UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI PEDIATRIA

SCUOLA DI DOTTORATO DI RICERCA IN MEDICINA DELLO SVILUPPO E SCIENZE DELLA PROGRAMMAZIONE

INDIRIZZO EMATOONCOLOGIA ED IMMUNOLOGIA

XXII CICLO

"Microarray Analysis: a Leading Tool in the Classification and Biological Characterization of Pediatric Onco-Hematological Diseases"

Direttore della Scuola: Ch.mo Prof. Giuseppe Basso

Supervisore: Dott.ssa Geertruy te Kronnie

Dottorando: Luca Trentin

2007-2009

Ἢ τὰν ἢ επὶ τᾶς.

Contents

pag.

Summary **11**

1. Introduction **17**

1.1 Gene Expression Profile Analysis: a 10–year-overview in the diagnosis of Acute Leukemia 18

1.1.1 "Class Prediction" through Gene Expression Profile Analysis 19

1.1.2 "Class discovery" through Gene Expression Profile Analysis 20

1.1.3 Gene Expression Profile Analysis for the prediction of response to therapy 21

1.2 References 23

2. Scope of this thesis **29**

2.1 References 29

3. Gene Expression Profile Analysis in leukemia classification: the MILE study **31**

3.1 Introduction 31

3.2 References 32

3.3 New data on robustness of gene expression signatures in leukemia: comparison of three distinct total RNA preparation procedures 33

4. Gene Expression Profile Analysis as a prognostic tool in Juvenile Myelomonocytic Leukemia (JMML) **35**

4.1 Introduction 35

4.2 References 36

4.3 Gene expression based classification is an independent predictor of clinical outcome in juvenile myelomonocytic leukemia 37

5. "Class Discovery" through Gene Expression Profile Analysis: the identification of a new leukemia subgroup among *MLL/AF4*-rearranged patients **39**

5.1 Introduction 39

5.1.2 Structure and function of normal *MLL* 40

5.1.3 MLL fusion proteins: structure and mechanisms of action 41

5.1.4 MLL fusion proteins and *HOXA* gene expression deregulation 43

5.2 Gene Expression Profile Analysis of B-cell precursors (BCPs) ALL pediatric patients with *MLL/AF4* 43

5.2.1 Two independent gene signatures in pediatric t(4;11) acute lymphoblastic leukemia patients 45

5.3 Molecular analysis of “HOXA high” and “HOXA low” *MLL/AF4* patients 47
5.3.1 Materials and methods 52
5.4 References 55

6. Identification of new a putative molecular target gene in *MLL/AF4*-rearranged patients through Gene Expression Profile analysis: the suppressor of cytokine signaling 2 (*SOCS-2*) 63
6.1 Introduction 63
6.2 RS4;11: *in vitro* model for silencing experiments 64
6.3 Efficiency of SOCS-2 silencing 64
6.4 Apoptosis induction in RS4;11 cell line through SOCS-2 silencing 65
6.5 Materials and methods 68
6.6 References 70

7. Discussion 73
7.1 References 75

Acknowledgments 77

Appendix 1 I
Appendix 2 III
Appendix 3 V

Summary

Gene Expression Profile (GEP) analysis through microarrays represents a powerful tool for the classification, the prediction and the identification of several leukemia subclasses. In this thesis, we have reported the results we have obtained applying the microarray technology to the study of pediatric onco-hematological diseases.

A huge amount of reports have highlighted the robustness of GEP in the classification of leukemia in both children and adults and these results support the application of microarrays in future routine diagnostic settings.

Since the quality of RNA used for the experiments is one of the critical factors when performing microarrays analysis and seeing that all laboratories commonly use their own distinctive RNA isolation protocol, we have questioned the influence of the three most frequently used extraction protocols in the gene expression profile analysis of pediatric leukemia. Our data have showed that different sample preparation procedures do not impair samples classification and that the underlying biological characteristics of the pediatric acute leukemia classes largely exceed the variations between different RNA preparation protocols.

We have then applied GEP analysis to the study of *MLL/AF4*-rearranged B-cell precursor acute lymphoblastic leukemia (BCP-ALL). Among the *MLL/AF4* BCPs leukemia samples, we have identified the presence of two subgroups of patients characterized by a distinctive gene expression signature in which the down-regulation of the *HOXA* genes is a particularly outstanding factor. *HOXA* genes deregulation, indeed, is commonly believed to be a key mechanism of MLL-fusion gene mediated leukemogenesis. Apart from the differential *HOXA* genes expression level in these subgroups of patients, no transcriptional deregulation of other known MLL-related genes (i.e. *MENIN*, *HOXC8* and *MEIS1*) could be identified.

We have also performed a microRNA expression profile analysis of the patients characterized by the up- or down-regulation of *HOXA* genes and we have demonstrated that they are characterized also by a distinct microRNA signature. Interestingly, patients displaying a low expression value of *HOXA* genes do not express the miR-196b which is located within the *HOXA* cluster and which is involved in leukemogenesis.

Furthermore, we have used microarray analysis to study Juvenile Myelomonocytic Leukemia (JMML). Remarkably, we could distinguish two distinct subgroups among the analyzed patients and this subdivision resulted to have a high prognostic value in the identification of subgroups of patients with distinct clinical outcome. The same result is not reproducible if the usual clinical factors (fetal Hb, age at diagnosis and platelet count) are applied.

Finally, we have focused on the role of the suppressor of cytokine signaling 2 (*SOCS-2*). This gene is reported to be up-regulated in stem cells and we have identified *SOCS-2* as one of the most up-regulated genes in *MLL/AF4* patients irrespective of *HOXA* gene expression level, when comparing t(4;11) samples with normal bone marrow controls. Transient silencing of SOCS-2 in the lymphoid

cell line RS4;11 showed that SOCS-2 depletion induces apoptosis in silenced cells and that this process is characterized by the concurrent increased expression of TP53 and BAX. Thus, *SOCS-2* up-regulation seems to be a mechanism in RS4;11 cells to impair apoptosis activation. We have analyzed SOCS-2 expression levels in patients belonging to several different ALL subclasses and have found that *SOCS-2* is up-regulated in all but T-lineage leukemia. This finding suggests that *SOCS-2* up-regulation could be a common mechanism in ALLs to prevent induction of apoptosis.

Sommario

L'analisi del profilo d'espressione genica mediante microarray rappresenta uno strumento utile per la classificazione delle leucemie in ambito diagnostico, l'identificazione di nuove sottoclassi di malattia e l'associazione di profili d'espressione genica con la prognosi. I molteplici lavori pubblicati nell'ambito delle malattie onco-ematologiche sia nell'adulto che nel bambino hanno evidenziato la robustezza della tecnologia microarray ed auspicano, quindi, l' utilizzo dei microarrays in affiancamento alle metodiche "*gold standard*" per la diagnosi di leucemia. Considerando che la qualità dell'RNA di partenza è un fattore determinante per la buona riuscita di un esperimento di studio dell'espressione genica, abbiamo valutato se diverse metodiche di isolamento dell'RNA avessero una qualche influenza sulla variazione del profilo d'espressione genica. I risultati ottenuti nel nostro studio, analizzando diversi sottotipi di leucemie pediatriche, hanno evidenziato che le metodiche impiegate per l'estrazione dell'RNA non vanno ad influire sul profilo d'espressione genico e che quest'ultimo rimane, comunque, ben identificabile a prescindere dalla metodologia usata per l'isolamento dell'RNA.

Applicando, poi, l'analisi microarrays alle leucemie a cellule precursori B e con traslocazione *MLL/AF4*, abbiamo individuato, all'interno di questo sottotipo di leucemia ritenuto fino ad ora omogeneo, due sottogruppi di pazienti caratterizzati da un differente profilo d'espressione genica in cui spiccava la diversa espressione dei geni *HOXA*. Questo risultato è alquanto sorprendente poiché la maggiore espressione dei geni *HOXA* è una caratteristica distintiva delle leucemie con riarrangiamento del gene *MLL*. Non abbiamo identificato nessuna altra variazione d'espressione di geni (per es. *MENIN*, *HOXC8* e *MEIS1*) comunemente associati con le leucemie con riarrangiamento del gene *MLL*.

Anche l'analisi del profilo dell'espressione dei microRNA ha dimostrato che questi pazienti possono essere suddivisi in due sottogruppi ben distinti ed, inoltre, ha evidenziato che i pazienti con bassa espressione dei geni *HOXA* non esprimono il microRNA mir-196b, che è localizzato nel medesimo *cluster* dei geni *HOXA* e che è coinvolto nei processi di leucemogenesi.

Lo studio del profilo d'espressione genica di pazienti affetti da leucemia mielomonocitica giovanile (JMML) ci ha, poi, consentito di dividere i campioni analizzati in due sottogruppi. Questa suddivisione è associata, in modo altamente significativo, con la prognosi di malattia. Il medesimo risultato prognostico non è conseguibile prendendo in considerazione i fattori prognostici clinici standard (emoglobina fetale, età alla diagnosi e conta piastrinica).

Infine, abbiamo studiato il ruolo del gene *SOCS-2* nelle leucemie con riarrangiamento *MLL/AF4*. Questo gene, *up*-regolato nelle cellule staminali, è uno dei geni maggiormente espressi nei pazienti con *MLL/AF4* rispetto ai controlli normali. Il silenziamento di SOCS-2 nelle cellule RS4;11 determina un aumento dell'apoptosi rispetto alle cellule silenziate con un siRNA di controllo ed una simultanea maggiore espressione di TP53 e BAX. L'*over*-espressione di SOCS-2 nelle cellule RS4;11 sembra

essere, quindi, un meccanismo in grado di aumentare la sensibilità di queste cellule all'apoptosi. L'analisi dell'espressione di *SOCS-2* in più pazienti affetti da leucemia linfoblastica acuta (LLA) ha evidenziato che *SOCS-2* è *over*-espresso in tutte le LAL tranne le LAL a cellule T. Questo dato suggerisce che l'azione anti-apoptotica di *SOCS-2* potrebbe essere comune in più sottotipi di leucemia.

1. Introduction

Human acute leukemias arise from blood cell progenitors developing in the lymphoid or myeloid pathway or from primitive stem cells with multilineage potential. The analysis of clonal chromosomal abnormalities in acute myeloid (AML) and acute lymphoblastic leukemia (ALL) of both B- and T-lineage has identified a massive quantity of different genetic aberrations (**Fig. 1**). In AML, chromosomal alterations are identified in 55% of patients at diagnosis and molecular mutations can be detected in about 85% of cases with normal karyotype (1). In B-lineage ALL the frequency of chromosomal abnormalities is about 90% (2, 3) and 50-70% in T-lineage ALL (4, 5). Moreover, the simultaneous presence of mutations of the genes *FLT3* and *NPM1* in normal karyotype AML expands the range of genetic aberrations in AML (6).

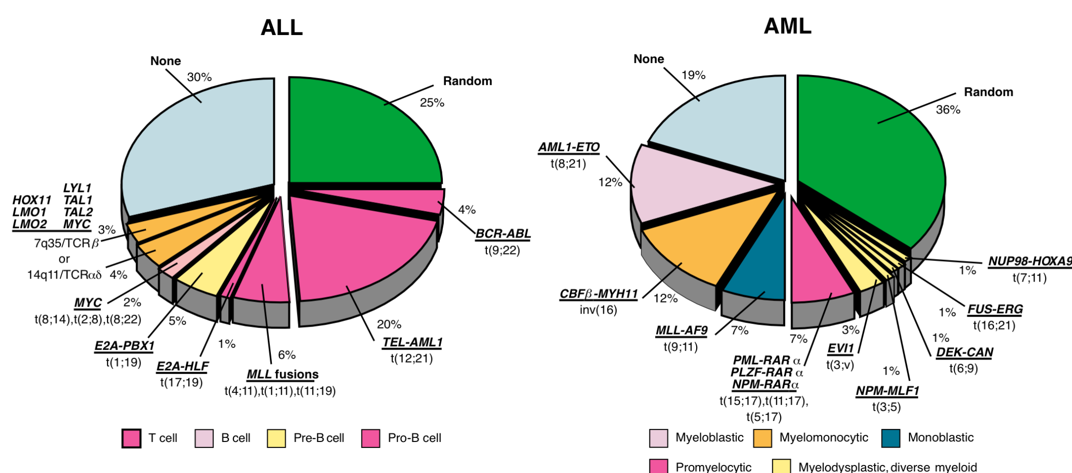


Figure 1. Distribution of translocation-generated oncogenes among the acute leukemia of children and young adults. “Random” refers to sporadic translocations; “None” refers to leukemia that lacks identifiable gene abnormalities (from Look 1997) (7).

Several approaches, such as morphology, cytogenetics, immunophenotyping, cytochemistry, fluorescence in situ hybridization (FISH) and molecular assays, are required for a proper diagnosis of all leukemias subtypes. Moreover, The precise classification of the specific leukemia subtypes, a more accurate patients risk group stratification and the early detection of patient’s therapy response are the main objectives of the current studies on leukemias.

Since its introduction in leukemia subtypes distinction in 1999 (8), microarray technology has proved to be a powerful tool for the study of leukemia as it can quantify the expression of ten of thousands genes in a single experimental approach. Furthermore, the analysis of the data obtained through microarray technology allows the identification and clustering of cases according to similar gene expression signatures.

The pivotal work reported by Golub et al (1999) identifies a limited number of genes (50) whose gene expression level is able to separate 27 patients with ALL from 11 patients with AML (**Fig. 2**);

moreover ALL are further separated in T-lineage and B-lineage. This study suggests three potential applications of GEP: "class discovery", "class prediction" and "class comparison". "Class discovery" focuses on the identification of new subgroups of patients and involves no pre-specified classes in the data collection whereas "class prediction" refers to the possibility to use gene expression data to predict already defined subgroups. "Class comparison" refers to the identification of genes differentially expressed among pre-specified groups and it is usually performed after a "class prediction " or a "class discovery" analysis.

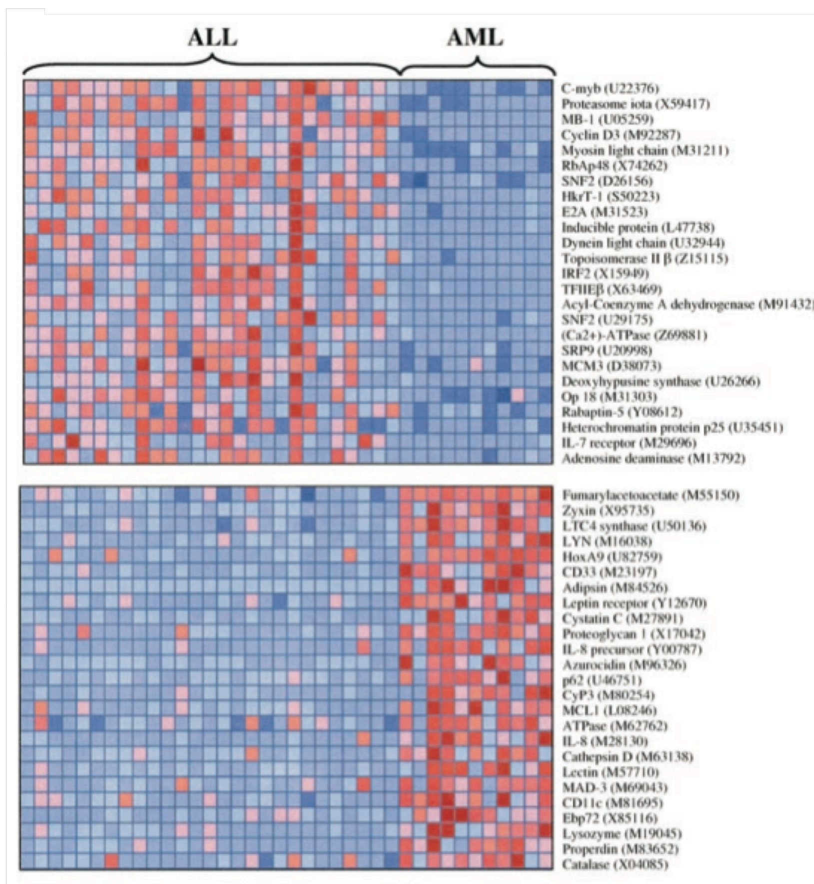


Figure 2. The 50 genes most highly correlated with the ALL-AML class distinction are shown. The top panel shows genes highly expressed in ALL, the bottom panel shows genes more highly expressed in AML (from Golub et al, 1999) (8).

1.1 Gene Expression Profile Analysis: a 10-year-overview in the diagnosis of Acute Leukemia

The data provided by Golub and his colleagues were subsequently confirmed by several works and represent the starting point for a widespread use of GEP in the study of leukemia. In the last ten years, more than two thousand reports applying microarray technology in hematological malignances have been published; this introductory section will focus on a part of these works showing that microarray technology represents a new tool for the classification of leukemia (class prediction), for the

identification of the de-regulated genes characterizing different leukemia sub-entities (class comparison) and for the detection of new biological and clinically relevant subtypes in leukemia (class discovery) in both children and adults.

1.1.1 "Class Prediction" through Gene Expression Profile Analysis

Since Golub et al (1999) first demonstrated that GEP can separate ALL from AML, great efforts have been made to prove that the GEP study can predict genetically well-defined leukemia sub-groups in both ALLs and AMLs. The studies by Schoch et al (2002)(9) and Ross et al (2003 and 2004)(10, 11) have demonstrated that the favorable AML subgroups harboring the t(8;21), t(15;17) and inv(16) genetic abnormalities can be divided according to their distinct gene expression profile both in children and adults and can also be predicted with up to 100% accuracy (12).

High prediction accuracy has also been achieved in the identification of the well-defined subgroup of leukemias harboring the translocation of the *MLL*/11q23 gene. Armstrong et al (2002) (13) have demonstrated that MLL-rearranged ALLs have a distinct gene expression signature and are characterized by the expression of early hematopoietic progenitor cell multilineage markers and specific *HOXA* genes.

The specific gene signature that distinguishes MLL-rearranged acute lymphoblastic leukemias with respect to all other ALLs and AMLs has also been widely investigated (14-16) and genes such as *HOXA9*, *MEIS1* and *PBX3* have resulted to be the genes most differentially regulated in both lymphoblastic and myeloid MLL-rearranged leukemias with respect to other leukemia subgroups.

Schichman et al (1994) (17) have proved that, in addition to chromosomal translocation, the *MLL* gene can also be altered by an internal partial tandem duplication (PTD). MLL-PTDs are typically found in AML cases that have either normal cytogenetics or a trisomy of chromosome 11; GEP analysis has revealed that AML with MLL-PTD do not cluster with MLL chimeric fusion gene cases (18). Thus, the pathogenic mechanisms of partial duplications and of chimeric gene fusions of the *MLL* gene seem to differ significantly.

The T-precursor ALL, the mature B-ALL with *IgH-c-MYC*/t(8;14), the B-ALL with *TEL-AML1*/t(12;21), *E2A-PBX1*/t(1;19) and the precursor B-ALL with *BCR-ABL*/t(9;22) ALL present also a distinct gene signature and can be properly classified by GEP (14, 19-23). Furthermore, ALL with hyperdyploid karyotype (> 50) (21) and AML with complex aberrant karyotype (defined by the occurance of at least three clonal chromosomal abnormalities) can also be properly separated by GEP from other genetically defined subtypes. In particular, complex AMLs showed a significantly higher expression of *RAD21* and a differential expression of genes localized on chromosomes such as 5q, 7 and 17p, which are frequently involved in karyotype aberrations identified in this genetic AML subgroup (9, 24, 25).

Nevertheless, some others AML subtypes presenting chromosomal abnormalities such as trisomy 8 and defects involving the 3q, have showed to be not well predictable by GEP in representative cohorts of AMLs (26-28). Also patients presenting a c-ALL/pre-B-ALL immunophenotype and without the t(9;22) are not precisely identified by GEP (21, 29).

1.1.2 "Class Discovery" through Gene Expression Profile Analysis

As reported, GEP analysis can identify several leukemia subgroups and this classification reflects also the division of the different leukemia sub-groups in prognostically unfavorable groups (i.e. T-lineage ALL, the precursor B-lineage subtypes with MLL-rearrangements, the *BCR-ABL* gene fusion) and prognostically favorable precursor B-subtypes (i.e. *TEL-AML1*-positive ALL, hyperdiploid ALL and E2A-rearranged ALL). Nevertheless, about 25% of cases are currently genetically unclassified ALL (B-other) at diagnosis.

The application of GEP to a cohort of 190 children has identified a new high-risk subtype of ALL that comprises 15–20% of all precursor B-ALL cases whose gene expression pattern is very similar to that of *BCR-ABL*-positive ALL (30). The same results have been described by Mullighan et al (2009) (31) in an independent study. The authors analyzed the GEP of 221 pediatric patients with high risk B-cell progenitor ALL. Making use of a new prognostic predictor based on single nucleotide polymorphism array analysis, the authors identified a group of patients characterized by the mutation of *IKZF1* and a very poor outcome. The GEP analysis revealed that these patients were characterized by an increased expression of hematopoietic stem cell genes and reduced expression of B-cell lineage genes. Remarkably, this signature is very similar to the gene signature of *BCR-ABL* positive patients (another high-risk ALL subtype with a high frequency of *IKZF1* deletion).

Also Bullinger et al (2004) (27) have used GEP analysis to detect new leukemia subgroups and better classify these diseases. They identified a specific gene signature, which was able to divide normal karyotype AML into two clusters with different prognosis in a cohort of 166 cases of AML. One cluster presented the high expression of several transcriptional regulators (i.e. *GATA2*, *DNMT3A* and *DNMT3B*), whereas the second cluster was characterized by granulocytic or monocytic differentiation and immune response. The prognostic significance of this gene signature was further validated in an independent study (32).

Wouters et al (2007) (33) have identified a new biologically and clinically relevant subgroup in AMLs characterized by the *CEBPA* silencing and an immature myeloid/T-lymphoid phenotype.

Furthermore, within a group of 166 AML, the patients with t(8;21) and inv(16) could be split into subgroups based merely on their gene expression signature, which suggests that also well-defined subgroups may represent heterogeneous entities in term of gene expression (27). For example, a subgroup of patients with shorter overall survival and characterized by an elevated white blood cells

count and *FLT3*-internal tandem duplication (ITD) was identified by GEP among the favorable group of AML with inv(16)/*CBFB-MYH11*.

A high heterogeneity has also been reported when studying the gene expression profile of patients with down-syndrome (DS). DS-ALLs are mostly of B-cell precursor (BCP) origin and similar in the age of diagnosis and immunophenotype to the two most common genetic subtypes of childhood ALL, i.e. high hyperdiploid (HD) and *TEL-AML1* ALLs. An unsupervised gene expression analysis has revealed that DS-ALLs are markedly less homogenous than the other ALL genetic subtypes suggesting that DS leukemia should not be considered as a unique molecular entity (34). Moreover, the comparison of the gene signature of the analyzed patients revealed that 62% of the DS-ALL samples were characterized by the high expression of type I cytokine receptor *CRLF2*.

1.1.3 Gene Expression Profile Analysis for the prediction of response to therapy

The current cure rate of 80% in childhood acute lymphoblastic leukemia (ALL) attested to the effectiveness of risk-directed therapy developed through clinical trials in the last ten years (35, 36). Children whose ALL cells exhibit *in vitro* resistance to anti- leukemic agents have a substantially worse prognosis than children whose ALL cells are drug-sensitive. However, little is known about the genetic basis of resistance to chemotherapy. Multidrug-resistant genes and genes involved in cell-cycle progression, DNA repair, drug metabolism and apoptosis have been associated with the prognosis of ALL, but their role in determining the sensitivity of ALL cells to individual antileukemic agents is not known.

Gene expression profile analysis has been used for the identification of genes that are differentially expressed in primary ALL cells that exhibit either resistance or sensitivity to prednisone, vincristine, aspareginase or daunorubicin (37). In the study reported by Holleman et al (2004) genes resulted to be differentially expressed among sensitive and resistant patients and all but 3 genes were not previously associated with the resistance to the four drugs.

Prediction of patients' response to therapy has also been achieved in another group of pediatric ALL samples; as reported by Willenbrock et al (2004) (23), GEP analysis was able to predict 5-year event free survival or relapse with a classification accuracy of 78%, regardless of immunophenotype in 45 pediatric ALL patients. Bhojwani et al (2008) (38) used GEP to identify children with acute lymphoblastic leukemia at initial diagnosis and who were at risk due to inferior response to therapy. They detected a 24-probe set signature highly predictive of day 7 marrow status was identified in 99 children with high risk ALL at diagnosis and also identified a 47-probe set signature predictive of long-term outcome not only in the 99 analyzed specimens but also in three large independent data sets of patients with childhood ALL treated on different protocols.

Hoffmann et al (2008) (39) and Chiaretti et al (2004) (40) have identified the genes that can predict the response to induction treatment or to specific compound. In particular, Hofmann et al (2002) (41) found 95 genes that were the most predictive of sensitivity to imatinib, a tyrosine kinase inhibitor used for the treatment of t(9;22) ALL. Resistant leukemic cells expressed high levels of Bruton's tyrosine kinase and two ATP synthetases (*ATP5A1* and *ATP5C1*) and showed significantly reduced expression of the pro-apoptotic gene *BAK1* and the cell-cycle control gene *p15 INK4b*.

Furthermore, Cheok et al (2003) (42) recognized through microarray analysis the genes involved in resistance to methotrexate and mercaptopurine. A recent gene expression study on MLL-rearranged infant patients has linked high-level *MCL-1* to resistance to prednisolone (43). This finding represents an important step towards the comprehension of the mechanisms of resistance to prednisone in MLL-rearranged infant patients. Prednisone is, indeed, the spearhead drug of ALL treatment regimens (44) and MLL-infant patients show a poor response to prednisone *in vivo* (defined as the presence of ≥ 1000 leukemic blast/μl after a 7-day window of prednisone mono-therapy (45, 46) in contrast to children older than 1 year of age (non-infants) diagnosed with ALL and who show a poor *in vivo* response to prednisone in less than 10% of the cases (47).

1.2 References

1. Marcucci G, Mrozek K, Bloomfield CD. Molecular heterogeneity and prognostic biomarkers in adults with acute myeloid leukemia and normal cytogenetics. Curr Opin Hematol 2005; 12: 68-75.

2. Fenaux P, Lai JL, Morel P, Nelken B, Taboureau O, Deminatti M, et al. Cytogenetics and their prognostic value in childhood and adult acute lymphoblastic leukemia (ALL) excluding L3. Hematol Oncol 1989; 7: 307-17.

3. Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. N Engl J Med 2004; 350: 1535-48.

4. Raimondi SC, Behm FG, Roberson PK, Pui CH, Rivera GK, Murphy SB, et al. Cytogenetics of childhood T-cell leukemia. Blood 1988; 72: 1560-6.

5. Heerema NA, Sather HN, Sensel MG, Kraft P, Nachman JB, Steinherz PG, et al. Frequency and clinical significance of cytogenetic abnormalities in pediatric T-lineage acute lymphoblastic leukemia: a report from the Children's Cancer Group. J Clin Oncol 1998; 16: 1270-8.

6. Falini B, Nicoletti I, Martelli MF, Mecucci C. Acute myeloid leukemia carrying cytoplasmic/mutated nucleophosmin (NPMc+ AML): biologic and clinical features. Blood 2007; 109: 874-85.

7. Look AT. Oncogenic transcription factors in the human acute leukemias. Science 1997; 278: 1059-64.

8. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999; 286: 531-7.

9. Schoch C, Kohlmann A, Schnittger S, Brors B, Dugas M, Mergenthaler S, et al. Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. Proc Natl Acad Sci U S A 2002; 99: 10008-13.

10. Ross ME, Zhou X, Song G, Shurtleff SA, Girtman K, Williams WK, et al. Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. Blood 2003; 102: 2951-9.

11. Ross ME, Mahfouz R, Onciu M, Liu HC, Zhou X, Song G, et al. Gene expression profiling of pediatric acute myelogenous leukemia. Blood 2004; 104: 3679-87.

12. Verhaak RG, Wouters BJ, Erpelinck CA, Abbas S, Beverloo HB, Lugthart S, et al. Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. Haematologica 2009; 94: 131-4.

13. Armstrong SA, Staunton JE, Silverman LB, Pieters R, den Boer ML, Minden MD, et al. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat Genet 2002; 30: 41-7.

14. Ferrando AA, Armstrong SA, Neuberg DS, Sallan SE, Silverman LB, Korsmeyer SJ, et al. Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: dominance of HOX dysregulation. Blood 2003; 102: 262-8.

15. Kohlmann A, Schoch C, Dugas M, Schnittger S, Hiddemann W, Kern W, et al. New insights into MLL gene rearranged acute leukemias using gene expression profiling: shared pathways, lineage commitment, and partner genes. Leukemia 2005; 19: 953-64.

16. Zangrando A, Dell'orto MC, Te Kronnie G, Basso G. MLL rearrangements in pediatric acute lymphoblastic and myeloblastic leukemias: MLL specific and lineage specific signatures. BMC Med Genomics 2009; 2: 36.

17. Schichman SA, Caligiuri MA, Strout MP, Carter SL, Gu Y, Canaani E, et al. ALL-1 tandem duplication in acute myeloid leukemia with a normal karyotype involves homologous recombination between Alu elements. Cancer Res 1994; 54: 4277-80.

18. Haferlach T, Schnittger S, Kern W, Hiddemann W, Schoch C. Genetic classification of acute myeloid leukemia (AML). Ann Hematol 2004; 83 Suppl 1: S97-100.

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

20. Moos PJ, Raetz EA, Carlson MA, Szabo A, Smith FE, Willman C, et al. Identification of gene expression profiles that segregate patients with childhood leukemia. Clin Cancer Res 2002; 8: 3118-30.

21. Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell 2002; 1: 133-43.

22. Kohlmann A, Schoch C, Schnittger S, Dugas M, Hiddemann W, Kern W, et al. Molecular characterization of acute leukemias by use of microarray technology. Genes Chromosomes Cancer 2003; 37: 396-405.

23. Willenbrock H, Juncker AS, Schmiegelow K, Knudsen S, Ryder LP. Prediction of immunophenotype, treatment response, and relapse in childhood acute lymphoblastic leukemia using DNA microarrays. Leukemia 2004; 18: 1270-7.

24. Lindvall C, Furge K, Bjorkholm M, Guo X, Haab B, Blennow E, et al. Combined genetic and transcriptional profiling of acute myeloid leukemia with normal and complex karyotypes. Haematologica 2004; 89: 1072-81.

25. Rucker FG, Bullinger L, Schwaenen C, Lipka DB, Wessendorf S, Frohling S, et al. Disclosure of candidate genes in acute myeloid leukemia with complex karyotypes using microarray-based molecular characterization. J Clin Oncol 2006; 24: 3887-94.

26. Virtaneva K, Wright FA, Tanner SM, Yuan B, Lemon WJ, Caligiuri MA, et al. Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. Proc Natl Acad Sci U S A 2001; 98: 1124-9.

27. Bullinger L, Dohner K, Bair E, Frohling S, Schlenk RF, Tibshirani R, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. N Engl J Med 2004; 350: 1605-16.

28. Haferlach T, Kohlmann A, Schnittger S, Dugas M, Hiddemann W, Kern W, et al. AML M3 and AML M3 variant each have a distinct gene expression signature but also share patterns different from other genetically defined AML subtypes. Genes Chromosomes Cancer 2005; 43: 113-27.

29. Chiaretti S, Li X, Gentleman R, Vitale A, Wang KS, Mandelli F, et al. Gene expression profiles of B-lineage adult acute lymphocytic leukemia reveal genetic patterns that identify lineage derivation and distinct mechanisms of transformation. Clin Cancer Res 2005; 11: 7209-19.

30. Den Boer ML, van Slegtenhorst M, De Menezes RX, Cheok MH, Buijs-Gladdines JG, Peters ST, et al. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. Lancet Oncol 2009; 10: 125-34.

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

32. Radmacher MD, Marcucci G, Ruppert AS, Mrozek K, Whitman SP, Vardiman JW, et al. Independent confirmation of a prognostic gene-expression signature in adult acute myeloid leukemia with a normal karyotype: a Cancer and Leukemia Group B study. Blood 2006; 108: 1677-83.

33. Wouters BJ, Jorda MA, Keeshan K, Louwers I, Erpelinck-Verschueren CA, Tielemans D, et al. Distinct gene expression profiles of acute myeloid/T-lymphoid leukemia with silenced CEBPA and mutations in NOTCH1. Blood 2007; 110: 3706-14.

34. Hertzberg L, Vendramini E, Ganmore I, Cazzaniga G, Schmitz M, Chalker J, et al. Down syndrome acute lymphoblastic leukemia: a highly heterogeneous disease in which aberrant expression of CRLF2 is associated with mutated JAK2: a report from the iBFM Study Group. Blood 2009;.

35. Pui CH, Evans WE. Acute lymphoblastic leukemia. N Engl J Med 1998; 339: 605-15.

36. Pui CH, Campana D, Evans WE. Childhood acute lymphoblastic leukaemia--current status and future perspectives. Lancet Oncol 2001; 2: 597-607.

37. Holleman A, Cheok MH, den Boer ML, Yang W, Veerman AJ, Kazemier KM, et al. Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment. N Engl J Med 2004; 351: 533-42.

38. Bhojwani D, Kang H, Menezes RX, Yang W, Sather H, Moskowitz NP, et al. Gene expression signatures predictive of early response and outcome in high-risk childhood acute lymphoblastic leukemia: A Children's Oncology Group Study [corrected]. J Clin Oncol 2008; 26: 4376-84.

39. Hoffmann K, Firth MJ, Beesley AH, Freitas JR, Ford J, Senanayake S, et al. Prediction of relapse in paediatric pre-B acute lymphoblastic leukaemia using a three-gene risk index. Br J Haematol 2008; 140: 656-64.

40. Chiaretti S, Li X, Gentleman R, Vitale A, Vignetti M, Mandelli F, et al. Gene expression profile of adult T-cell acute lymphocytic leukemia identifies distinct subsets of patients with different response to therapy and survival. Blood 2004; 103: 2771-8.

41. Hofmann WK, de Vos S, Elashoff D, Gschaidmeier H, Hoelzer D, Koeffler HP, et al. Relation between resistance of Philadelphia-chromosome-positive acute lymphoblastic leukaemia to the tyrosine kinase inhibitor STI571 and gene-expression profiles: a gene-expression study. Lancet 2002; 359: 481-6.

42. Cheok MH, Yang W, Pui CH, Downing JR, Cheng C, Naeve CW, et al. Treatment-specific changes in gene expression discriminate in vivo drug response in human leukemia cells. Nat Genet 2003; 34: 85-90.

43. Stam RW, Den Boer ML, Schneider P, de Boer J, Hagelstein J, Valsecchi MG, et al. Association of high-level MCL-1 expression with in vitro and in vivo prednisone resistance in MLL-rearranged infant acute lymphoblastic leukemia. Blood 2009;.

44. Pieters R, den Boer ML, Durian M, Janka G, Schmiegelow K, Kaspers GJ, 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.

45. Pieters R, Schrappe M, De Lorenzo P, Hann I, De Rossi G, Felice M, et al. A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. Lancet 2007; 370: 240-50.

46. Dordelmann M, Reiter A, Borkhardt A, Ludwig WD, Gotz N, Viehmann S, et al. Prednisone response is the strongest predictor of treatment outcome in infant acute lymphoblastic leukemia. Blood 1999; 94: 1209-17.

47. Riehm H, Reiter A, Schrappe M, Berthold F, Dopfer R, Gerein V, et al. Corticosteroid-dependent reduction of leukocyte count in blood as a prognostic factor in acute lymphoblastic leukemia in childhood (therapy study ALL-BFM 83). Klin Padiatr 1987; 199: 151-60.

2. Scope of the thesis

Since its introduction, microarray technology has resulted to be a powerful tool for the classification of the diverse leukemia subtypes, for the identification of new subgroups of patients characterized by presenting so far unknown mutations, for the recognition of patients with dismal outcome already at diagnosis and for the identification of genes associated with therapy resistance.

This thesis aims to use the above-mentioned distinctive properties of microarray technology for the study of hematological diseases in pediatric patients.

In the first part of the thesis, we have provided new evidence of the robustness of the gene expression profile technology. The works previously published by Kohlmann et al (2005; 2008)(1, 2) and Mitchell et al (2004)(3) have demonstrated that this technique is characterized by both a high inter-platform comparability and reproducible capacity. These results support the application of microarrays in future routine diagnostic settings; we have analyzed the influence of the three most frequently used RNA extraction protocols in the gene expression profile analysis of pediatric leukemia and thus focused on the potential introduction of microarrays in the diagnosis of leukemia.

As described in the introduction, the gene expression profile technology has been frequently to identify some heterogeneity among an already established disease subgroup or to predict the patient's outcome. In the second part of the thesis we have reported the results we obtained using microarray technology for the study of *MLL/AF4* rearranged B-cell precursor acute lymphoblastic leukemia (BCP-ALL) and juvenile myelomonocytic leukemia (JMML) which are usually diagnosed in early childhood and are characterized by a very poor outcome.

Finally, as GEP technology can also be considered a useful tool for the discovery of new potential target genes, we have used the data collected from the GEP analysis of *MLL/AF4* BCPs ALL patients to identify a particularly up-regulated gene that has not been broadly described so far in relation to leukemogenesis, *SOCS-2*, that we analyzed for its role in *MLL/AF4* BCPs ALL.

2.1 References

1. Kohlmann A, Schoch C, Dugas M, Rauhut S, Weninger F, Schnittger S, et al. Pattern robustness of diagnostic gene expression signatures in leukemia. Genes Chromosomes Cancer 2005; 42: 299-307.

2. Kohlmann A, Kipps TJ, Rassenti LZ, Downing JR, Shurtleff SA, Mills KI, et al. An international standardization programme towards the application of gene expression profiling in routine leukaemia

diagnostics: the Microarray Innovations in LEukemia study prephase. Br J Haematol 2008; 142: 802-7.

3. Mitchell SA, Brown KM, Henry MM, Mintz M, Catchpoole D, LaFleur B, et al. Inter-platform comparability of microarrays in acute lymphoblastic leukemia. BMC Genomics 2004; 5: 71.

3. The Gene Expression Profile Analysis in leukemia classification: the MILE study

3.1 Introduction

The data collected so far and concerning the classification of leukemia in both children and adults through GEP analysis, suggests that this technology is quite robust, replicable and should be introduced into the routinely established gold standard practice for the diagnosis of leukemia. According to these notions, the European Leukemia Network (ELN, WP13) between years 2005 and 2008 performed the Microarray Innovation in LEukemia (MILE) study in order to evaluate the role of microarray technology in the diagnosis and prediction of the various subtypes of leukemia and myelodysplastic syndromes (MDS) in both children and adults using standardized protocols and instruments.

This study includes 11 laboratories across three continents: 7 from the European Leukemia Network, including the Laboratory of Pediatric Onco-Hematology of Padova, 3 from the US and 1 from Singapore. Each centre was trained on an identical microarray (Affymetrix HG-U133 Plus 2.0) and used the same equipment and kits for target preparation (i.e. cDNA synthesis, *in vitro* transcription, labeling and washing procedures); a total of 3252 cases (leukemias and MDS) were analyzed.

The study included a biomarker discovery phase and a classification phase. The biomarker discovery phase was used to generate a whole-genome gene expression profile from recognized categories of clinically relevant leukemias and myelodysplastic syndromes. During this phase an algorithm called DC model was developed; this algorithm allowed the assignment of a diagnostic sample to one of the following 18 classes using only a small subset of genes corresponding to 534 probe sets: class 1: mature B-ALL with t(8;14), class 2: pro-B-ALL with t(11q23)/MLL, class 3: c-ALL/pre-B-ALL with t(9;22), class 4: T-ALL, class 5: ALL with t(12;21), class 6: ALL with t(1;19), class 7: ALL with hyperdiploid karyotype, class 8: c-ALL /pre-B ALL without specific genetic abnormalities, class 9: AML with t(8;21), class 10: AML with t(15;17), class 11: AML with inv(16), class 12: AML with t(11q23)/MLL, class 13: AML with normal karyotype or other abnormalities, class 14: AML with complex aberrant phenotype, class 15: CLL, class 16: CML, class 17: MDS and class 18: non leukemic and healthy bone marrow.

The classification phase of the MILE study was performed using a customized array representing the 1480 probe sets used by the algorithm to classify diagnostic specimens. This phase represented an independent and blinded test set for the algorithm developed during MILE stage one.

A classification scheme aiming to accurately addressing acute leukemia only resulted in 95.5% median sensitivity and 99,5% median specificity for the 14 classes included in the classifier (class1-14). Lower accuracies were observed for the samples diagnosed as class 7-8 in ALL as well as class 12 and class 14 in AML. In addition, lower sensitivities were also observed for class 17.

The data collected during the MILE study on the 3252 patients confirms that microarrays can accurately classify acute and chronic leukemias samples into known diagnostic sub-categories. Moreover, this study has showed that standardized methods in the gene expression profile analysis and their operating procedures may improve current diagnostic techniques, considering that 52 cases were correctly diagnosed by microarray analysis if compared to the initial diagnoses (1).

As stated before, during the MILE study, all the centers collaborating to the project used the same platform and procedures and this standardization was decided in view of the possible future introduction of microarray analysis into the standard diagnostic techniques. Nevertheless, this is not the rule as each laboratory commonly uses different protocol and can vary the various steps of the sample-preparation procedure, such as RNA extraction protocol or the amount of RNA used for the cDNA synthesis.

Since the quality of RNA is one of the major concerns in microarrays experiments, our study aims to address to what extent distinct total RNA template isolation techniques impair the precision and reproducibility of gene expression data from the same sample and whether the underlying characteristic leukemia-specific gene expression signatures are affected by the RNA preparation procedure.

3.2 References

1. Haferlach T, Kohmann A, Wieczorek Lea. The clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: Report on 3334 cases from the international MILE study group. J Clin Oncol 2009; doi:10.1200/JCO.2009.23.4732.

3.3

New data on robustness of gene expression signatures in leukemia: comparison of three distinct total RNA preparation procedures

Marta Campo Dell'Orto°, Andrea Zangrando°, Luca Trentin°, Rui Li*, Wei-min Liu*, Geertruy te Kronnie°, Giuseppe Basso° and Alexander Kohlmann*

° University of Padua, Laboratory of Molecular Diagnostic, Department of Pediatric Oncology, Padua, Italy
* Roche Molecular Systems, Inc, Department of Genomics and Oncology, Pleasanton, CA, USA

BMC Genomic 2007, 8:188

See Appendix 1

4. Gene Expression Profile Analysis as a prognostic tool in Juvenile Myelomonocytic Leukemia (JMML)

4.1 Introduction

Juvenile myelomonocytic leukemia (JMML) is rare childhood hematopoietic stem cell disease characterized by the excessive proliferation of monocytic and granulocytic cells.

JMML is characterized by the hypersensitivity of the myeloid progenitor cells to granulocyte–macrophage colony-stimulating factor (GM-CSF) in culture (1), which is caused by a selective inability of these cells to modulate the RAS-dependent signaling pathways. This hypersensibility is also partially due to mutually exclusive mutations in *NRAS*, *KRAS*, *NF1*, *CBL* and *PTPN11*. These mutations are found in more than 75% of patients with JMML (2, 3). The patients' karyotype is usually normal, whereas monosomoy 7 is present in 25% of cases of JMML and other abnormalities are detected in 10% of cases (4). To date, hematopoietic stem cell transplantation (HSCT) is the only available treatment. The median survival time without HSCT is approximately one year; low platelet count, age more than two years at diagnosis and high fetal hemoglobin (HbF) at diagnosis are the main predictors of short survival (5, 6).

We performed the GEP analysis of 44 patients with JMML and, using the DC model classifier developed during the MILE study, we distinguished two distinct subgroups among the analyzed patients (class prediction). This subdivision resulted to have a high prognostic value in the identification of subgroups of patients with distinct clinical outcome; the same result is not reproducible if the usual clinical factors are applied.

We then compared the expression value of the genes used by the DC model algorithm among the two recognized sub-groups (class comparison) and we identified twenty-seven genes involved in the B-cell lineage differentiation and proliferation.

4.2 References

1. Emanuel PD, Bates LJ, Castleberry RP, Gualtieri RJ, Zuckerman KS. Selective hypersensitivity to granulocyte-macrophage colony-stimulating factor by juvenile chronic myeloid leukemia hematopoietic progenitors. Blood 1991; 77: 925-9.

2. Muramatsu H, Makishima H, Jankowska AM, Cazzolli H, O'Keefe C, Yoshida N, et al. Mutations of E3 ubiquitin ligase Cbl family members but not TET2 mutations are pathogenic in juvenile myelomonocytic leukemia. Blood 2009;.

3. Niemeyer CM, Kratz CP. Paediatric myelodysplastic syndromes and juvenile myelomonocytic leukaemia: molecular classification and treatment options. Br J Haematol 2008; 140: 610-24.

4. Niemeyer CM, Locatelli F. Childhood Leukemias. 2006; New York, Cambridge University Press.

5. Niemeyer CM, Arico M, Basso G, Biondi A, Cantu Rajnoldi A, Creutzig U, et al. Chronic myelomonocytic leukemia in childhood: a retrospective analysis of 110 cases. European Working Group on Myelodysplastic Syndromes in Childhood (EWOG-MDS). Blood 1997; 89: 3534-43.

6. Passmore SJ, Chessells JM, Kempski H, Hann IM, Brownbill PA, Stiller CA. Paediatric myelodysplastic syndromes and juvenile myelomonocytic leukaemia in the UK: a population-based study of incidence and survival. Br J Haematol 2003; 121: 758-67.

4.3

Gene expression profile classification predicts clinical outcome in juvenile myelomonocytic leukaemia

Silvia Bresolin[1]*, Marco Zecca[2]*, Christian Flotho[3], Luca Trentin[1], Andrea Zangrando[1], Laura Sainati[1], Jan Stary[4], Barbara de Moerloose[5], Henrik Hasle[6], Charlotte M. Niemeyer[3], Geertruy te Kronnie[1], Franco Locatelli[2] and Giuseppe Basso[1]

1. Laboratory of Oncohematology, Department of Pediatrics, University of Padova, Italy.
2. Pediatric Hematology and Oncology, Fondazione IRCCS Policlinico San Matteo, University of Pavia, Italy.
3. Division of Pediatric Hematology-Oncology, Depament of Pediatrics and Adolescent Medicine, University of Freiburg, Germany.
4. Department of Pediatric Hematology and Oncology, Charles University, Prague, Czech Republic.
5. Department of Pediatric Hemato-oncology, Ghent University Hospital, Belgium.
6. Department of Pediatrics, Aarhus University Hospital Skejby, Aarhus, Denmark.

* contributed equally

Accepted for publication, JCO.

See Appendix 2

5. “Class Discovery” through Gene expression Profile Analysis: the identification of a new leukemia subgroup among *MLL/AF4*-rearranged patients.

5.1 Introduction

Mixed lineage leukemia (MLL) was recognized more than 30 years ago and is characterized by the presence of both lymphoid and myeloid antigens (1) as well as by a chromosomal translocation at 11q23 (2). The gene spanning the translocation breakpoint was identified as the human homolog of the *Drosophila melanogaster* gene named trithorax (Trx) and because of this relationship and the association with leukemia, the human gene was labeled as either *HRX* (human trithorax), *ALL-1* (acute lymphblastic leukemia-1) or *MLL* (mixed lineage leukemia).

The MLL gene is highly promiscuous as it is translocated in malignancies such as ALL (both B- and T-lineage), AML and myelodysplastic syndrome (MDS) in both children and adults (3) with more than 70 different partner genes mapped on quite all chromosomes (4). Nevertheless, the five most frequent *MLL* rearrangements accounting for approximately 80% of all MLL-translocation-bearing leukemias, are: t(4;11) (q21;q23) or *MLL/AF4*; t(9;11)(p22;q23) or *MLL/AF9*; t(11;19)(q23;p13.3) or *MLL/ENL*; t(10;11)(p12;q23) or *MLL/AF10*; and t(6;11)(q27;q23) or *MLL/AF6* (**Fig. 3**)

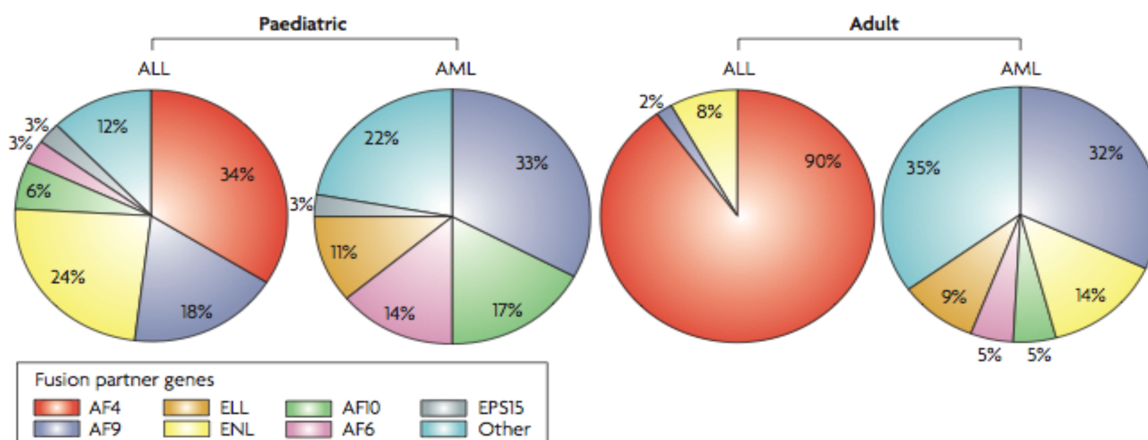


Figure 3. Distribution of the major MLL fusion partner genes in *de novo* childhood and adult leukemia. Mixed lineage leukemia (MLL) rearrangements are found in approximately 5% of acute lymphoblastic leukemia (ALL), in approximately 5–10% of acute myeloid leukemia (AML) and virtually in all cases of mixed lineage leukemia (MLL). Major *MLL* fusion partner genes are *AF4*, which is predominantly found in ALL; *AF9*, which is predominantly found in AML; and *ENL*, which is found in both ALL and AML (from Krivtsov et al, 2007)(5).

MLL rearrangements are found in >70% of infant (less then one years old at the time of diagnosis) leukemia (6) but are less frequent in leukemia involving older children. *MLL* translocations are also found in approximately 10% of AMLs in adult patients, and in therapy-related leukemias (t-leukemias) that develop in patients previously treated with topoisomerase II inhibitors for other malignancies (7). Overall, leukemia with *MLL* rearrangements are found in approximately 10% of all

human leukemia. Patients with *de novo* or therapy-related MLL leukemia display a particular dismal outcome (8, 9).

5.1.2 Structure and function of normal *MLL*

The *MLL* gene spans approximately 90 kb on chromosome 11q23 with nearly 12 kb of coding sequence (consisting of 36 exons) and encodes a protein product of 3969 amino acid residues with a molecular weight of almost 430 kDa (10-12). MLL is proteolytically cleaved by Taspase1 into a 320-kd N-terminal fragment (MLL^N) and a 180-kd C-terminal fragment (MLL^C), which interact with each other in order to avoid degradation and confer stability to the MLL protein (13). MLL presents several functional domains (**Fig. 4**) (14).

The MLL^N terminal fragment, which displays a transcriptional-repression activity, possesses 3 AT hooks involved in the MLL-DNA-binding activity and a transcriptional repression domain (RD) (15). The RD consists of two distinct domains, called RD1 and RD2, and binds several other repressor proteins belonging to the PcG and histone deacetylases (i.e. HDAC1 and HDAC2) (16). Moreover, RD1 contains a DNA methyltransferase (MT) homology domain, known also as CXXC domain, which is essential for the transformation potential of MLL fusion protein, as it has a specific DNA-binding activity for unmethylated CpG DNA sequences (17). The MLL^N fragment also includes multiple PHD (plant homeodomain) zinc finger domains, which are thought to be involved in the transcriptional regulation and chromatin-remodeling activity of MLL (18).

Differently, MLL^C possesses a transcriptional-activation domain that binds to CBP (CREB-binding protein), a histone acetylase that promotes transcriptional activation by acetylating histones H3 and H4 at target gene loci (19). The C-terminal–located SET (su[var]3-9, enhancer of zeste, trithorax) domain, a highly evolutionarily conserved region, has a lysine-directed histone methyltransferase activity that methylates lysine 4 of histone H3, with the results of maintaining active chromatin (20, 21).

The MLL protein is a component of a large multiprotein super-complex with at least 27 proteins, including components of the human transcription complexes TFIID, SWI/ SNF, NuRD, hSNFsH, and Sin3A (20). The complex acetylates, deacetylates and methylates nucleosome-attached histones, resulting in chromatin remodeling. MLL is believed to have a role in the assembly of the complex and in the binding of the multiprotein complex to the promoters of the target genes (20, 21). Two proteins have also been identified as essential for the correct functioning and targeting of the MLL complex: MENIN (22) and LEDGF (lens epithelium derived growth factor) (23).

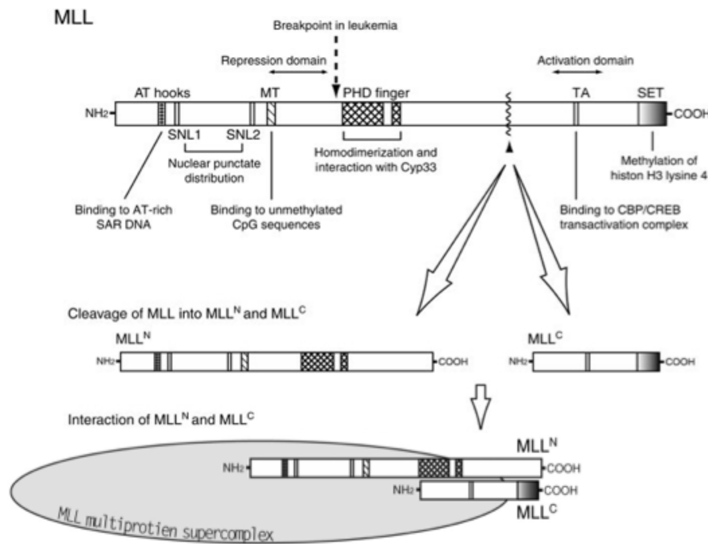


Figure 4. Protein structure of MLL. The different functional domains are shown. MLL is proteolytically cleaved by Taspase1 into MLL^N and MLL^C moieties which interact together to stabilize the protein complex. MLL is a component of a larger multiprotein supercomplex with at least 27 proteins. Dotted vertical arrow indicates the location of the breakpoint occurring in MLL related leukemia. MT: DNA methyltransferase homology domain; PHD: PHD zinc fingers; TA: transactivation domain; SET: SET domain; SNL1 and SNL2: speckled nuclear localization signals 1 and 2 (from Eguchi et al, 2005 and Eguchi et al, 2003)(3, 24)

The MLL multiprotein complex is involved in the regulation of transcription of thousands of genes (25). One of the critical functions of MLL is to maintain *HOX* genes expression (26) such as *HOXA9*, *HOXC8*, *HOXA7* and *HOXC9* (26-28). Moreover, *MLL* has a critical role in the control of the development of axial skeleton (27) and hematopoietic system (29-31). In *MLL*-null embryos, the hematopoietic cells of the yolk sac and the early fetal liver have reduced clonogenic capacity and chimeric mice with *MLL*-/- cells possess adult lymphoid or myeloid cells, which suggest a role for MLL at a multipotent hematopoietic stem cell (HSC) stage.

5.1.3 MLL fusion proteins: structure and mechanisms of action

The *MLL* gene is the frequent target of chromosomal translocation and most of the genomic breakpoints are located within a 8,3 kb fragment encompassed exons 8 to 13 (32). *MLL* has been found in 73 different translocations and 54 partner genes have been cloned (http://atlasgeneticsoncology.org/Genes/MLL.html; last update 5/08). All fusions are in frame and code for a chimeric protein with novel aberrant properties (24). Regardless of the *MLL* fusion partner genes (FPGs), all chimeric proteins include, at the N-terminus, the DNA-binding AT hooks and MT homology domains of MLL but not the PHD zinc finger, and SET domain (**Fig. 5**).

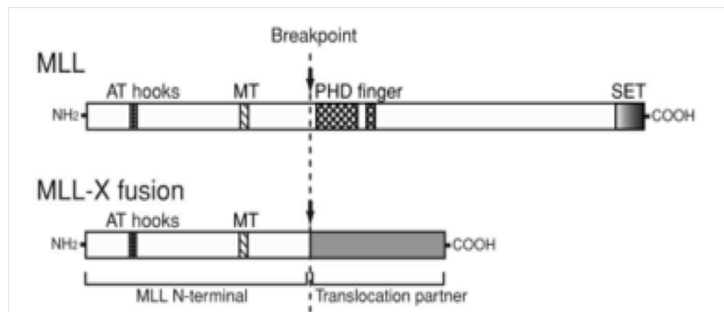


Figure 5. Simplified representation of a generic MLL fusion protein. The MLL^N terminal fragment of MLL is fused in frame with the C-terminal portion of a generic fusion partner protein (from Eguchi et al, 2005)(3).

The differences between the various *MLL* fusions therefore reside in the domains contributed by fusion partners. Some of the fusion partner genes encode nuclear proteins with transcriptional regulation activity, whereas some FPGs are cytoplasmatic with diverse functions. Neverthless, all the MLL chimeric proteins localized to the nucleus, regardless of whether the fusion gene itself is nuclear or cytoplasmatic (33, 34). Despite the huge number of FPGs, the most frequent MLL translocation partner genes can be grouped in 4 clusters: the first comprises genes encoding for nuclear protein such as AF4, AF9, AF10, ENL and ELL; the second one includes cytoplasmic proteins such as GAS7, EEN, AF6, AFX; the third group includes the septins (SEPT2, SEPT5, SEPT6, SEPT9 and SEPT11) and the fourth includes histone acetyltransferases p300 and CBP (5).

The transforming capacity of MLL fusion proteins is believed to be mediate through four major mechanisms. As the MLL fusion proteins lack the SET domain, the first mechanism is believed to be related to a perturbation of the normal H3K4 methylation. Furthermore, functional studies of the respective fusion proteins have revealed that the minimal domains required for transformation display transactivation properties (35, 36). This data supports that a common mechanism of transformation in at least some MLL's nuclear protein fusion partner is that the partner gene provides a transcriptional domain to the MLL fusion protein (37, 37, 38). A third mechanism of gene expression activation relates to the property of the *MLL* partner gene to direct oligomerization of the fusion gene (38). Finally, accumulating evidences suggest that many MLL fusion partners belong to a network involved in transcriptional regulation through chromatin remodeling (39). For example, the MLL-fusion partner AF10 associates with the DOT1L histone methyltransferase which methylates lysine 79 residues in histone H3 (H3K79) (40).

Moreover, Cozzio et al (2003)(41) have provided important insights into MLL fusion-protein mediated leukemogenesis: when assessing the transformation potential of MLL/ENL in committed myeloid progenitors, the authors demonstrated that mouse myeloid leukemia can originate not only from HSCs but also from committed myeloid progenitors that lack self-renewal potentials. Thus, MLL/ENL appears to be able to re-activate some parts of the hematopoietic stem-cell program. This finding has also been confirmed by the analysis of the MLL/AF9 transformation potential in granulocyte macrophage progenitors (GMP) (42). Moreover, the analysis of the genes activated

immediately after the induction of MLL/AF9 has showed that only a small subset including the *HoxA* cluster genes (i.e. *HoxA5*, *HoxA9*, *HoxA10*) *Meis1* and *Mef2c* is up-regulated (43), whereas the full signature evolves over time probably as a consequence of additional events (44). Therefore, the MLL fusion does not seem to induce de-differentiation but, rather, appears to generate a cell with stem-cell like properties and gene expression programmes at an inappropriate stage of the hematopoietic differentiation.

5.1.4 MLL fusion proteins and *HOXA* gene expression deregulation

Several gene expression studies on MLL-rearranged leukemia (ALL and AML) have demonstrated that MLL leukemia represents a distinct leukemia subtype and have also provided insights into the gene expression profiles of these diseases and their downstream targets (45-48). Among these targets, *HOXA7*, *HOXA9*, *HOXA10*, *HOXC8* and *MEIS1* are involved in the MLL fusion gene mediated oncogenesis (45-47, 49).

HOX genes belong to the homeobox (*Hox*) gene family and encode a group of transcription factors that control both embryonic development and hematopoietic cell differentiation (50). In general, HOX transcription factors are not only master controls of embryonic development but they also direct normal hematopoietic differentiation. *HOX* expression is high in stem cells and early precursors and needs to be down-regulated for maturation. Therefore, a continuous ectopic *HOX* expression through MLL-fusion protein can block differentiation and create a rapidly proliferating pre-leukemic precursor pool.

Besides, MLL wild type is a positive regulator of *HOX* genes expression (26, 51) and over-expression of *HoxA9* has proved to alter normal hematopoiesis in mice, thus leading to development of leukemia (52). Even though the up-regulation of *HOXA* genes is considered a key mechanism of leukemic transformation initiated by MLL fusion proteins, whether the over-expression of *HOXA* genes is essential for leukemia onset and progression or not is still controversial.

Several studies on myeloid MLL-rearranged leukemia using *in vivo* and *in vitro* models suggest a direct role of *HOXA* genes in the development of leukemia (42, 53, 54), even though some works (55, 56) have questioned the role of *HOXA* genes in myeloid MLL related leukemias.

5.2 Gene Expression Profile Analysis of B-cell precursors (BCPs) ALL pediatric patients with *MLL/AF4*

The most prevalent *MLL* fusion gene in pediatric infant patients is the translocation *MLL/AF4* or t(4;11) (q21;q23), which fuses *MLL* with the *AF4* gene located on chromosome 4 band q21.3-5 (57-

59). The *AF4* gene encodes a serine/proline-rich protein containing a nuclear localization signal and a guanoside triphosphate (GTP-binding domain). It is localized to the nucleus and it is probably involved in the control of gene transcription. Moreover, *AF4*-deficient mice exhibit imperfect T-cell development and modest alteration in B-cell development (60). The t(4;11) translocation generates two fusion genes, *MLL/AF4* and *AF4/MLL*; the significance of either fusion genes for leukemogenesis is not yet completely understood. Gaussmann et al 2007(61) have demonstrated that both fusion proteins are involved in the pathological disease mechanism of t(4;11) leukemia; the MLL/AF4 fusion protein increases resistance against apoptosis, whereas AF4/MLL fusion protein is able to growth transform cells.

The critical requirement of *HOXA* genes expression (particularly *HOXA9*) in BCPs ALL, and particularly in MLL/AF4 leukemia, has been poorly investigated so far, although the high transcriptional level of *HOXA9* has proved to be a specific feature also in lymphoid leukemia bearing *MLL* rearrangements (62) and *HOXA9* suppression has shown to interfere with cell viability in cell lines harboring the *MLL/AF4* fusion gene (54).

We have analyzed the GEP of 20 *MLL/AF4* positive B-cell precursors ALL pediatric patients at diagnosis; after a pairwise subtraction approach using the gene expression profile of three normal bone marrow samples, we could first identify a common gene signature for all the analyzed patients. Secondarily, considering four distinct biological discriminators (*HOXA* gene expression, age at diagnosis, fusion gene transcripts and chromosomal breakpoints) we could divide patients into two distinct subgroups: the first one comprised infant patients with low *HOXA* genes expression and the *MLL* breakpoints within introns 11/12. The second one comprised non-infant patients with high *HOXA* expression and *MLL* breakpoints within introns 9/10. Thus, we demonstrated once again that GEP technology could be successfully used for the identification (class discovery) of new leukemia subtypes among already defined leukemia entities. Moreover, the outstanding differential expression of *HOXA* genes in the analyzed t(4;11) specimens assigns a less prominent role to *HOXA* genes in MLL-related ALL leukemogenesis, which has also been observed in models of MLL-related acute myeloid leukemia (55, 56). Moreover, the presence of two separate clusters among t(4;11)-positive patients has been recently confirmed (63). These findings are particularly remarkable, as *HOXA* over-expression is believed to be a hallmark of MLL-rearranged leukemia (45, 47).

5.2.1

Two independent gene signatures in pediatric t(4;11) acute lymphoblastic leukemia patients

Luca Trentin[1*], Marco Giordan[1*], Theo Dingermann[2], Giuseppe Basso[1], Geertruy te Kronnie[1], Rolf Marschalek[2]

[1] Hemato-Oncology, Dept. of Pediatrics, University of Padova, Via Giustiani 3, Padova, Italy;

[2] Institute of Pharmaceutical Biology / ZAFES / DCAL / CEF, JWG-University Frankfurt, Biocenter, Max-von-Laue-Str. 9, Frankfurt/Main, Germany.

* contributed equally

Eur J Haematol. 2009 Nov; 83(5):406-19.

See Appendix 3

5.3 Molecular analysis of “HOXA high” and “HOXA low” *MLL/AF4* patients

The remarkable difference in *HOXA* gene expression identified through GEP in 20 MLL/AF4 patients has been further investigated.

At first, using the GEP raw data of the 20 analyzed samples, we searched for any differential expression of genes commonly associated to *MLL*; apart from the differential *HOXA* genes expression level (*HOXA9, HOXA7, HOXA4, HOXA5, HOXA10*), no transcriptional deregulation of other known MLL-related genes (i.e. *MENIN*, *HOXC8* and *MEIS1*) could be identified (Fig. 5 A-E). The analysis of the transcriptional expression of *MEIS1* (Fig. 6 A-C) showed that only one of the “HOXA high” patients displayed a weaker *MEIS1* transcription level.

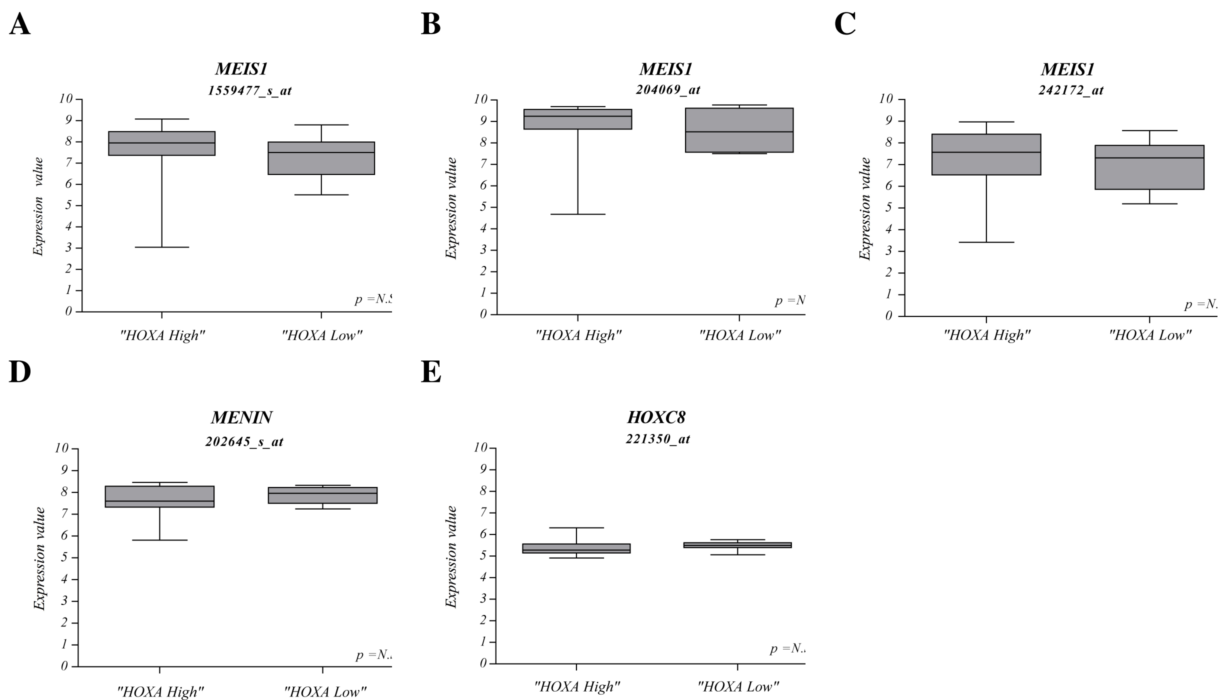


Figure 6. The normalized expression values of the probe sets relative to the *MEIS1* (A-C), *MENIN* (D), and *HOXC8* (E) genes. The normalized expression value of the selected genes is reported in log2 scale as measured by the HG-U133 Plus 2.0 array in 20 t(4;11) pediatric patients at diagnosis.

The up-regulation of *MEIS1* in both groups of patients is particularly intriguing: *MEIS1* alone is not able to induce leukemia, whereas its co-expression with *HOXA9* promotes leukemia initiation and maintenance. Moreover, Faber et al regard *MEIS1* as a downstream target of *HOXA9*, raising the question whether continued *HOXA9* expression is required to maintain high levels of *MEIS1* in blast cells or not. Our results demonstrated that *MEIS1* over-expression in t(4;11) leukemic cells is maintained irrespective of the presence of up-regulated *HOXA9* gene.

Then, we also validated the differential *HOXA* expression in BCPs ALL t(4;11) through *HOXA9* quantitative reverse transcriptase-PCR (qRT-PCR) (**Fig. 7**) in a larger cohort of patients. We analyzed 26 BCPs ALL t(4;11) patients and the qRT-PCR confirmed the *HOXA9* down-regulation in 9 patients (p-value=0.0002).

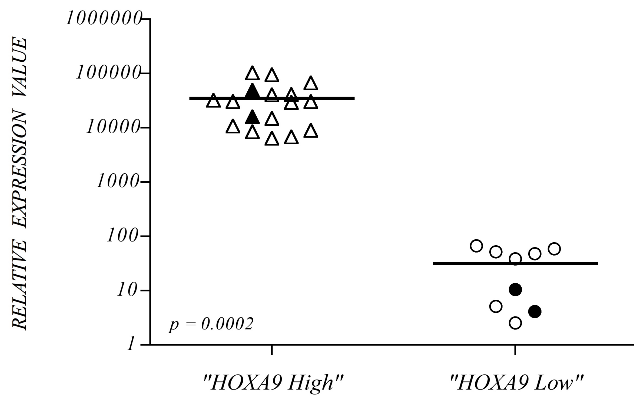


Figure 7. *HOXA9* real-time quantitative PCR analysis was performed on 22 t(4;11) patients at diagnosis and 2 paired samples both at diagnosis and relapse. The pair belonging to the "HOXA high" group is shown as black-filled triangles and the pair belonging to the "HOXA low" group is shown as black-filled circles. Fifteen patients at diagnosis (8 "HOXA high" and 7 "HOXA low") belong to the initial cohort of samples analyzed by GEP. The relative *HOXA9* gene expression value has been calculated in the 26 samples using the $2^{-\Delta\Delta Ct}$ method and the normal CD19+ as calibrator. *GUS B* was used as reference gene. The real time quantitative PCR analysis confirms the presence of a highly significant difference in *HOXA9* gene expression among the 26 specimens. The relative *HOXA9* expression value is reported in log10 scale. All the p-values have been calculated using the two-sample Welch t-statistics (unequal variance) in PRISM 4 Version 4.0 (GraphPad Software).

We then questioned whether the outstanding difference in *HOXA9* mRNA expression between "HOXA high" and "HOXA low" patients could also be recognized at protein level.

We analyzed the HOXA9 protein expression level in three *MLL/AF4* positive patients (i.e. two "HOXA low" and one "HOXA high") at diagnosis, whose gene expression profiles were also available. We analyzed also the HOXA9 expression of the RS4;11 cell line that carries the *MLL/AF4* fusion gene and one normal bone marrow (BM) control (**Fig. 8**).

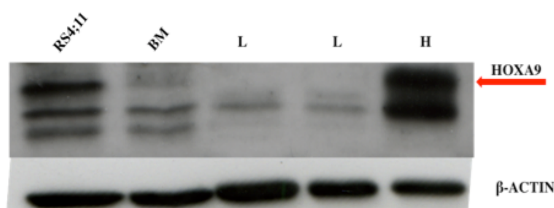


Figure 8. Western Blot analysis of 3 *MLL/AF4* pediatric patients. The immunoblots show the HOXA9 and β-ACTIN expression level in one cell line (RS4;11), 3 MLL/AF4 positive patients and 1 normal control. β-ACTIN was used as loading control. L: "HOXA low" patient; H: "HOXA high" patient; BM: normal bone marrow control. The proteins of the 3 patients and of the normal control were extracted from total BM and 20 μg of lysate were loaded.

As shown in **Fig. 8**, the two "HOXA low" patients lacked HOXA9 expression to the same as the normal BM control, whereas the "HOXA high" patient expressed HOXA9 to an extent not seen also in the RS4;11 cell line.

On the whole, this data proved that *MLL/AF4* leukemia develops regardless of the contribution of the *HOXA* genes (mRNA and protein) expression and this prompts new questions on the function of *HOXA* genes in MLL-mediated leukemogenesis.

We then aimed to identify a rationale for the absence of *HOXA* genes expression in patients with *MLL/AF4*.

At first, we collected the available information about the karyotype of the analyzed patients; no particular abnormalities affecting the chromosomal region encompassing the *HOXA9* locus on chromosome 7 could be identified. Moreover, an independent study by Bardini et al (2009)(64) analyzed the single nucleotide polymorphism (SNP) profile of 28 pediatric *MLL/AF4* patients using the 100K SNP arrays; this study revealed that only a limited numbers of deletions/amplifications could be identified in these patients. Thus, it seems unlikely that the differential *HOXA* cluster genes expression is due to additional genetic abnormalities.

We also searched for any clinical differences between the two recognized subgroups. The two different identified signatures didn't reflect a diverse clinical disease course and both groups of patients presented a fully developed leukemia; we were not able to identify, indeed, any significant correlations between the two subgroups and clinical data such as gender, age at diagnosis, risk group stratification, disease status and white blood cells count (WBCs) at diagnosis. Nevertheless, age at diagnosis in all "HOXA low" patients is below six months, whereas age at diagnosis in "HOXA high" patients is heterogeneously distributed.

Besides confirming the presence of two distinct gene signatures among *MLL/AF4* infant ALL, the recent work by Stam et al (2009)(63) has showed that patients lacking *HOXA* genes expression are at extreme high risk of disease relapse. We did not identify the same correlation, which may be due to the limited number of patients present in our study.

Moreover, we analyzed the clusters of differentiation (CD) expression values in "HOXA low" and "HOXA high" patients; we used only data obtained making use of the same flow cytometry instrument in order to avoid possible differences in fluorescence detection. We compared the CD45, CD34, CD58, CD19, CD24 and CD52 expression value in 16 patients (10 "HOXA high" and 6 "HOXA low"). We identified a higher expression of CD34 in "HOXA low" patients in comparison with "HOXA high" patients (**Fig. 9**). This significant difference ($p = 0.0095$) was recognizable when considering the percentage (%) of CD34+ blast cells (**Fig. 9 A**); this statistically significant difference was not achieved when considering the mean intensity expression value of CD34 between the same subgroups of patients (**Fig. 9 B**). Nevertheless, a trend toward higher CD34 expression could be all the same recognized in "HOXA low" patients.

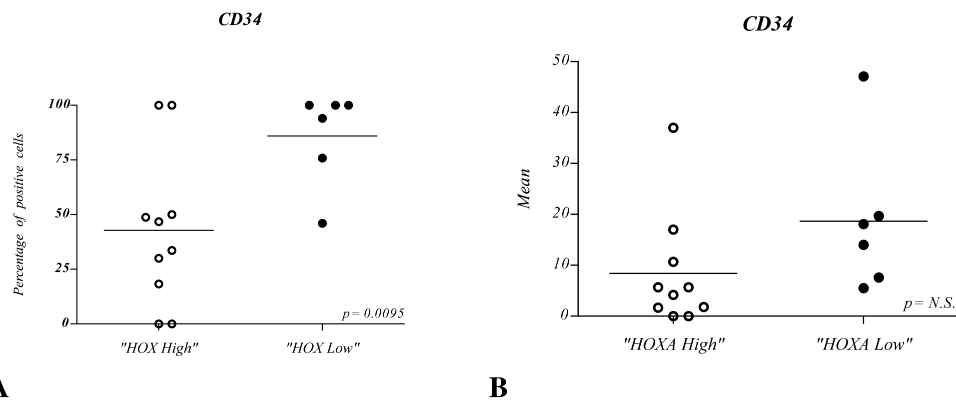


Figure 9. CD34 expression value in 16 *MLL/AF4* patients at diagnosis. A: CD34 expression value considering the percentage (%) of CD34+ blast cells. 11 B: CD34 expression value considering the mean intensity value of CD34 + blast cells. Each dot represents a patient. The p-values have been calculated using the two-sample Welch t-statistics (unequal variance) in PRISM 4 Version 4.0 (GraphPad Software).

As *HOXA9* was found to be down-regulated through its promoter methylation in several malignancies such as ovarian carcinomas (65) and neuroblastoma (66), we have analyzed the methylation status of the *HOXA9* promoter in 2 “HOXA high” and 2 “HOXA low” MLL/AF4 patients; for one “HOXA low” patient the paired material at diagnosis and relapse was also available for analysis. Besides, we used a mix of normal peripheral blood (PB) as negative control and a cell line, kindly provided by Dott. Franco Fais, harboring the *MLL/AF4* fusion gene and expressing low transcriptional level of *HOXA* genes as positive control; this cell line, indeed, resulted to be methylated in the *HOXA9* promoter region. As shown in **Fig. 10**, the amplified *HOXA9* promoter region was not methylated in the analyzed samples regardless of the *HOXA9* high or *HOXA9* low expression level. After DNA bisulphite conversion, the MSP products were also sequenced in order to verify the accuracy of the amplified promoter sequence confirming that promoter sequences are not methylated in “HOXA low” patients.

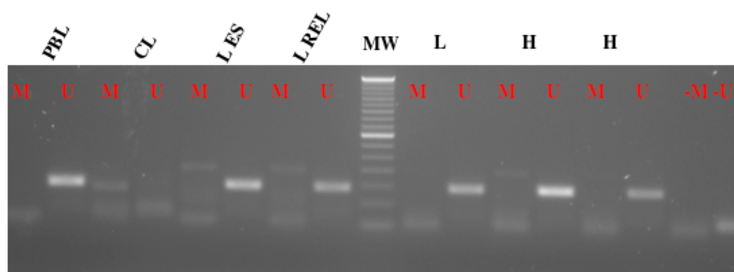


Figure 10. Methylation Specific PCR (MSP) for *HOXA9*. PBL: mix of 5 different PB normal controls; CL: cell line with low *HOXA* genes expression level. L ES: “HOXA low” patient at diagnosis; L REL: “HOXA low” patient at relapse; MW: marker 50 bp; L: “HOXA low” patient; H: “HOXA high” patient. M: methylated; U: unmethylated; -M and -U: MSP negative controls. *HOXA9* MSP-M 127 bp; *HOXA9* MSP-U 139bp.

Therefore, we conclude that *HOXA9* silencing in these *MLL/AF4* patients is not due to promoter methylation.

The recent work by Popovic et al 2009(67) has demonstrated that wild-type *MLL* regulates the expression of the microRNA mir-196b located in a highly evolutionary conserved region between *HOXA9* and *HOXA10* genes at chromosome band 7p15.2 in human. The study indicates that *MLL* regulates hsa-mir-196b in a pattern similar to that of the surrounding genes and that mir-196b is over-expressed specifically in primary leukemia samples from patients with *MLL*-fusion gene but not from other subtypes of leukemia and expression of MLL fusion proteins in primary bone marrow cells causes over-expression of mir-196b. Furthermore, mir-196b expression increases proliferation and survival, and also partially blocks differentiation of normal bone marrow hematopoietic progenitor cells. Thus, up-regulation of mir-196b by MLL fusion genes is an important component in leukemia development caused by MLL-fusion proteins.

On the basis of this data, we analyzed the microRNA expression profile of 3 "HOXA high", 3 "HOXA low" patients, 2 normal BM controls and 2 cell lines (i.e. the RS4;11 and the cell line with low *HOXA* genes expression level), as we aimed to evaluate whether the mir-196b was expressed or not in "HOXA high" and "HOXA low" patients independently from the *HOXA* genes expression level.

Interestingly, mir-196b resulted to be down-regulated in the 3 "HOXA low" patients compared to the 3 "HOXA high" patients and resembled the microRNA's expression level detectable in normal BM controls (**Fig. 11**).

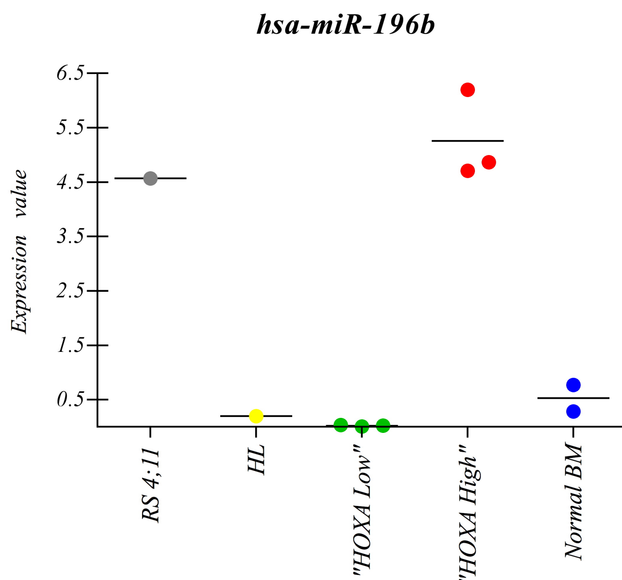


Figure 11. The expression level of hsa-mir-196b in 3 "HOXA low", 3 "HOXA high", 2 normal BM controls and 2 cell lines is illustrated. Each dot represents a sample; the expression value is expressed in log2 scale. HL: cell line with low HOXA genes expression value

Thus, leukemia develops in the “HOXA low” *MLL/AF4* patients not only without the involvement of HOXA de-regulation but also without the aberrant mir-196b expression mediated by the *MLL/AF4* fusion gene.

Moreover, we questioned whether it was possible to divide the 3 “HOXA high” and the 3 “HOXA low” patients on the basis of differentially expressed microRNAs. We performed a class comparison analysis between the two subgroups and we could clearly divide the two subgroups of patients using the microRNAs whose expression varied by at least 2.0 fold (**Fig. 12**). Remarkably, the same division was obtained also performing an unsupervised analysis using all the microRNA probes spotted in the array.

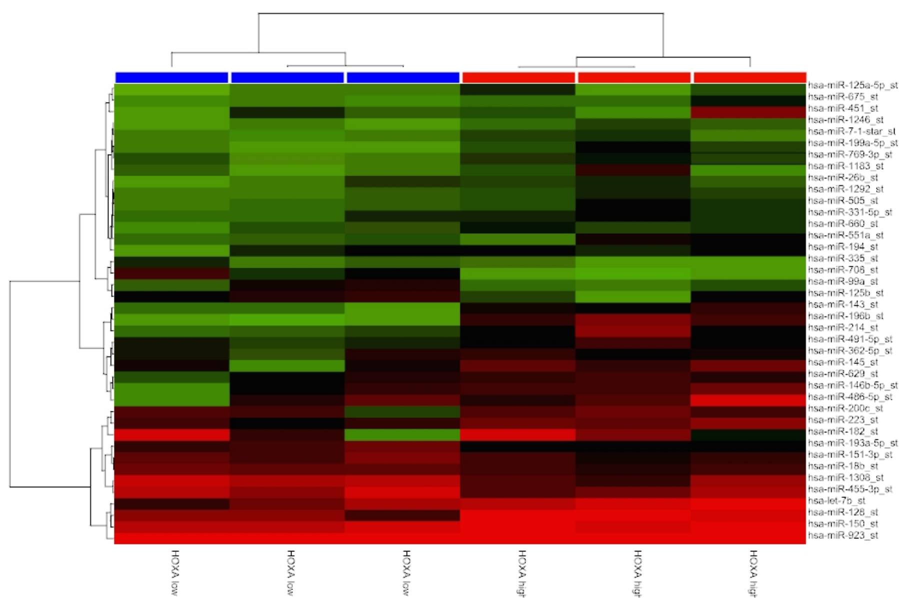


Figure 12. Heat map showing the differentially expressed microRNAs between “HOXA high” and “HOXA low” patients. The red and blue bars indicate “HOXA high” and “HOXA low” patients respectively. Each column represents a patient and each row represents an individual microRNA. Expression levels are normalized for each microRNA and the expression level is indicated by color: intensity of red is proportional to the degree of expression above the mean and intensity of green is proportional to the degree of expression below the mean. For the heat map generation, only microRNAs with fold change ≥ 2.0 were used.

Thus, the two subgroups of patients (i.e. “HOXA high” and “HOXA low”) identified using GEP analysis are also characterized by a specific microRNA expression signature.

5.3.1 Materials and Methods

Patients and cell lines.

Bone marrow aspirates or peripheral blood (PB) samples of 26 patients with t(4;11) (q21;q23) B-cell precursor (BCP) acute lymphoblastic leukemia (ALL) were studied. All samples were t(4;11)

(q21;q23) positive. Mononuclear cells from patients and two peripheral blood (PB) samples used as normal controls were isolated using Ficoll-Hypaque (Amersham Biosciences) density gradient centrifugation. CD19+ B cells were isolated from control PB by positive selection using Milthenyi CD19 Microbeads (Milthenyi Biotec). The purity of the isolated CD19+ B cells was over 98%, as determined by flow cytometric analysis.

The RS4;11 cell line was obtained from DSMZ (Braunschweig, Germany); the cell line with low *HOXA* gene expression level was a kind gift of Dott. Franco Fais (University of Genoa). Both cell lines were maintained in RPMI 1640 (Biochrom AG, Berlin, Germany) with 10% FCS, penicillin (100 units/ml) and streptomycin (100 μg/ml) and maintained at 37°C in a humidified atmosphere with 5% CO_2.

RNA isolation and real-time PCR analysis.

Total RNA was isolated using TRIzol reagent (Invitrogen). The RNA quality and concentration were assessed on the Agilent 2100 Bioanalyzer (Agilent Technologies) and the NanoDrop ND-1000 spectrofotometer (NanoDrop Technologies), respectively. cDNA was synthesized using SuperScripII reverse transcriptase (Invitrogen) and 1μg of RNA. Real-time PCR analysis was performed on an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems) using Platinum SYBR Green qPCR SuperMix UDG (Invitrogen). *HOXA9* and *GUS B* primers sequences are the following: Fw 5'-GGTGACTGTCCCACGCTTGAC-3', Rw 5'-GAGTGGAGCGCGCATGAAG-3', Fw 5'-GAAAATATGTGGTTGGAGAGCTCATT-3', Rw 5'- CCGAGTGAAGATCCCCTTTTTA-3'.

Western Blot.

20 μg from total protein fraction (Buffer-Biosource by Invitrogen-Gibco) obtained from the 2 cell lines and 3 BM patients were used to perform protein analysis. Protein concentration was determined using the BCA method (Pierce). Samples were ran on a 15% SDS-plyacrylamide gel and then transferred to 0,2 μm polyvylidene difluoride membrane (GE-Healthcare) for immunodetection with specific antibodies followed by horseradish peroxidase-conjugated goat anti-rabbit or mouse IgG (Upstate). Anti-HOXA9 antibody was obtained by Millipore (catalog# 07-178) and anti β–ACTIN was obtained by Sigma Aldrich. The specific bands of the target proteins were visualized by enhanced chemoluminescence (ECL advance) according to the manufacturer's instructions (GE-Healthcare).

Sodium bisulfite modification of DNA and methylation-specific plymerase chain reaction (MSP).

Gemonic DNA was extracted using the Purogene kit (Gentra); 2 μg of DNA was subjected to bisulfite conversion using the EZ DNA Methylation-Gold Kit (Zymo Research). Methylation analysis was carried out by methylation specific PCR (MSP) using 200 ng of modified DNA and *HOXA9* specific primers. HOXA9-MSP-M Fw 5'- GGTTAATGGGGGCGCGGGCGTC-3'; HOXA9-MSP-M Rw5'-AACGCCTAACCCGCCCGACCCG-3'; HOXA9-MSP-U Fw 5'-

GTATGGTTAATGGGGGTGTGGGTGTT-3'; HOXA9-MSP-U Rw5'-CCATACCCAACACCTAACCCACCCAACCCA-3' (65).

MicroRNA Expression and data analysis.

1 μg of total RNA was used for the experiments. The RNA was labeled using the FlashTag Kit (Genisphere) following the manufacturer instruction. Briefly, the process begins with a tailing reaction followed by ligation of the biotinylated signal molecule to target RNA sample. The biotin-labeled RNA was then hybridized on miRNA arrays (Affymetrix). Hybridization, washing, staining and scanning protocols were performed on Affymetrix GeneChip instruments (Hybridization Oven 640, Fluidics Station 450Dx, Scanner GCS3000Dx, respectively), following the manufacturer's instructions. All arrays used for the following analysis passed the overall quality control. MicroRNA expression data was analyzed with Partek Genomics Suite Software.

5.4 References

1. Stass SA, Mirro J,Jr. Lineage heterogeneity in acute leukaemia: acute mixed-lineage leukaemia and lineage switch. Clin Haematol 1986; 15: 811-27.

2. Djabali M, Selleri L, Parry P, Bower M, Young BD, Evans GA. A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias. Nat Genet 1992; 2: 113-8.

3. Eguchi M, Eguchi-Ishimae M, Greaves M. Molecular pathogenesis of MLL-associated leukemias. Int J Hematol 2005; 82: 9-20.

4. Meyer C, Kowarz E, Hofmann J, Renneville A, Zuna J, Trka J, et al. New insights to the MLL recombinome of acute leukemias. Leukemia 2009; 23: 1490-9.

5. Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer 2007; 7: 823-33.

6. Biondi A, Cimino G, Pieters R, Pui CH. Biological and therapeutic aspects of infant leukemia. Blood 2000; 96: 24-33.

7. Felix CA. Secondary leukemias induced by topoisomerase-targeted drugs. Biochim Biophys Acta 1998; 1400: 233-55.

8. Chen CS, Sorensen PH, Domer PH, Reaman GH, Korsmeyer SJ, Heerema NA, et al. Molecular rearrangements on chromosome 11q23 predominate in infant acute lymphoblastic leukemia and are associated with specific biologic variables and poor outcome. Blood 1993; 81: 2386-93.

9. Balgobind BV, Raimondi SC, Harbott J, Zimmermann M, Alonzo TA, Auvrignon A, et al. Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study. Blood 2009; 114: 2489-96.

10. Nilson I, Lochner K, Siegler G, Greil J, Beck JD, Fey GH, et al. Exon/intron structure of the human ALL-1 (MLL) gene involved in translocations to chromosomal region 11q23 and acute leukaemias. Br J Haematol 1996; 93: 966-72.

11. Marschalek R, Nilson I, Lochner K, Greim R, Siegler G, Greil J, et al. The structure of the human ALL-1/MLL/HRX gene. Leuk Lymphoma 1997; 27: 417-28.

12. Rasio D, Schichman SA, Negrini M, Canaani E, Croce CM. Complete exon structure of the ALL1 gene. Cancer Res 1996; 56: 1766-9.

13. Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P, Korsmeyer SJ. Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. Mol Cell Biol 2003; 23: 186-94.

14. 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: 3710-8.

15. Zeleznik-Le NJ, Harden AM, Rowley JD. 11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. Proc Natl Acad Sci U S A 1994; 91: 10610-4.

16. 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: 8342-7.

17. Allen MD, Grummitt CG, Hilcenko C, Min SY, Tonkin LM, Johnson CM, et al. Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltransferase. EMBO J 2006; 25: 4503-12.

18. 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: 3589-97.

19. 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: 2249-58.

20. Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, Wassell R, et al. ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol Cell 2002; 10: 1119-28.

21. Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, et al. MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell 2002; 10: 1107-17.

22. Yokoyama A, Somervaille TC, Smith KS, Rozenblatt-Rosen O, Meyerson M, Cleary ML. The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. Cell 2005; 123: 207-18.

23. Yokoyama A, Cleary ML. Menin critically links MLL proteins with LEDGF on cancer-associated target genes. Cancer Cell 2008; 14: 36-46.

24. Eguchi M, Eguchi-Ishimae M, Greaves M. The role of the MLL gene in infant leukemia. Int J Hematol 2003; 78: 390-401.

25. Milne TA, Dou Y, Martin ME, Brock HW, Roeder RG, Hess JL. MLL associates specifically with a subset of transcriptionally active target genes. Proc Natl Acad Sci U S A 2005; 102: 14765-70.

26. 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: 10632-6.

27. Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ. Altered Hox expression and segmental identity in Mll-mutant mice. Nature 1995; 378: 505-8.

28. Yokoyama A, Wang Z, Wysocka J, Sanyal M, Aufiero DJ, Kitabayashi I, et al. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. Mol Cell Biol 2004; 24: 5639-49.

29. Hess JL, Yu BD, Li B, Hanson R, Korsmeyer SJ. Defects in yolk sac hematopoiesis in Mll-null embryos. Blood 1997; 90: 1799-806.

30. Yagi H, Deguchi K, Aono A, Tani Y, Kishimoto T, Komori T. Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. Blood 1998; 92: 108-17.

31. Fidanza V, Melotti P, Yano T, Nakamura T, Bradley A, Canaani E, et al. Double knockout of the ALL-1 gene blocks hematopoietic differentiation in vitro. Cancer Res 1996; 56: 1179-83.

32. Ayton PM, Cleary ML. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. Oncogene 2001; 20: 5695-707.

33. Yano T, Nakamura T, Blechman J, Sorio C, Dang CV, Geiger B, et al. Nuclear punctate distribution of ALL-1 is conferred by distinct elements at the N terminus of the protein. Proc Natl Acad Sci U S A 1997; 94: 7286-91.

34. Joh T, Yamamoto K, Kagami Y, Kakuda H, Sato T, Yamamoto T, et al. Chimeric MLL products with a Ras binding cytoplasmic protein AF6 involved in t(6;11) (q27;q23) leukemia localize in the nucleus. Oncogene 1997; 15: 1681-7.

35. DiMartino JF, Miller T, Ayton PM, Landewe T, Hess JL, Cleary ML, et al. A carboxy-terminal domain of ELL is required and sufficient for immortalization of myeloid progenitors by MLL-ELL. Blood 2000; 96: 3887-93.

36. Luo RT, Lavau C, Du C, Simone F, Polak PE, Kawamata S, et al. The elongation domain of ELL is dispensable but its ELL-associated factor 1 interaction domain is essential for MLL-ELL-induced leukemogenesis. Mol Cell Biol 2001; 21: 5678-87.

37. So CW, Cleary ML. Common mechanism for oncogenic activation of MLL by forkhead family proteins. Blood 2003; 101: 633-9.

38. 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: 99-110.

39. Erfurth F, Hemenway CS, de Erkenez AC, Domer PH. MLL fusion partners AF4 and AF9 interact at subnuclear foci. Leukemia 2004; 18: 92-102.

40. Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, et al. hDOT1L links histone methylation to leukemogenesis. Cell 2005; 121: 167-78.

41. Cozzio A, Passegue E, Ayton PM, Karsunky H, Cleary ML, Weissman IL. Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors. Genes Dev 2003; 17: 3029-35.

42. Krivtsov AV, Twomey D, Feng Z, Stubbs MC, Wang Y, Faber J, et al. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature 2006; 442: 818-22.

43. Somervaille TC, Cleary ML. Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. Cancer Cell 2006; 10: 257-68.

44. Faber J, Armstrong SA. Mixed lineage leukemia translocations and a leukemia stem cell program. Cancer Res 2007; 67: 8425-8.

45. Armstrong SA, Staunton JE, Silverman LB, Pieters R, den Boer ML, Minden MD, et al. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat Genet 2002; 30: 41-7.

46. Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell 2002; 1: 133-43.

47. Ferrando AA, Armstrong SA, Neuberg DS, Sallan SE, Silverman LB, Korsmeyer SJ, et al. Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: dominance of HOX dysregulation. Blood 2003; 102: 262-8.

48. Kohlmann A, Schoch C, Dugas M, Schnittger S, Hiddemann W, Kern W, et al. New insights into MLL gene rearranged acute leukemias using gene expression profiling: shared pathways, lineage commitment, and partner genes. Leukemia 2005; 19: 953-64.

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

50. Sauvageau G, Lansdorp PM, Eaves CJ, Hogge DE, Dragowska WH, Reid DS, et al. Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. Proc Natl Acad Sci U S A 1994; 91: 12223-7.

51. Milne TA, Martin ME, Brock HW, Slany RK, Hess JL. Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9 locus, promoting transcription and multiple histone modifications. Cancer Res 2005; 65: 11367-74.

52. Kroon E, Krosl J, Thorsteinsdottir U, Baban S, Buchberg AM, Sauvageau G. Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. EMBO J 1998; 17: 3714-25.

53. Zeisig BB, Milne T, Garcia-Cuellar MP, Schreiner S, Martin ME, Fuchs U, et al. Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. Mol Cell Biol 2004; 24: 617-28.

54. Faber J, Krivtsov AV, Stubbs MC, Wright R, Davis TN, van den Heuvel-Eibrink M, et al. HOXA9 is required for survival in human MLL-rearranged acute leukemias. Blood 2009; 113: 2375-85.

55. 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: 1823-8.

56. 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: 3192-9.

57. Gu Y, Nakamura T, Alder H, Prasad R, Canaani O, Cimino G, 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.

58. McCabe NR, Burnett RC, Gill HJ, Thirman MJ, Mbangkollo D, Kipiniak M, et al. Cloning of cDNAs of the MLL gene that detect DNA rearrangements and altered RNA transcripts in human leukemic cells with 11q23 translocations. Proc Natl Acad Sci U S A 1992; 89: 11794-8.

59. Domer PH, Fakharzadeh SS, Chen CS, Jockel J, Johansen L, Silverman GA, et al. Acute mixed-lineage leukemia t(4;11)(q21;q23) generates an MLL-AF4 fusion product. Proc Natl Acad Sci U S A 1993; 90: 7884-8.

60. Isnard P, Core N, Naquet P, Djabali M. Altered lymphoid development in mice deficient for the mAF4 proto-oncogene. Blood 2000; 96: 705-10.

61. Gaussmann A, Wenger T, Eberle I, Bursen A, Bracharz S, Herr I, et al. Combined effects of the two reciprocal t(4;11) fusion proteins MLL.AF4 and AF4.MLL confer resistance to apoptosis, cell cycling capacity and growth transformation. Oncogene 2007; 26: 3352-63.

62. Rozovskaia T, Feinstein E, Mor O, Foa R, Blechman J, Nakamura T, et al. Upregulation of Meis1 and HoxA9 in acute lymphocytic leukemias with the t(4 : 11) abnormality. Oncogene 2001; 20: 874-8.

63. Stam RW, Schneider P, Hagelstein JA, van der Linden MH, Stumpel DJ, de Menezes RX, et al. Gene expression profiling-based dissection of MLL translocated and MLL germline acute lymphoblastic leukemia in infants. Blood 2009;.

64. Bardini, M., Spinelli,R., Corral,L., Mangano, E., Fazio,G.,Cifola, I., Biondi,A., Battaglia, C., Cazzaniga, G. INFANT ALL PATIENTS WITH T(4;11)/MLL-AF4 FUSION HAVE A DIFFERENT GENOTYPIC PROFILE THAN OLDER ALL CHILDREN. Haematologica 2008; 93: 63 Abs.0159.

65. Wu Q, Lothe RA, Ahlquist T, Silins I, Trope CG, Micci F, et al. DNA methylation profiling of ovarian carcinomas and their in vitro models identifies HOXA9, HOXB5, SCGB3A1, and CRABP1 as novel targets. Mol Cancer 2007; 6: 45.
66. Margetts CD, Morris M, Astuti D, Gentle DC, Cascon A, McRonald FE, et al. Evaluation of a functional epigenetic approach to identify promoter region methylation in phaeochromocytoma and neuroblastoma. Endocr Relat Cancer 2008; 15: 777-86.
67. Popovic R, Riesbeck LE, Velu CS, Chaubey A, Zhang J, Achille NJ, et al. Regulation of mir-196b by MLL and its overexpression by MLL fusions contributes to immortalization. Blood 2009; 113: 3314-22.

6. Identification of a new putative molecular target gene in *MLL/AF4*-rearranged patients through Gene Expression Profile Analysis: the suppressor of cytokine signaling 2 (*SOCS-2*)

6.1 Introduction

Several studies (1-3) have suggested that MLL-fusion genes could re-activate a stem cell-like gene expression program in target cells. When performing a “class comparison” analysis between the *MLL/AF4* patients and the normal BM control samples, we identified several up- and down-regulated genes and among these genes we searched for those known to be up- or down-regulated in HSCs (4-6). One of the most up-regulated stem-like genes in *MLL/AF4* patients was *SOCS-2* (suppressor of cytokine signaling 2), regardless of *HOXA* genes expression levels.

SOCS-2 belongs to a group of proteins identified in 1997 (7) including eight members: (cytockine-inducible SH2 containing protein) CIS and SOCS1 through SOCS7. All these proteins are characterized by a central SH2 domain, a conserved C-terminal domain named SOCS box and a vatriable N-terminal domain (8). A small kinase inhibitor domain (KIR) is present in the N-terminal region of SOCS-1 and SOCS-3.

SOCS proteins are implicated in the tight regulation of the *JAK-STAT* pathway (9) and this pathway is involved in the activation of several signal cascades such as those activating *RAS, PI3K* and in the transcriptional activation of numerous target genes (10).

SOCS mRNA and protein levels are constitutively low in un-stimulated cells but their expression is rapidly induced upon cytokine stimulation, suggested to create a negative feedback loop (8). In particular, *SOCS-2* is able to regulate the JAK/STAT signaling pathway in several systems *in vitro* and has been associated with the regulation of growth hormone (GH), insulin like growth factor 1(IGF-1), prolactin (PRL), interleukin 2 (IL-2), interleukin 3 (IL-3), erythropoietin (EPO), leukemia inhibitor factor (LIF), epidermal growth factor (EGF), leptin and IFN-α-dependent signaling pathways, either positive or negative (8).

The mechanisms regulating SOCS-2 functions have been only partially revealed: although it can inhibit GH (11) and PRL responses (12), it can also potentiate the signals induced by cytokines. SOCS-2, indeed, is believed to have a dual role, since low concentration of SOCS-2 inhibits GH action whereas higher concentration of SOCS-2 enhanced GH signaling (13). Accordingly, mice over-expressing and mice lacking SOCS-2 display a gigantism phenotype (14). Moreover, the work by Tannahill et al (2005) (15) supports the hypothesis that expression of SOCS-2 can enhance cytokine responses (16-18), most likely through degradation of other SOCS proteins, especially CIS and SOCS-3 (13).

Research has also been carried out on the role of *SOCS-2* in the regulation of hematopoiesis. Despite data suggesting that SOCS-2 transgenic mice do not show leukemia development or

maturation arrest in hematopoietic cells of any lineage (19), high levels of *SOCS-2* mRNA have been found in patients with chronic myeloid leukemia during the onset of blast crisis (20) and *SOCS-2* has been identified as a downstream target of the *BCR/ABL1* fusion gene (21).

SOCS-2 is involved in osteoblast differentiation as well. Bone mineral density analysis in SOCS-2 deficient mice has revealed that, the absence of SOCS-2 induces a reduction in the trabecular and cortical volume of the bone mineral density (BMD). *SOCS-2* induces also the differentiation of mesenchymal cells into myoblasts or osteoblasts through the up-regulation of JunB; this data is particularly intriguing because of the known relationship between hematopoiesis and osteogenesis (22) and prompts for further investigation into the role of *SOCS-2* in the hematopoiesis and development of leukemia.

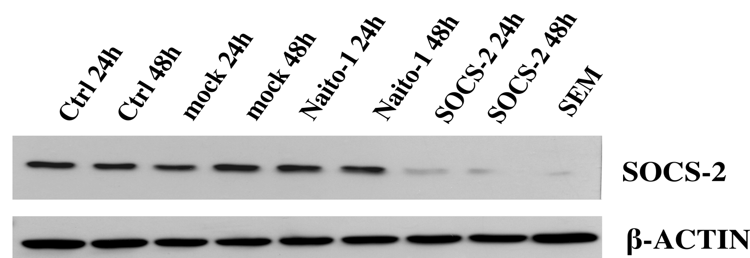
Gene expression analysis of ALL patients with *MLL/AF4* revealed that these patients are characterized by a high expression of *SOCS-2* mRNA compared to normal bone marrow controls. We used RNA interference (RNAi) to specifically inhibit *SOCS-2* expression in t(4;11) cells and we demonstrated that depletion of SOCS-2 induces apoptosis in the silenced RS4;11 cells through the over-expression of TP53 and BAX.

6.2 RS4;11: *in vitro* model for silencing experiments

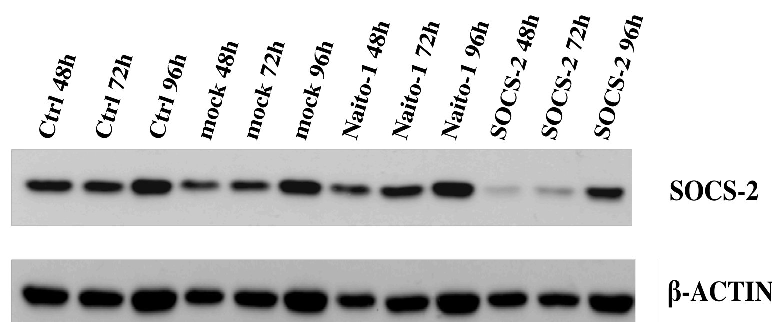
We analyzed the gene expression profile of 3 MLL-rearranged cell lines (RS4;11, SEM and MV4;11) in order to identify the most suitable ones for the analysis. The RS4;11 cell line resulted to express *SOCS-2* at levels similar to that in patients and was therefore used in further experiments. Remarkably, the SEM cell line expressed *SOCS-2* mRNA at a very low level and was used as a control for SOCS-2 antibody during the Western Blot analysis.

6.3 Efficiency of SOCS-2 silencing

A chimera small interfering RNA (siRNA) designed according to the most recent RNAi technology in order to avoid the siRNA off-target effects (including microRNA-based influence) and interferon induction (23) was used to silence *SOCS-2*. Two separate time course experiments were performed; in both cases, $3*10^6$ cells were transefcted with 1,75 μM siRNA anti-SOCS-2 and with 1,75 μM siRNA control (Naito-1). As transfection control (mock) $3*10^6$ cells were treated only with the transfection solution. As shown in **Fig. 13 A-B,** the SOCS-2 amount reached its minimum at 48h after siRNA transfection and recovered to normal level at day 4. The silencing efficiency of SOCS-2 in the RS4;11 was calculated after normalization with β-ACTIN and ranged between 80 and 85% at 48h post-transfection.



A



B

Figure 13. Time course of SOCS-2 depletion. Total proteins were isolated at the indicated time points after transfection with 1,75 μM siRNA anti-SOCS-2 and with 1,75 μM siRNA control (Naito-1). Total proteins were also isolated from cells transfected without siRNA treatment (mock); 15 μg of total protein have been used for the immunoblot analysis. β-ACTIN was used as loading control and for normalization. The SEM cell line was used as antibody negative control. **A**: immunoblot analysis of the first time course analysis of SOCS-2 depletion at 24h and 48h post single transfection; **B**: immunoblot analysis of the second time course analysis of SOCS-2 silencing at 48h, 72h and 96h post single transfection.

6.4 Apoptosis induction in RS4;11 through SOCS-2 silencing

RS4;11 cells were transfected twice with both siRNAs in order to obtain a nearly continuous depletion of SOCS-2 protein; a similar approach was used by Thomas et al (2005) (24) for the targeting of MLL/AF4 fusion protein with transient short interfering RNA. The immunoblot analysis revealed that SOCS-2 was significantly down-regulated (**Fig. 14 A-B**) compared with SOCS-2 expression levels in cells transfected with the Naito-1 control siRNA; the most significantly depletion of SOCS-2 was achieved at 48h post-second transfection.

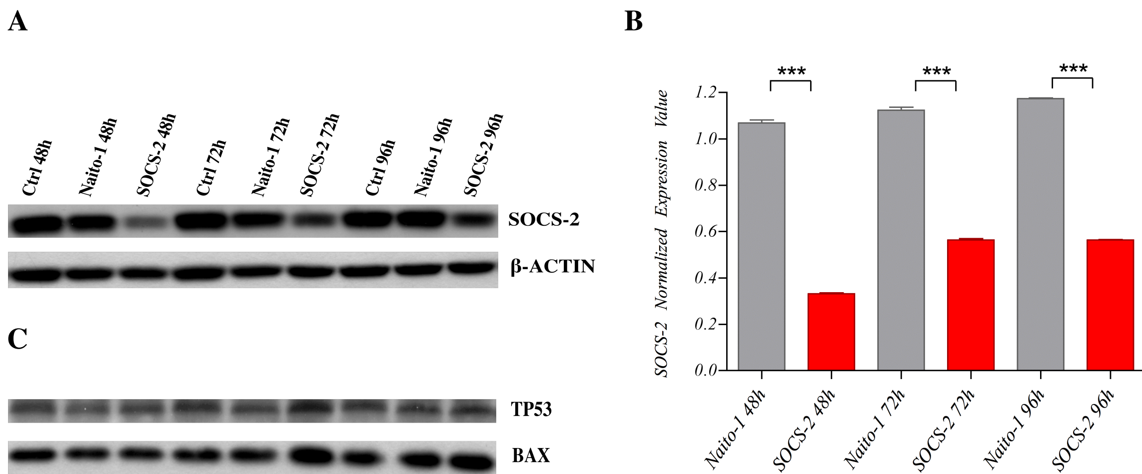


Figure 14. Effects of SOCS-2 silencing in RS4;11. Immunoblot analysis of SOCS-2, TP53 and BAX after SOCS-2 depletion is shown in A and C. The depletion of SOCS-2 determines the increased expression of TP53 and BAX at the reported time points after the second transfection (C). The SOCS-2 depletion is highly significant at every considered time points. The expression value of SOCS-2 normalized to the respective β–ACTIN is reported in B. All the p-values have been calculated using the two-sample Welch t-statistics (unequal variance) in PRISM 4 Version 4.0 (GraphPad Software). ***: p-value ≤ 0,001.

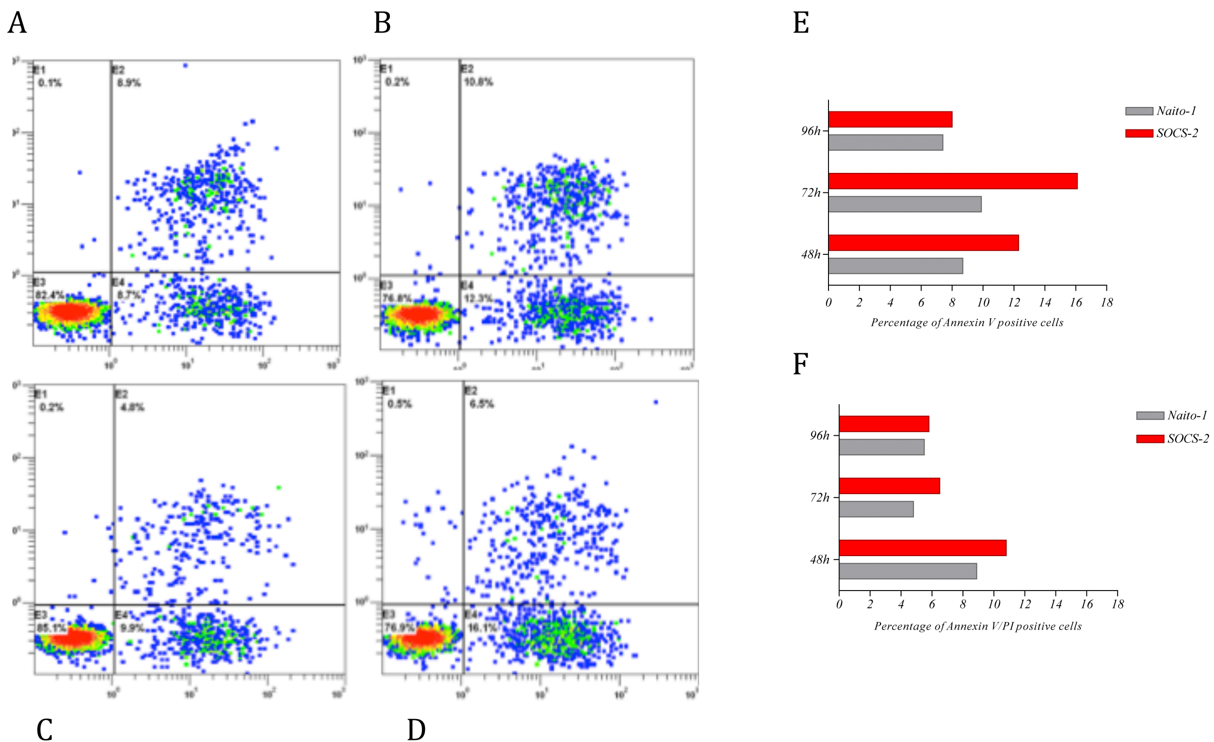


Figure 15. SOCS-2 depletion induces apoptosis in RS4;11 cell line. The induction of apoptosis after SOCS-2 depletion in RS4;11 cell is reported. Plots A-D show the annexin V/PI acquisition data after RS4;11 transfection with the Naito-1 control siRNA at 48h (A) and 72h (C) post second transfection. The annexin V/PI acquisition data of RS4;11 treated with the SOCS-2 siRNA is shown in plot B (48h) and D (72h). X axis: Annexin V; Y axis: PI. The graph bars depicted in figure E and F represent the percentage of annexin V (E) and annnexin V/PI (F) positive cells at the indicated time points post second transduction.

At 48h, 72h and 96h post second-trasfection, $5*10^5$ cells were collected and labeled with annexin V and propidium iodide (PI). SOCS-2 depletion induced an increase of annexin-V and annexin V/PI-positive cells at 48h and 72h post-transfection (**Fig. 15**), suggesting a correlation between SOCS-2 depletion and rate of apoptosis.

According to this data, we sought to investigate if the increased apoptotic rate could be associated with a major expression of TP53 which is known to promote apoptosis through the up-regulation of several pro-apoptotic genes (i.e. *BAX*) and to down-regulate anti-apoptotic genes such as those belonging to the *BCL-2* family. We found that TP53 was significantly up-regulated in SOCS-2 silenced RS4;11 cells at 48h and 72h post transfection (p-value ≤ 0.001) (**Fig. 14 C and Fig. 16 A**), whereas BAX was significantly up-regulated in SOCS-2 silenced cells at every time point (**Fig. 14 C and Fig. 16 B**).

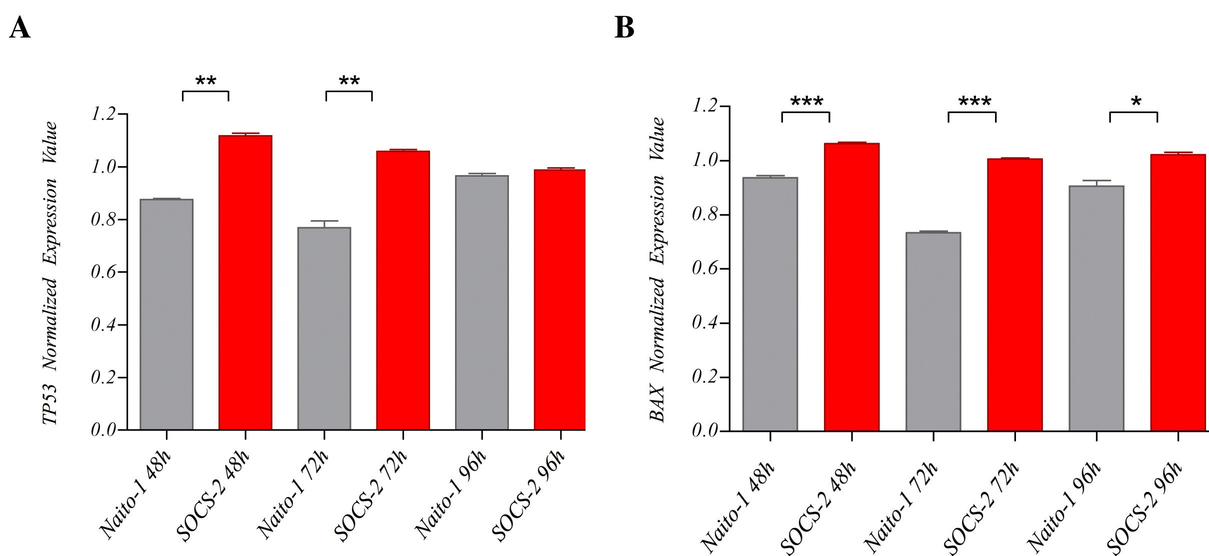


Figure 16. Induction of TP53 and BAX after SOCS-2 silencing. The expression values of TP53 and BAX after the second transfection are shown. The graph bars indicate the normalized expression values of the proteins TP53 (A) and BAX (B) as calculated at the indicated time points after the second transfection. Each protein expression value has been normalized to the corresponding β–ACTIN value. All the p-values have been calculated using the two-sample Welch t-statistics (unequal variance) in PRISM 4 Version 4.0 (GraphPad Software). ***: p-value ≤ 0,001; **: 0,001< p< 0,01; *: 0,01 <p< 0,05; NS: not significant.

Thus, this data suggests that SOCS-2 depletion may contribute to the activation of a TP53-mediated pro-apoptotic program in the RS4;11 cell line; we also analyzed the gene expression profiles of samples belonging to different ALLs subclasses. All but T-ALLs express high levels of *SOCS-2* mRNA signifying that SOCS-2 up-regulation could be a common mechanism in apoptosis resistance.

6.5 Materials and Methods

Cell lines

The RS4;11, the SEM and MV4;11 cell lines were obtain from DSMZ (Braunschweig, Germany); both cell lines were maintained in RPMI 1640 (Biochrom AG, Berlin, Germany) with 10% FCS, penicillin (100 units/ml) and streptomycin (100 μg/ml) and maintained at 37°C in a humidified atmosphere with 5% CO_2.

RNA isolation and Microarray analysis

Total RNA was isolated using TRIzol reagent (Invitrogen) followed by RNA purification on RNeasy columns (Qiagen). The RNA quality and concentration were assessed on the Agilent 2100 Bioanalyzer (Agilent Technologies) and the NanoDrop ND-1000 spectrofotometer (NanoDrop Technologies), respectively. For microarray experiments *in vitro* transcription, hybridization and biotin labelling were performed according to Affymetrix One Cycle Target Labeling protocol. GeneChip Human Genome U133 Plus 2.0 arrays (Affymetrix) were used. Microarray data (.CEL files) were generated using default Affymetrix microarray analyisis parameters (GCOS 1.2 software).

siRNA treatment

The chimeric siRNA for *SOCS-2* and the control siRNA Naito-1 were obtained by Abnova. For the silencing experiment $3*10^6$ cells were transfected (AMAXA device, solution V, program L-017) with 1,75 μM siRNAs; 48h post first transfection cells have been collected and transfected again with 1,75 μM siRNAs. Cells have been collected for the following experiments at 48h, 72h and 96h post second transfection.

Western Blot

20 μg from total protein fraction (Buffer-Biosource by Invitrogen-Gibco) obtained from the cells were used to perform protein analysis. Protein concentration was determined using the BCA method (Pierce). Samples were ran on a 15% SDS-plyacrylamide gel and then transferred to 0,2 μm polyvylidene difluoride membrane (GE-Healthcare) for immunodetection with specific antibodies followed by horseradish peroxidase-conjugated goat anti-rabbit or mouse IgG (Upstate). Anti-SOCS-2, anti TP53 and anti BAX antibodies were obtained by Cell Signaling (catalog# 2779, 9282 and 2772 respectively) and anti β–ACTIN antibody was obtained by Sigma Aldrich. The specific bands of the target proteins were visualized by enhanced chemoluminescence (ECL advance) according to the manufacturer’s instructions (GE-Healthcare).

Apoptosis assay

Apoptosis was examined with a human Annexin V kit (ROCHE) according to manufacturer's instructions; briefly $5*10^5$ cells were washed with Hanks' salt solution (Biochrom AG) at the indicated time points after transfection followed by incubation in the presence of annexinV/PI solition for 15 minutes at room temperature in the dark. The cells were then washed again with Hanks' salt solution (Biochrom AG) and immediately analyzed by flow cytometry using a flow cytometer (Beckman Couter Cytomics FC 500).

6.6 References

1. Cozzio A, Passegue E, Ayton PM, Karsunky H, Cleary ML, Weissman IL. Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors. Genes Dev 2003; 17: 3029-35.

2. Krivtsov AV, Twomey D, Feng Z, Stubbs MC, Wang Y, Faber J, et al. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature 2006; 442: 818-22.

3. Faber J, Armstrong SA. Mixed lineage leukemia translocations and a leukemia stem cell program. Cancer Res 2007; 67: 8425-8.

4. Ivanova NB, Dimos JT, Schaniel C, Hackney JA, Moore KA, Lemischka IR. A stem cell molecular signature. Science 2002; 298: 601-4.

5. Georgantas RW,3rd, Tanadve V, Malehorn M, Heimfeld S, Chen C, Carr L, et al. Microarray and serial analysis of gene expression analyses identify known and novel transcripts overexpressed in hematopoietic stem cells. Cancer Res 2004; 64: 4434-41.

6. Toren A, Bielorai B, Jacob-Hirsch J, Fisher T, Kreiser D, Moran O, et al. CD133-positive hematopoietic stem cell "stemness" genes contain many genes mutated or abnormally expressed in leukemia. Stem Cells 2005; 23: 1142-53.

7. Starr R, Willson TA, Viney EM, Murray LJ, Rayner JR, Jenkins BJ, et al. A family of cytokine-inducible inhibitors of signalling. Nature 1997; 387: 917-21.

8. Rico-Bautista E, Flores-Morales A, Fernandez-Perez L. Suppressor of cytokine signaling (SOCS) 2, a protein with multiple functions. Cytokine Growth Factor Rev 2006; 17: 431-9.

9. Krebs DL, Hilton DJ. SOCS proteins: negative regulators of cytokine signaling. Stem Cells 2001; 19: 378-87.

10. Darnell JE,Jr. STATs and gene regulation. Science 1997; 277: 1630-5.

11. Greenhalgh CJ, Alexander WS. Suppressors of cytokine signalling and regulation of growth hormone action. Growth Horm IGF Res 2004; 14: 200-6.

12. Harris J, Stanford PM, Sutherland K, Oakes SR, Naylor MJ, Robertson FG, et al. Socs2 and elf5 mediate prolactin-induced mammary gland development. Mol Endocrinol 2006; 20: 1177-87.

13. Greenhalgh CJ, Metcalf D, Thaus AL, Corbin JE, Uren R, Morgan PO, et al. Biological evidence that SOCS-2 can act either as an enhancer or suppressor of growth hormone signaling. J Biol Chem 2002; 277: 40181-4.

14. Metcalf D, Greenhalgh CJ, Viney E, Willson TA, Starr R, Nicola NA, et al. Gigantism in mice lacking suppressor of cytokine signalling-2. Nature 2000; 405: 1069-73.

15. Tannahill GM, Elliott J, Barry AC, Hibbert L, Cacalano NA, Johnston JA. SOCS2 can enhance interleukin-2 (IL-2) and IL-3 signaling by accelerating SOCS3 degradation. Mol Cell Biol 2005; 25: 9115-26.

16. Favre H, Benhamou A, Finidori J, Kelly PA, Edery M. Dual effects of suppressor of cytokine signaling (SOCS-2) on growth hormone signal transduction. FEBS Lett 1999; 453: 63-6.

17. Dif F, Saunier E, Demeneix B, Kelly PA, Edery M. Cytokine-inducible SH2-containing protein suppresses PRL signaling by binding the PRL receptor. Endocrinology 2001; 142: 5286-93.

18. Pezet A, Favre H, Kelly PA, Edery M. Inhibition and restoration of prolactin signal transduction by suppressors of cytokine signaling. J Biol Chem 1999; 274: 24497-502.

19. Greenhalgh CJ, Bertolino P, Asa SL, Metcalf D, Corbin JE, Adams TE, et al. Growth enhancement in suppressor of cytokine signaling 2 (SOCS-2)-deficient mice is dependent on signal transducer and activator of transcription 5b (STAT5b). Mol Endocrinol 2002; 16: 1394-406.

20. Schultheis B, Carapeti-Marootian M, Hochhaus A, Weisser A, Goldman JM, Melo JV. Overexpression of SOCS-2 in advanced stages of chronic myeloid leukemia: possible inadequacy of a negative feedback mechanism. Blood 2002; 99: 1766-75.

21. Hakansson P, Nilsson B, Andersson A, Lassen C, Gullberg U, Fioretos T. Gene expression analysis of BCR/ABL1-dependent transcriptional response reveals enrichment for genes involved in negative feedback regulation. Genes Chromosomes Cancer 2008; 47: 267-75.

22. Patt HM, Maloney MA. Bone formation and resorption as a requirement for marrow development. Proc Soc Exp Biol Med 1972; 140: 205-7.

23. Ui-Tei K, Naito Y, Zenno S, Nishi K, Yamato K, Takahashi F, et al. Functional dissection of siRNA sequence by systematic DNA substitution: modified siRNA with a DNA seed arm is a

powerful tool for mammalian gene silencing with significantly reduced off-target effect. Nucleic Acids Res 2008; 36: 2136-51.

24. Thomas M, Gessner A, Vornlocher HP, Hadwiger P, Greil J, Heidenreich O. Targeting MLL-AF4 with short interfering RNAs inhibits clonogenicity and engraftment of t(4;11)-positive human leukemic cells. Blood 2005; 106: 3559-66.

7. Discussion

The introduction of microarray technology has been a major step towards the comprehensive biologic characterization of different hematological diseases.

The increasing amount of data showing the potential of gene expression profile analysis in the classification of leukemia subclasses suggests that microarray analysis should be introduced into the routinely established "gold standard" techniques for the diagnosis of leukemia. This concept is also supported by works demonstrating that this technique is robust and characterized by a high inter-platform comparability. Our work focusing on the analysis of the impact of different RNA preparation procedures on gene expression data shows that different sample preparation procedures do not impair samples classification and that the underlying biological characteristics of the pediatric acute leukemia classes largely exceed the variations between different RNA preparation protocols. However, inter-laboratory reproducibility and gene expression data mining are still two of the major concerns in the application of microarray technology (1); looking to a future application of microarray technology in the diagnosis of leukemia, this issue should be overtaken, as shown by the MILE study, using standardized protocols, instruments and software analysis tools.

Although microarray analysis represents a powerful tool for the classification and prediction of nearly all leukemia subgroups, it is currently not able to predict some less well-defined leukemia subclasses (i.e. the B-others); this finding may lead to the interpretation that these kinds of leukemia do not represent a homogeneous subgroup but rather a heterogeneous assortment of different leukemia subtypes and suggests that more platforms such as microRNA, CGH and SNPs arrays should be used in to reveal peculiar features of these kinds of leukemia.

Our studies on JMML and *MLL/AF4* BCPs ALL clearly confirmed that microarray technology can be applied to predict patients outcome and to reveal new subgroups of patients among an already defined disease entity.

The study on JMML patients can be considered a remarkable example of the microarray capability to identify two gene expression signatures underlying different clinical outcomes. For example, the 10-year probability of survival after diagnosis for JMML patients classified as AML-like and non AML-like is significantly different (7% and 74%, respectively $P = 0.005$).

Although our study on *MLL/AF4* the B-cell precursors (BCPs) ALL does not identify two subgroups characterized by different outcomes, it has allowed us to divide the homogeneous BCPs ALL class with t(4;11) into two distinct subgroups and to identify a group of patients in which leukemia arises without the canonical contribution of *HOXA* genes over-expression. This is a striking finding as *HOXA* genes de-regulation is commonly considered a key feature of MLL-rearranged AML and ALL leukemia. The reason why the *HOXA* genes cluster is not expressed in these patients is still unclear; further experiments on the *HOXA* promoter sequence and chromatin immunoprecitipation (CHIP) experiments should clarify this issue. Remarkably, the presence of two different gene

signature among BCP-ALL *MLL/AF4* patients has been recently confirmed by an independent study on 29 *MLL/AF4* pediatric infant patients; furthermore, it seems that this subdivision has also a prognostic value, as patients lacking *HOXA* expression are at extremely high risk of disease relapse (2).

Patients with *MLL/AF4* ALL are characterized by a very poor outcome and resistance to conventional treatment such as prednisone. *MLL/AF4* ALL patients are characterized by the up-regulation of *SOCS-2* and the SOCS-2 depletion in RS4;11 cells determines the initial activation of a pro-apoptotic cascade through over-expression of TP53. We have analyzed also the *SOCS-2* expression profile in patients belonging to several different ALL and have found that *SOCS-2* is up-regulated in all but T-lineage. This finding may suggest that *SOCS-2* up-regulation could be a common mechanism to prevent induction of apoptosis in ALLs. Silencing experiments in different ALL cell lines should be performed in order to confirm this concept.

The first work by Golub and colleagues applying microarray analysis in leukemia highlighted the potential of gene expression profile analysis for the identification of specific gene expression signatures distinguishing AML and ALL subgroups. Moreover, this study, demonstrates that microarray technology could be applied not only for descriptive studies (3, 4) but also for a more comprehensive and systematic approach to cancer and leukemia classification based on the simultaneous expression monitoring of thousands of genes. Furthermore, this work points to microarray technology as a new promising tool for the identification of the abnormally activated biological processes involved in the development of leukemia. Nowadays, these concepts have been confirmed and expanded by a huge amount of reports and microarray technology is constantly applied as a frontline tool not only for the biological characterization of different leukemia subgroups, but also for outcome prediction and for the identification of new leukemia entities.

Nevertheless, microarray technology does not seem to be sufficient for the complete characterization of leukemia subgroups; a clear example is provided by the low accuracy observed during the MILE study in the classification of AML with 11q23 rearrangements, hyperdyploid katyotype and the MDS or leukemic patients characterized by the absence of known molecular markers or translocations (i.e. B-others). Moreover, whereas gene expression microarray technology provides information on transcriptionally active genes in a context-dependent moment of the cells, it does not contain any data concerning the genes' regulatory states, gene sequence polymorphisms or chromosomal abnormalities. Furthermore, the gene expression signatures obtained through microarray analysis are not simple to interpret with respect to the biology of the underlying disease. Thus, microarray data should be integrated with information deriving from different platforms such as single nucleotide polymorphism (SNP) arrays, array-based comparative genomic hybridization (CGH) and epigenomic arrays. Certainly, the elaboration of such a huge amount of data is a great future challenge. However, it is likely that the combination of gene expression profile with complementing technologies will provide new opportunities to answer questions that cannot be resolved by GEP alone.

7.1 References

1. Ioannidis JP, Allison DB, Ball CA, Coulibaly I, Cui X, Culhane AC, et al. Repeatability of published microarray gene expression analyses. Nat Genet 2009; 41: 149-55.

2. Stam RW, Schneider P, Hagelstein JA, van der Linden MH, Stumpel DJ, de Menezes RX, et al. Gene expression profiling-based dissection of MLL translocated and MLL germline acute lymphoblastic leukemia in infants. Blood 2009;.

3. Khan J, Simon R, Bittner M, Chen Y, Leighton SB, Pohida T, et al. Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. Cancer Res 1998; 58: 5009-13.

4. DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, et al. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat Genet 1996; 14: 457-60.

Γένοιο οἷος εἶ.

Pindaro

Acknowledgements

Papà Pietro, mamma Bruna, Cristina
Enrico, Chiara, Martino, Nicola, Alessia, Francesco, Giovanni, Miki

Giuseppe Basso, Geertruy te Kronnie

Silvia, Elena, Marta, Benedetta, Marco, Andrea,
Beppe, Marti, Manu, Katia, Silvia D., Chiara, Elena S., Sanja, France,
Alberto, Eva, Alessandra, Ilaria, Maddy, Giulia, Manara, Alessandra, Barbara, Grazia,
Anna Maria, Anna, Barbara, Elena F, Massimo

Appendix 1

Appendix 2

Appendix 3