 hpf wild type zebrafish embryos or a whole wild-type zebrafish adult for RT-PCR with gene-specific primers from exons 1 and 35, I cloned a 12412 bp cDNA, the 5' end of which is overlapping with the product the 5' RACE (fig. 7.3.2B). A complete 12412 sequence contig has been generated that, together with the product of the 5' RACE, contains all but 46 bases at the 3' end of the complete zebrafish *mll* cDNA (fig. 7.3.2C).

Next I combined the 12412 bp zebrafish *mll* cDNA sequence and more 5' coding sequence that was generated together with the 46 missing 3' bases taken from Entrez Gene 557048 in order to compare the zebrafish *mll* cDNA and predicted protein to human *MLL* and its protein product. The human *MLL* cDNA (GenBank Accession no. L04284) contains 11910 bases and 36 exons, while there are 12657 bases and 35 exons in its zebrafish ortholog. ClustalW analysis of the zebrafish mll protein predicted by the cDNA clones indicated that zebrafish mll contains all of the same important functional domains as the human protein (fig. 8.3.2).

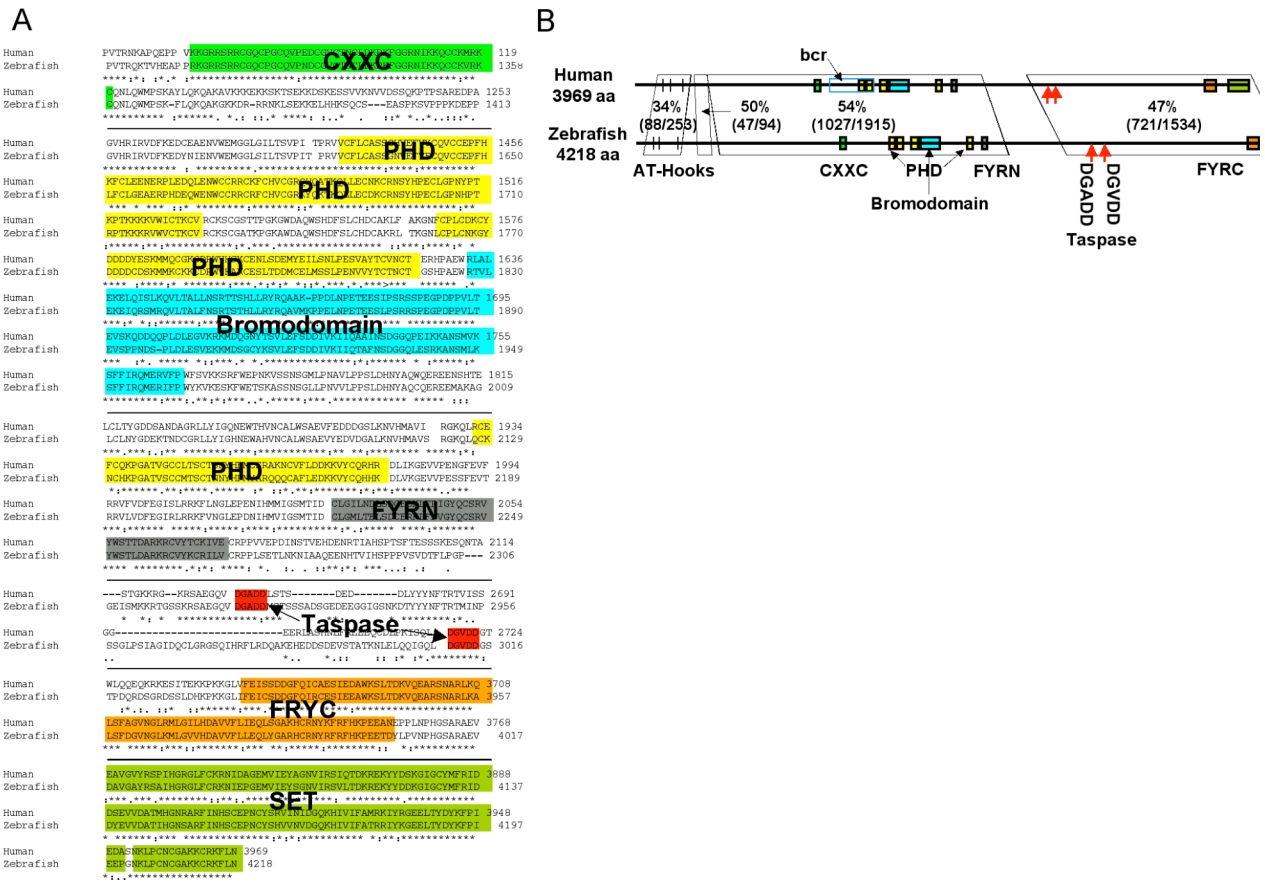


Figure 8.3.2. Conserved protein domains in zebrafish mll and human MLL. A. ClustalW alignment of human MLL protein sequence (GenBank accession no. AAA58669) and sequence of predicted zebrafish mll protein derived by assembling the cloned 12412 bp fragment and more 5' coding sequence in zebrafish *mll* cDNA with 46 missing 3' bases taken from Entrez Gene 557048. Colored regions indicate protein domains that both species have conserved. (N.B. Only regions of high sequence conservation are shown in the figure) B. Protein domain alignment generated using SMART (http://smart.embl-heidelberg.de/) and NCBI BLAST programs. Percent amino acid sequence identity determined by NCBI BLAST is indicated, and the domains found by SMART are color-coded to match the highlighted regions in the ClustalW alignment shown in A. Both proteins contain the same critical functional domains.

Human MLL is a 3969 amino acid protein, while there are 4218 amino acids in zebrafish mll with 45.7% sequence identity overall to the human protein. The highest amino acid sequence identity (54%) is in the central portion of the protein containing the plant homology domains, bromodomain and FYRN sequence, but there also is high amino acid sequence identity (50%) in a less well defined amino terminal region, and in the more carboxyl terminal portion of the protein where the taspase cleavage sites, FYRC and SET domain are located. The relationships of the zebrafish mll protein encoded by the *mll* cDNA sequence was further analyzed by ClustalW alignment with human MLL, as well as chimpanzee, dog, cow, mouse,

opossum, chicken and pufferfish mll protein sequences and the protein sequence of *Drosophila* trx in order to generate the phylogram tree shown in fig. 9.3.2.

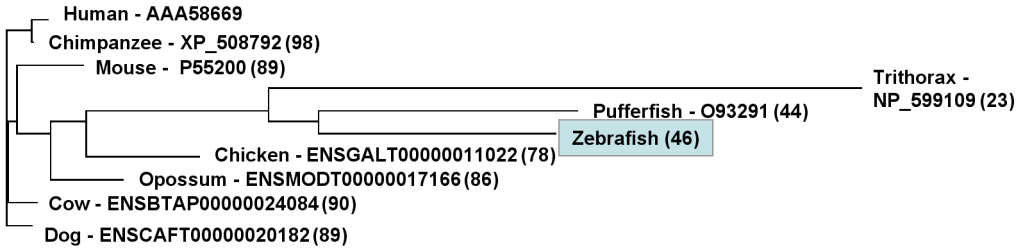


Figure 9.3.2. Phylogram tree of teleost, bird and mammalian MLL proteins and Drosophila trx. ClustalW was used to produce a consensus phylogram tree from alignment of MLL protein sequences of different species using human as the reference. GenBank accession numbers for human, mouse, pufferfish and trithorax isoform A are indicated. Protein sequence of zebrafish mll was derived from molecular cloning shown in figure 3.7 plus additional 46 3' bases from Entrez Gene 557048. Branch lengths represent the amount of evolutionary change. Numbers in parentheses represent amino acid identity to the human MLL as determined using EMBOSS Pairwise Alignment Algorithm Needle (http://www.ebi.ac.uk/emboss/align/).

As expected, this analysis revealed a closer relationship between the human, chimpanzee, mouse and other mammalian MLL proteins relative to chicken and the other non-mammalian vertebrates zebrafish and pufferfish and the more distant fly. Although the phylogram tree analysis suggests that the zebrafish mll ortholog reflects evolutionary divergence of mammals from teleosts, the high degree of protein sequence identity and conservation of critical functional domains from zebrafish to human (fig. 8.3.2) predicts that zebrafish may be a useful model system for elaborating the role of MLL in normal and malignant hematopoiesis. In the next experiments the temporal expression of *mll* RNA in zebrafish development was first examined by Northern blot analysis and quantitative RT-PCR (qRT-PCR) analysis. As shown in fig. 10.3.2, Northern blot analysis using total RNAs detected an abundant 12.6 kb signal consistent with *mll* transcript expression in the embryos at 2 hpf, and weak 12.6 kb signals at 24 hpf, 48 hpf, 72 hpf, 5 dpf and in the adult.

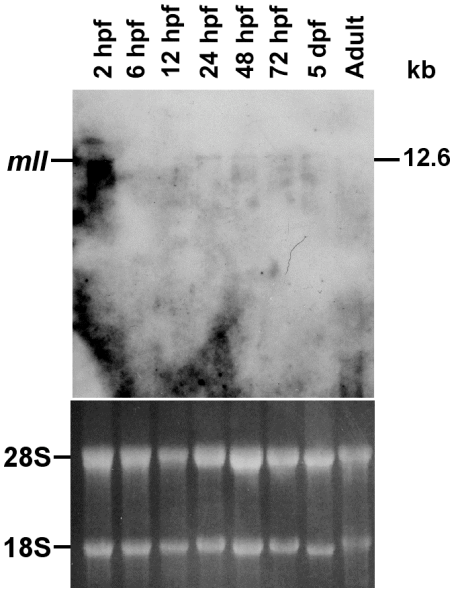


Figure 10.3.2. Northern blot analysis of *mll* RNA expression in wild type adult zebrafish and zebrafish embryos. Twenty mg of total RNA per lane from whole wild type adult or pooled wild type zebrafish embryos collected at indicated times were analyzed with 12.4 kb fragment of zebrafish *mll* cDNA. Conditions for electrophoresis, transfer, nick translation, and hybridization were those employed for human RNAs[42] except that no blocking DNA was used in hybridization. Corresponding ethidium-stained gel is at bottom. Hybridization with zebrafish *a1 tubulin* cDNA was used as control (data not shown).

Quantitative RT-PCR analysis, as showed in fig. 11.3.2, indicates that *mll* RNA is maternally supplied during the earliest timepoints in the development of the embryo. There was peak in zygotic *mll* expression at 12 hpf in the embryo and the highest relative expression occurred in the zebrafish adult (fig. 11.3.2).

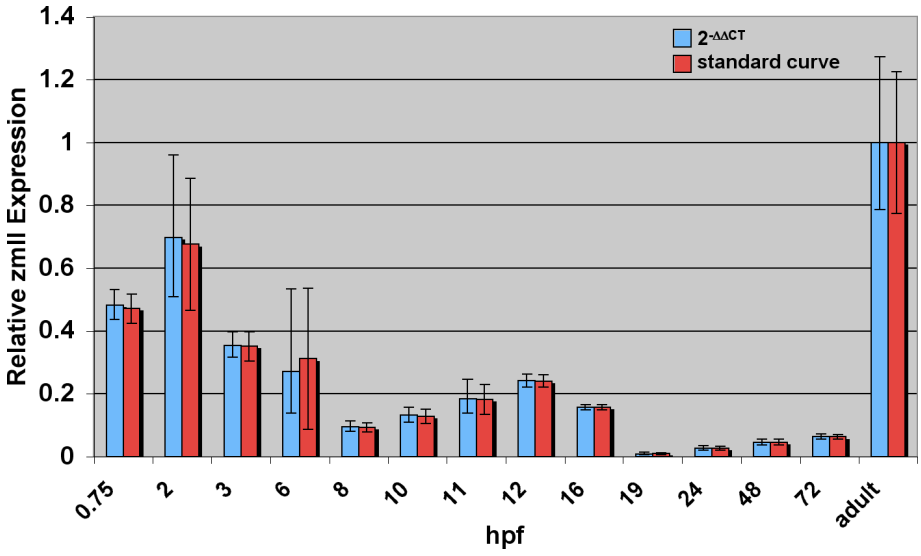


Figure 11.3.2. Quantitative real-time PCR analysis of *mll* mRNA expression during zebrafish development and adult. Embryos were pooled and sacrificed at the indicated times. Total RNAs were extracted using Trizol reagent and the RNAs were treated with DNase. Gene specific primers were designed to generate single gene-specific amplicons of 100-200 nucleotides. Sense and antisense *mll* specific primers 5'-CAA CCC TCA GGA GGA AGA TG-3' and 5'- CCT GCA GAA CAA ACC TCT GC-3', respectively, from positions 11921-11940 in exon 32 and positions 12086-12067 in exon 34 in the 3' region of *mll* cDNA corresponding to the SET domain (GenBank no. EF462416) were used to generate a plasmid subclone for construction of a standard curve[47] and the same primers were used for quantitative RT-PCR. Sense and antisense primers to amplify the beta bactin1 housekeeping gene were 5'-CGA GCA GGA GAT GGG AAC C-3' and 5'-CAA CGG AAA CGC TCA TTG C-3', respectively, corresponding to nucleotides 722-740 and 823-805 in intron 2 (GenBank no. NM_131031). To generate standard the curves, random hexamer primed first strand cDNA from the whole adult fish was amplified with the *mll* or *bactin1* specific primers, and the PCR products were used to generate plasmid subclones containing the relevant *mll* or *bactin1* amplicon in the TOPO TA vector (Invitrogen). Standard curves were constructed after performing quantitative real-time PCR on triplicate 10-fold serial dilutions of the linearized plasmids (10^9 to 10^2 copies per reaction) using SYBR green and the ABI 7900 HP detection system. The copy number for each reaction was calculated with the SDS software package (ABI). The standard curves had linear ranges between 10^2 and 10^8 molecules/ml and both slopes were -3.3. 1 ug of total RNAs from the embryos at the specified timepoints and from the zebrafish adult were used to synthesize random hexamer primed first strand cDNAs using Superscript II reverse transcriptase and a 1 ul aliquot from each cDNA reaction was analyzed in triplicate by quantitative real-time PCR using the same *mll* or *zbactin1* primers that were used to generate the standard curves. The mean *mll* copy number was normalized to the mean *zbactin1* copy number at each timepoint to determine normalized *mll* copy number from the standard curves. The red bars compare the normalized *mll* expression data derived from the standard curves by the absolute quantification method at each timepoint in embryogenesis to the normalized *mll* expression in the adult, with expression values in the embryos shown as fractions of the adult calibrator sample. In addition, the 2^{-ddCt} method[48] was used to analyze the relative changes in *mll* expression as a function of the age of the embryo compared to the adult with expression in the adult calibrator sample set to 1 (blue bars). Analysis of the data by the absolute (standard curve) and by the relative (2^{-ddCt}) quantitative methods both gave the same results.

It previously has been demonstrated that zygotic gene expression in zebrafish does not begin until 3 hpf and that maternal transcripts are degraded by 5 hpf, after which all transcripts are zygotic[45,46]. Therefore, the detection of a signal on Northern blot analysis, and the relative high levels of *mll* expression evaluated by quantitative RT-PCR at the earliest timepoint (0.75-3 hpf) is consistent with the presence of maternally supplied *mll* transcripts in the embryo. To analyze the spatiotemporal expression of *mll* during early embryonic development, whole-mount *in situ* hybridization was performed from two-cell stage to three-day-old embryos using antisense probe. *mll* transcripts were already detected at two-cell stage, thus pointing to a maternal origin of the transcript (data not shown). A specific pattern of zygotic *mll* expression can be detected from 10 hpf. At this

stage *mll* is strongly expressed in the head with marked staining in the midbrain-hindbrain boundary (fig. 12.3.2A-H). At this developmental point *mll* shows also a peculiar expression in few specific areas of the tail (fig. 12.3.2A-E,I). In particular, *mll* transcripts are localized in three stripes of cells within the last somites of the tail suggesting a segmental patterning of *mll*-expressing cells (fig. 12.3.2I). At later stages (36-72 hpf), *mll* expression is maintained only in the head (fig. 12.3.2F).

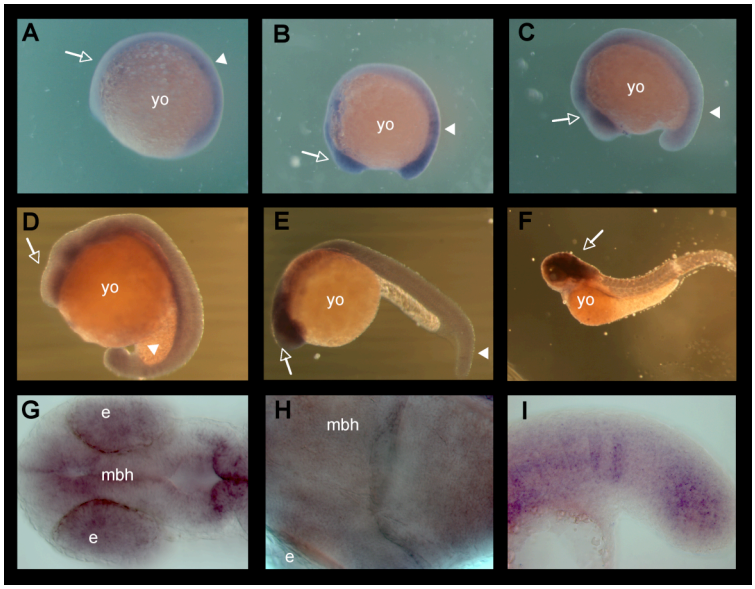


Figure 12.3.2. Expression of *mll* in zebrafish wild-type embryos. Lateral views showing the expression of z*mll* in developing embryos at 10 hpf (A), 11-12 hpf (B,C), 14 hpf (D), and 24-48 hpf (E,F) respectively. *mll* expression is observed in the developing ventral forebrain, midbrain and hindbrain (arrows) and in particular somites in the caudal tail (arrowhead). Dorsal and lateral views of high magnifications of 24 hpf embryos, showing expression of *mll* in midbrain-hindbrain boundary and in the caudal tail. Anterior is to the left. Abbreviations: mbh, midbrain-hindbrain boundary; e, eye; yo, yolk. The *mll* antisense probe was generated by PCR using sense 5'-CAA AGG GAC AGT GGG AGA GA-3' and antisense 5'-ACT TAA GAG GCA CAA TAC CAT TAA CC-3' primers, respectively. This template contained 957 bp of coding sequences (SET domain) and 142 bp of adjacent 3'-UTR. For synthesis of RNA probes the 1099 bp PCR fragment was cloned in a pCRII-TOPO plasmid. The plasmids were cut and transcribed with Sp6/T7 (Roche). Embryos were staged by morphological features, and developmental stages accordingly to Kimmel et al.[49]. Whole-mount *in situ* hybridization was carried out as previously described[50]. In Brief, embryos were permeabilized with Proteinase K (10ug/ml, Sigma) and hybridized overnight at 65°C with the DIG-labeled antisense or sense riboprobes. After several washes at high stringent temperature, NBT/BCIP (Roche) staining was performed according to the manufacturer's instructions. Images were obtained by a Nikon SMZ 1500 digital camera system.

In conclusion, these results indicate that there is a single *mll* gene with highly conserved functional similarity to human *MLL*. The temporal pattern of expression, including maternal supply of transcript to the embryo, indicates that *mll* is important from early embryogenesis through the entire lifespan of the fish.

The localization of *mll* transcripts in the CNS as well as in peculiar restricted domain suggests a function for *mll* in zebrafish development. The high evolutionary conservation of critical domains creates the starting point to use zebrafish for studying MLL in hematopoiesis and leukemia.

Finally, the functional validation of a zebrafish *mll* ortholog will provide the opportunity to use this model system for studies of leukemogenesis because defining the molecular programs controlled by MLL in normal cells and deregulated by the translocations is an avenue to new treatments.

General references

1. Hsu K, Kanki JP, Look AT. Zebrafish myelopoiesis and blood cell development. Curr Opin Hematol 2001;8(4):245-51.
2. de Jong JL, Zon LI. Use of the zebrafish system to study primitive and definitive hematopoiesis. Annu Rev Genet 2005;39:481-501.
3. Barbazuk WB, Korf I, Kadavi C, et al. The syntenic relationship of the zebrafish and human genomes. Genome Res 2000;10(9):1351-8.
4. Hogan BM, Pase L, Hall NE, Lieschke GJ. Characterisation of duplicate zinc finger like 2 erythroid precursor genes in zebrafish. Dev Genes Evol 2006.
5. Juarez MA, Su F, Chun S, Kiel MJ, Lyons SE. Distinct roles for SCL in erythroid specification and maturation in zebrafish. J Biol Chem 2005;280(50):41636-44.
6. Galloway JL, Wingert RA, Thisse C, Thisse B, Zon LI. Loss of gata1 but not gata2 converts erythropoiesis to myelopoiesis in zebrafish embryos. Dev Cell 2005;8(1):109-16.
7. Rhodes J, Hagen A, Hsu K, et al. Interplay of pu.1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish. Dev Cell 2005;8(1):97-108.
8. Gering M, Yamada Y, Rabbitts TH, Patient RK. Lmo2 and Scl/Tal1 convert non-axial mesoderm into haemangioblasts which differentiate into endothelial cells in the absence of Gata1. Development 2003;130(25):6187-99.
9. Nishikawa K, Kobayashi M, Masumi A, et al. Self-association of Gata1 enhances transcriptional activity in vivo in zebra fish embryos. Mol Cell Biol 2003;23(22):8295-305.
10. Blake T, Adya N, Kim CH, et al. Zebrafish homolog of the leukemia gene CBFB: its expression during embryogenesis and its relationship to scl and gata-1 in hematopoiesis. Blood 2000;96(13):4178-84.
11. Willett CE, Kawasaki H, Amemiya CT, Lin S, Steiner LA. Ikaros expression as a marker for lymphoid progenitors during zebrafish development. Dev Dyn 2001;222(4):694-8.
12. Burns CE, DeBlasio T, Zhou Y, Zhang J, Zon L, Nimer SD. Isolation and characterization of runxa and runxb, zebrafish members of the runt family of transcriptional regulators. Exp Hematol 2002;30(12):1381-9.
13. Song HD, Sun XJ, Deng M, et al. Hematopoietic gene expression profile in zebrafish kidney marrow. Proc Natl Acad Sci U S A 2004;101(46):16240-5.
14. Galloway JL, Zon LI. Ontogeny of hematopoiesis: examining the emergence of hematopoietic cells in the vertebrate embryo. Curr Top Dev Biol 2003;53:139-58.
15. Medvinsky A, Dzierzak E. Definitive hematopoiesis is autonomously initiated by the AGM region. Cell 1996;86(6):897-906.
16. Johnson GR, Moore MA. Role of stem cell migration in initiation of mouse foetal liver haemopoiesis. Nature 1975;258(5537):726-8.
17. Thompson MA, Ransom DG, Pratt SJ, et al. The cloche and spadetail genes differentially affect hematopoiesis and vasculogenesis. Dev Biol 1998;197(2):248-69.
18. Detrich HW, 3rd, Kieran MW, Chan FY, et al. Intraembryonic hematopoietic cell migration during vertebrate development. Proc Natl Acad Sci U S A 1995;92(23):10713-7.

19. Davidson AJ, Ernst P, Wang Y, et al. cdx4 mutants fail to specify blood progenitors and can be rescued by multiple hox genes. Nature 2003;425(6955):300-6.
20. Davidson AJ, Zon LI. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. Oncogene 2004;23(43):7233-46.
21. Bennett CM, Kanki JP, Rhodes J, et al. Myelopoiesis in the zebrafish, Danio rerio. Blood 2001;98(3):643-51.
22. Crowhurst MO, Layton JE, Lieschke GJ. Developmental biology of zebrafish myeloid cells. Int J Dev Biol 2002;46(4):483-92.
23. Willett CE, Cortes A, Zuasti A, Zapata AG. Early hematopoiesis and developing lymphoid organs in the zebrafish. Dev Dyn 1999;214(4):323-36.
24. Weinstein BM, Schier AF, Abdelilah S, et al. Hematopoietic mutations in the zebrafish. Development 1996;123:303-9.
25. Herbomel P, Thisse B, Thisse C. Ontogeny and behaviour of early macrophages in the zebrafish embryo. Development 1999;126(17):3735-45.
26. Lieschke GJ, Oates AC, Crowhurst MO, Ward AC, Layton JE. Morphologic and functional characterization of granulocytes and macrophages in embryonic and adult zebrafish. Blood 2001;98(10):3087-96.
27. Zapata A, Diez B, Cejalvo T, Gutierrez-de Frias C, Cortes A. Ontogeny of the immune system of fish. Fish Shellfish Immunol 2006;20(2):126-36.
28. Danilova N, Steiner LA. B cells develop in the zebrafish pancreas. Proc Natl Acad Sci U S A 2002;99(21):13711-6.
29. Lam SH, Chua HL, Gong Z, Lam TJ, Sin YM. Development and maturation of the immune system in zebrafish, Danio rerio: a gene expression profiling, in situ hybridization and immunological study. Dev Comp Immunol 2004;28(1):9-28.
30. Kalev-Zylinska ML, Horsfield JA, Flores MV, et al. Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis. Development 2002;129(8):2015-30.
31. Kataoka H, Ochi M, Enomoto K, Yamaguchi A. Cloning and embryonic expression patterns of the zebrafish Runt domain genes, runxa and runxb. Mech Dev 2000;98(1-2):139-43.
32. Lieschke GJ, Oates AC, Paw BH, et al. Zebrafish SPI-1 (PU.1) marks a site of myeloid development independent of primitive erythropoiesis: implications for axial patterning. Dev Biol 2002;246(2):274-95.
33. Schreiber-Agus N, Horner J, Torres R, Chiu FC, DePinho RA. Zebra fish myc family and max genes: differential expression and oncogenic activity throughout vertebrate evolution. Mol Cell Biol 1993;13(5):2765-75.
34. Kratz E, Eimon PM, Mukhyala K, et al. Functional characterization of the Bcl-2 gene family in the zebrafish. Cell Death Differ 2006;13(10):1631-40.
35. Onnebo SM, Condron MM, McPhee DO, Lieschke GJ, Ward AC. Hematopoietic perturbation in zebrafish expressing a tel-jak2a fusion. Exp Hematol 2005;33(2):182-8.
36. Langenau DM, Traver D, Ferrando AA, et al. Myc-induced T cell leukemia in transgenic zebrafish. Science 2003;299(5608):887-90.
37. Sabaawy HE, Azuma M, Embree LJ, Tsai HJ, Starost MF, Hickstein DD.TEL-AML1 transgenic zebrafish model of precursor B cell acute lymphoblastic leukemia. Proc Natl Acad Sci U S A. 2006 Oct 10;103(41):15166-71.

38. Caldas C, Kim MH, MacGregor A, Cain D, Aparicio S, Wiedemann LM. Isolation and characterization of a pufferfish MLL (mixed lineage leukemia)-like gene (fMll) reveals evolutionary conservation in vertebrate genes related to Drosophila trithorax. Oncogene 1998;16(25):3233-41.
39. Reaman GH. Biology and treatment of acute leukemia in infants in treatment of acute leukemias. In: Pui C-H, editor. New Directions in Clinical Research: Humana Press; 2003. p. 73-83.
40. Geer LY, Domrachev M, Lipman DJ, Bryant SH. CDART: protein homology by domain architecture. Genome Res 2002;12(10):1619-23.
41. Tatusov RL, Koonin EV, Lipman DJ. A genomic perspective on protein families. Science 1997;278(5338):631-7.
42. 157. Felix CA, Kim CS, Megonigal MD, et al. Panhandle polymerase chain reaction amplifies MLL genomic translocation breakpoint involving unknown partner gene. Blood 1997;90(12):4679-86.
43. Gu Y, Nakamura T, Alder H, et al. The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to Drosophila trithorax, to the AF-4 gene. Cell 1992;71(4):701-8.
44. Frohman MA, Dush MK, Martin GR. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci U S A 1988;85(23):8998-9002.
45. Chatterjee B, Chin AJ, Valdimarsson G, et al. Developmental regulation and expression of the zebrafish connexin43 gene. Dev Dyn 2005;233(3):890-906.
46. Christie TL, Mui R, White TW, Valdimarsson G. Molecular cloning, functional analysis, and RNA expression analysis of connexin45.6: a zebrafish cardiovascular connexin. Am J Physiol Heart Circ Physiol 2004;286(5):H1623-32.
47. Rutledge RG, Cote C. Mathematics of quantitative kinetic PCR and the application of standard curves. Nucleic Acids Res 2003;31(16):e93.
48. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;25(4):402-8.
49. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF.. Stages of embryonic development of the zebrafish. Dev Dyn 1995;203(3):253-310.
50. Thisse C, Thisse B, Schilling TF, Postlethwait JH. Structure of the zebrafish snail1 gene and its expression in wild-type, spadetail and no tail mutant embryos. Development 1993;119(4):1203-1215.

PUBBLICATIONS (2005-2008)

Hepatol Res. 2005 Sep;33(1):57-60.

Intron 2 [IVS2, T-C +4] HFE gene mutation associated with S65C causes alternative RNA splicing and is responsible for iron overload.

Floreani A, Navaglia F, Basso D, Zambon CF, Basso G, Germano G, Rizzotto ER, Guido M, Plebani M.
Department of Surgical and Gastroenterological Sciences, University of Padova, via Giustiniani 2, 35128 Padova, Italy.

A patient with congenital liver fibrosis revealed a high transferrin saturation index and iron overload on liver biopsy. He did not carry the most frequent HFE mutations: C282Y or H63D. Heterozygosity was detected for S65C. Unknown HFE mutations were also sought using a combined denaturing high performance liquid chromatography (DHPLC)/direct sequence approach and another point mutation, a transition T-C (nt 4910), at the fourth base of the donor splice site of intron 2 [HFE, intervening sequence (IVS) 2, T-C +4] was found. Family screening revealed that a daughter carried both S65C and [IVS2, T-C +4]. CONCLUSION:: The existence in our proband of a partly-altered HFE protein in the region encoded by exon 2 might be responsible for the histologically-demonstrated iron overload.

Leukemia. 2005 Sep;19(9):1687-9.

Comparative sequence analysis of incomplete DJH and TCR gene rearrangements in children with relapses of T-ALL.

Germano G, del Giudice L, Lo Nigro L, Polato K, Giarin E, Paganin M, Basso G.
Laboratory of Pediatric Onco-Hematology, Department of Pediatrics, University Hospital of Padua, Padua, Italy.

Clin Cancer Res. 2005 Nov 1;11(21):7720-7.

Immunogenotype changes prevail in relapses of young children with TEL-AML1-positive acute lymphoblastic leukemia and derive mainly from clonal selection.

Panzer-Grümayer ER, Cazzaniga G, van der Velden VH, del Giudice L, Peham M, Mann G, Eckert C, Schrauder A, Germano G, Harbott J, Basso G, Biondi A, van Dongen JJ, Gadner H, Haas OA.
Children's Cancer Research Institute, Vienna, Austria.

PURPOSE: Variations of the immunogenotype and TEL deletions in children with TEL-AML1+ acute lymphoblastic leukemia support the hypothesis that relapses derive from a persistent TEL-AML1+ preleukemic/leukemic clone rather than a resistant leukemia. We aimed at elucidating the relationship between the immunogenotype patterns at diagnosis and relapse as well as their clinical and biological relevance. PATIENTS AND METHODS: Immunoglobulin and T-cell receptor gene rearrangements were analyzed in 41 children with a TEL-AML1+ acute lymphoblastic leukemia and an early (up to 30 months after diagnosis; n = 12) or late (at 30 months or later; n = 29) disease recurrence by a standardized PCR approach. RESULTS: In 68% of the patients (group I), we identified differences in the immunogenotype patterns, whereas no changes were observed in the remaining 32% (group II). The divergence resulted more often from clonal selection than clonal evolution and consisted predominantly of losses (0-6, median 5) and/or gains (0-4, median 1) of rearrangements. The frequency and number of clonal immunoglobulin/T-cell receptor rearrangements in group I was higher at diagnosis (2-13, median 5) than at relapse (2-7, median 4), whereas it was the lowest in group II (1-5, median 3). Although group I children were younger at diagnosis, there was no correlation between particular immunogenotype patterns and remission duration. CONCLUSION: These findings imply that the clonal heterogeneity in younger children most likely reflects an ongoing high recombinatorial activity in the preleukemic/leukemic cells, whereas the more uniform repertoire observed in older children mirrors end-stage rearrangement patterns of selected cell clones that evolved during the prolonged latency period.

Haematologica. 2006 May;91(5 Suppl):ECR09.

Two consecutive immunophenotypic switches in a child with MLL-rearranged acute lymphoblastic leukemia.

Germano G, Pigazzi M, del Giudice L, Campo Dell'Orto M, Spinelli M, Zangrando A, Paolucci P, Ladogana S, Basso G.
Laboratory of Pediatric Onco-Hematology, Department of Pediatrics, University Hospital of Padua, Padua, Italy.

An 18-month-old girl was diagnosed with pre-pre-B ALL/t(4;11) leukemia, which during the treatment and after matched bone marrow transplantation (BMT), underwent two consecutive switches from lymphoid to myeloid lineage and vice versa. The high expression of HOXA9 and FLT3 genes remaining genotypically stable in a leukemia throughout phenotypic switches, suggests that this leukemia may have originated as a common B/myeloid progenitors.

Haematologica. 2007 Oct;92(10):1435-7.

cAMP response element binding protein (CREB) overexpression CREB has been described as critical for leukemia progression.

Pigazzi M, Ricotti E, Germano G, Faggian D, Aricò M, Basso G.
Laboratory of Pediatric Onco-Hematology, Department of Pediatrics, University Hospital of Padua, Padua, Italy.

CREB has been described as critical for leukemia progress. We investigated CREB expression in ALL and AML pediatric patients. CREB protein was significantly high ($p<0.001$) at diagnosis but not during remission. This study underlines the role of CREB in leukemia and suggests new insights into the transformation process.