UNIVERSIDADE DE LISBOA
FACULDADE DE MEDICINA DE LISBOA

The role of reciprocal regulation between TAL1 and miRNA expression in T-cell leukemia progression

Nádia Grangeon Cavaleiro de Sousa Correia

Orientador: Professor Doutor João Pedro Taborda Barata
Co-Orientador: Professor Doutor Francisco J. Enguita

Doutoramento em Ciências Biomédicas

Especialidade em Biologia Celular e Molecular

A impressão desta dissertação foi aprovada pelo Conselho Científico da Faculdade de Medicina de Lisboa na reunião de 28 de Outubro de 2014.

Preface

This thesis describes the research work under the scope of my PhD project developed between July of 2009 and July of 2014 at the *Instituto de Medicina Molecular* (Lisbon, Portugal) under the supervision of João T. Barata, PhD. During this period, part of the research work was also carried at the Cell Division and Cancer Group at CNIO (Madrid, Spain) under the supervision of Prof. Marcos Malumbres and at the Center for Medical Genetics, (Ghent University, Belgium) under the supervision of Prof. Frank Speleman.

The results presented in this thesis are the result of my own research work and it is clearly acknowledged in the text whenever results or reagents produced by others were utilized. I was financially supported by a scholarship from Fundação para a Ciência e Tecnologia (FCT, SFRH/BD/47575/2008) and the research project was financed by Liga Portuguesa Contra o Cancro (Terry Fox Award) and by FCT (project PTDC/BIM-ONC/1548/2012).

The work described here originated the following publication in a peer-reviewed international journal, in which I am the first author (Appendix I):

Correia NC, Durinck K, Leite AP, Ongenaert M, Rondou P, Speleman F, Enguita FJ, Barata JT. *Novel TAL1 targets beyond protein-coding genes: identification of TAL1-regulated microRNAs in T-cell acute lymphoblastic leukemia*. **Leukemia**. 2013 Jul;27(7):1603-6.

I also participated in other studies during the period of my PhD, which originated the following publications in peer-reviewed international journals:

Correia NC*, Gírio A*, Antunes I*, Martins LR*, Barata JT. *The multiple layers of non-genetic regulation of PTEN tumour suppressor*. **Eur J Cancer**. 2014 Jan;50(1):216-25. * co-first authors

In this review article I wrote the part concerning the transcriptional regulation, epigenetic silencing and post-transcriptional regulation of PTEN and contributed to the organization and global revision of the article.

Ribeiro D, Zenatti PP, Li W, Zuurbier L., Silva MC, Paganin M, Tritapoe J, Hixon JA, Silveira AB, Cardoso BA, Sarmento LM, Correia N, Toribio ML, Kobarg J, Horstmann M, Pieters R, Brandalise SR, Ferrando AA, Meijerink JP, Durum SK, Yunes JA, Barata JT. *Oncogenic IL7R gain-of-function mutations in childhood T-cell acute lymphoblastic leukemia*. **Nature Genetics** 2011, 43(10):932-9

In this manuscript I contributed with the analysis of cell viability in the presence of JAK-STAT pathway inhibitors; analysis of cell proliferation in absence of growth factors; analysis of cell transfection with siRNAs against JAK1, JAK3 and IL-2Rγ; and preparation of samples for immunoblot analysis upon incubation in the presence of JAK-STAT pathway inhibitors or in the absence of growth factors.

Acknowledgements

Dedico esta tese aos meus avós Elsa e Luís, a quem dediquei e dedicarei todos os sucessos da minha vida. Obrigada por sempre me terem acolhido e ajudado como a uma filha. A vós vos devo praticamente tudo o que sou.

Esta foi talvez a secção mais difícil de escrever. Primeiro porque a minha manifesta gratidão e a lista das pessoas merecedoras de tal é tão extensa que daria uma tese em si. Mas principalmente porque ao escrever estas palavras sinto que fecho um capítulo e me despeço de uma época tão frutífera, onde pude desenvolver pessoal e profissionalmente as minhas capacidades no melhor dos ambientes, rodeada das melhores pessoas. Portanto, não é sem grande emoção que o faço.

É desta forma que o meu principal agradecimento vai para o melhor chefe do mundo. Ter-te tido como chefe, João, elevou de tal modo a fasquia que temo nunca mais a conseguir igualar. Obrigada por tudo, pelos conselhos, pelas reprimendas, pelo optimismo, pelo orgulho, pela dedicação, rectidão e moralidade com que conduzes o teu trabalho, pela inspiração, pela confiança, pelo altruísmo. Obrigada pelos momentos de riso, pelos momentos de borga, pelos bolos e pelas correcções. És um verdadeiro orientador na verdadeira ascensão da palavra. Obrigada por me dares força para empreender um projecto arriscado e pela experiência que me proporcionou e obrigada por seres um exemplo a seguir. Obrigada também pelo grupo que formaste, pelos elementos que escolheste e por motivares e proporcionares o excelente ambiente pessoal e profissional que se sente na nossa unidade.

Também me considero sortudo quando escolhi o meu co-orientador, Paco. Obrigada pelos conselhos e disponibilidade. Obrigada por acreditares em mim e me fazeres sentir sempre capaz a empreender mais um desafio. Obrigada pelo input técnico e pela partilha de conhecimento.

I would also like to acknowledge to our collaborators, professor Marcos Malumbres, Maria Toríbio and Frank Speleman and all the member of their groups for receiving me in

their labs, for making me feel part of their team and for all the ideas and discussion about my work.

Aos meus colegas de laboratório, passados e presentes (Ana, Bruno, Leila, Inês, Leonor, Alice, Vanda, Daniel, Joana, Mariana, Milene, Tavanez, Isabel, Margarida, Rita) muito obrigada, vocês proporcionaram uma excelente experiência. À sua maneira cada um contribuiu com ideias, soluções, conselhos, ajuda, animação, diversão e amizade. Obrigada pelo ambiente de interajuda e por contribuírem para o enorme prazer em ir trabalhar todos os dias. Obrigada também à UBH e seus elementos com que partilhamos o laboratório e os bons momentos. Um agradecimento com especial carinho ao Bruno e à Ana Gírio que há sete anos atrás me receberam no seu projecto e me ensinaram os primeiros passos. Um agradecimento especial também à Alice, Vanda, Leonor e Rita pela ajuda que em deram na execução deste trabalho, pela companhia no trabalho "fora de horas", pelas discussões científicas, pela galhofa geral e pela amizade. Finda esta etapa, já não haverá nada fora do sítio no lab e vocês vão sentir a minha falta! Um grande obrigada à Rita por ter aceite participar nas últimas fases deste trabalho e pela extensa revisão desta tese.

Gostava também de agradecer aos colegas e colaboradores do IMM e das suas "facilities" que de uma maneira ou outra contribuíram para este trabalho, com ajuda técnica, conselhos, ou aquele reagente que fazia mesmo falta.

Um trabalho desta extensão e dedicação não se faz sem um forte apoio pessoal e emocional. Nos bastidores estão os leigos que desempenham um papel tão fundamental quanto invisível. Agradeço aos elementos mais próximos da minha família pelo apoio prestado. Agradeço especialmente aos meus tios Miguel, Luís e Maria José. Obrigada por terem sido mais do que tios. Obrigada também pelos meus priminhos lindos, pelos risos e brincadeiras. Obrigada pelo apoio durante o meu percurso académico e sem o qual não teria certamente ultrapassado muitos obstáculos pessoais.

Agradeço também ao meu pai. Creio que de ti herdei esta insaciável curiosidade, esta sede de conhecimento. Contigo também apreendi a conduta profissional, a rectidão de carácter. Obrigada a ti e à Isabel pelos meus irmãos e pela oportunidade de crescer com eles.

Um agradecimento especial aos meus sogros, Ilda e Jorge, por me terem acolhido como filha e por todo o apoio prestado. E claro, pelo fantástico filho que geraram!

Aos meus amigos de infância, da escola, dos escoteiros, dos intercâmbios, da faculdade, das danças e de Gent, agradeço os bons momentos passados, o espairecer, o companheirismo, os jantares (do chefe!), os scones, as férias, as viagens, as festas, os aniversários, as visitas enquanto estava fora, as mensagens de apoio, as sessões de skipe… vocês são a família que eu escolhi. Um forte obrigada às minhas princesas pelas visitas enquanto expatriada e pelos "scones" semanais. Um grande obrigada à comunidade de expatriados de Gent, portugueses e estrangeiros, que me fez sentir o tempo lá passado como uma das melhores experiências da minha vida. A special acknowledgement to my neighbors Amir and Niloofar and also to my new friends Maria e Filipa for being so kind, supportive and for enlarging my horizons.

Finalmente, à pessoa que assumiu a maior importância na minha vida na última década, o meu xuxu, agradeço tudo e o mais que possa vir. André este trabalho também é teu. Obrigada pelo apoio e dedicação, pelo carinho, pela ajuda, pelo companheirismo, pela motivação, pela segurança, pela amizade e amor. Devias ganhar o grau por extensão, tu que já és doutorado em paciência! Não teria conseguido sem ti. Só se pode crescer, só se tem forças para lutar e para enfrentar os nossos medos e fraquezas quando estamos felizes, quando nos sentimos amados e realizados. És o meu pilar, a minha rede de segurança e o meu trampolim. Contigo ao meu lado poderei conquistar o mundo, sem ter de o fazer para me sentir feliz. Tu, eu, o amor e uma cabana, para sempre… Muito muito obrigada!

Não encontro palavras suficientes para exprimir a minha gratidão e todos os elogios me parecem curtos. A todos os que tornaram este trabalho possível, fica expresso o meu sincero.

Obrigada!

Resumo

A Leucemia Linfoblástica Aguda de Células T (LLA-T) é um cancro pediátrico agressivo que resulta da expansão clonal de células progenitoras de linfócitos T, cuja diferenciação se encontra bloqueada em diferentes estádios de desenvolvimento. Embora os regimes quimioterápicos actuais sejam bastante eficazes, culminando numa taxa de cura de cerca de 80% em crianças, existe ainda um número significativo de doentes que recidivam. Para além disso, os regimes intensivos de quimioterapia estão normalmente associados a efeitos secundários consideráveis a médio e longo prazo. Por esta razão, urge desenvolver terapias com maior especificidade para as células leucémicas, reduzindo toxicidade e consequentes efeitos secundários. Para tal é necessário melhorar o nosso conhecimento sobre as causas, fisiologia e regulação da LLA-T, em particular através da identificação de alvos moleculares e vias fundamentais para a progressão da doença.

TAL1 é um factor de transcrição essencial para a função das células estaminais hematopoiéticas. No entanto, a expressão de TAL1 decresce rapidamente nos progenitores destinados a originar células T. Assim, quando expresso de forma aberrante nos precursores de células T, TAL1 desempenha claramente uma função oncogénica. De facto, a sobre-expressão de TAL1 ocorre em mais de 60% dos casos de LLA-T.

Os microRNAs (miRNAs) são espécies de RNAs não-codificantes de pequena dimensão que regulam negativamente a expressão genes codificantes de proteínas. A sua acção é mediada pela inibição da tradução ou aumento da degradação dos RNAs mensageiros, resultando num decréscimo da proteína alvo. O envolvimento dos miRNAs na regulação da carcinogénese é largamente reconhecido, nomeadamente devido à sua capacidade de inibição da expressão de oncogenes ou supressores tumorais, resultando na prevenção ou promoção, respectivamente, do desenvolvimento tumoral. Tal como para outros tipos de cancro, vários miRNAs foram identificados e a sua função descrita como importante em LLA-T.

O trabalho desenvolvido no âmbito desta tese tem como objectivo caracterizar a rede de interacções entre TAL1 e microRNAs. Desta forma, esperamos contribuir para a

compreensão dos mecanismos pelos quais TAL1 promove o desenvolvimento de leucemia. Através da identificação de novos genes promotores dos efeitos de TAL1 ou envolvidos na regulação deste oncogene esperamos identificar potenciais alvos moleculares para intervenção terapêutica em LLA-T. Para além disso, pretendemos contribuir para o conhecimento geral da biologia e fisiologia dessa doença.

Na maioria dos casos de LLA-T os mecanismos subjacentes à expressão ectópica de TAL1 estão ainda por elucidar. Uma das questões colocadas nesta tese visa esclarecer se a sobre-expressão de TAL1 em LLA-T pode resultar ou ser potenciada pela desregulação de determinados miRNAs. Esta hipótese implica por outro lado que a diminuição dos níveis de TAL1 durante o normal desenvolvimento de células T resulta de um decréscimo que é, em parte, dependente de microRNAs. De forma a verificar se assim é de facto, analisámos ratinhos com uma deleção condicional da enzima DICER, que consequentemente não expressam miRNAs em precursores T, e verificámos que timócitos de ratinhos deficitários para DICER expressam níveis aumentados do transcrito de *Tal1* em relação a ratinhos controlo. De seguida, realizámos estudos bioinformáticos para prever os miRNAs que se podem ligar ao transcrito humano do gene *TAL1* e compilámos uma lista de possíveis candidatos. Através de um sistema repórter baseado na expressão de luciferase, confirmámos as possíveis interacções miRNA/TAL1. Desta forma seleccionámos os miRNAs com relevância biológica que levaram a uma redução na expressão do repórter em pelo menos 25%: miR-101, miR-520d-5p, miR-140-5p, miR-448 e miR-485-5p. De seguida, validámos a interacção destes miRNAs com o transcrito de *TAL1* através da mutação dos elementos reconhecidos por miRNAs e desta forma confirmámos que os miR-101 e miR-140-5p regulam *TAL1* através dos locais previstos. De forma a avaliar a importância fisiológica da interacção miRNA/TAL1, sobre-expressámos os miRNAs candidatos em linhas celulares de LLA-T. Verificámos que a sobre-expressão de miR-520d, miR-101, miR-140, miR-485 e miR-448 em diferentes linhas celulares resultam na inibição da expressão de TAL1 a nível do transcrito e proteína. Por outro lado, também observámos que a inibição dos miR-520d-5p e miR-101 endógenos é responsável pelo incremento da proteína TAL1. Através da comparação da expressão dos miRNAs miR-101, miR-140-5p, miR-448 e miR-485-5p entre

células de pacientes com LLA-T e células de medula óssea normal, verificámos que a expressão dos microRNAs se encontra diminuída em relação aos controlos. Estes resultados vão ao encontro da nossa hipótese de que um decréscimo de expressão de miRNAs que regulam TAL1 pode ser, pelo menos parcialmente, responsável pela expressão aberrante de TAL1 em LLA-T. O nosso trabalho demonstra também que TAL1 é regulado de forma pos-transcricional por miRNAs.

Apesar de se conhecer há muito a relevância de TAL1 na biologia da LLA-T, pouco se sabe sobre os genes alvo da sua função como regulador transcricional e do contributo destes genes para o desenvolvimento leucémico. Uma outra questão que procurámos resolver no âmbito deste trabalho foi perceber se TAL1 regula a expressão de genes de miRNAs potencialmente relevantes para o desenvolvimento leucémico. Para tal, analisámos o perfil de expressão de miRNAs dependente de TAL1 e identificámos os miRNAs cuja expressão variou de forma significativa após sobre-expressão de TAL1 numa linha celular LLA-T. Os resultados obtidos foram validados através da sobre-expressão ou silenciamento de TAL1 em linhas celulares de LLA-T, confirmando que a expressão dos miR-135a, miR-223 e miR-330-3p é activada por TAL1, enquanto a expressão de miR-146b-5p e miR-545 é inibida. Para além disso verificámos que a o TAL1 se liga a uma região no genoma 3.5000 nucleótidos a jusante do local de início de transcrição do miR-223, indicando que existe uma regulação directa da transcrição deste miRNA por TAL1. O facto de TAL1 activar a expressão de miR-223, que outros demonstraram inibir o supressor tumoral FBXW7, sugere que o TAL1 se encontra no centro de uma rede transcricional que envolve a estabilização de proteínas com importantes funções oncogénicas em LLA-T. Curiosamente, a anotação funcional dos alvos já validados dos miRNAs regulados por TAL1 revelou a prevalência de processos biológicos relacionados com inflamação e cancro. Em conclusão, identificámos um conjunto de genes não-codificantes regulados por TAL1, implicando pela primeira vez genes de miRNAs na rede transcricional a jusante de TAL1 cujo papel pode ser relevante no contexto da hematopoiese normal e do desenvolvimento de LLA-T.

Finalmente direccionámos a nossa investigação para os efeitos moleculares e funcionais da desregulação do miR-146b-5p por TAL1 em LLA-T. Os nossos resultados

sugerem que TAL1 inibe transcricionalmente o miR-146b-5p de forma directa. Para além disso, apesar de não termos encontrado diferenças na viabilidade e proliferação das células, demonstrámos que a inibição do miR-146b-5p tem como consequência uma melhoria das capacidades de migração e invasão das células de LLA-T. Estes resultados sugerem que o miR-146b-5p poderá ter uma função supressora tumoral. Em concordância com esta possibilidade, as células leucémicas de pacientes com LLA-T expressam significativamente menos miR-146b-5p do que as células de dadores saudáveis, indicando que a diminuição da expressão deste miRNA poderá ser benéfica para as células tumorais. Para além disso, observámos que a sobre-expressão de miR-146b-5p em células humanas de LLA-T aumentou a sobrevivência de ratinhos num modelo de xenotransplante de leucemia humana. Desta forma, os nossos resultados sugerem que o miR-146b-5p é um alvo transcricional de TAL1 com relevância funcional. A redução dos níveis de expressão deste miRNA poderá contribuir para a LLA-T através da regulação da mobilidade das células leucémicas e da agressividade da doença.

Em conclusão, o trabalho apresentado nesta tese estabelece a existência de uma relação recíproca entre TAL1 e microRNAs que envolve a regulação epigenética de TAL1 por miRNAs a jusante e a regulação transcricional de miRNAs por TAL1 a montante. Por outras palavras, os nossos estudos contribuem para o esclarecimento dos mecanismos conducentes à expressão anómala de TAL1 em LLA-T bem como dos mecanismos tumorigénicos a jusante de TAL1, nomeadamente no que respeita à identificação de genes de miRNAs alvo de regulação transcricional por TAL1. A pertinência destes estudos para intervenção terapêutica para benefício dos pacientes de LLA-T é um desafio para o futuro, para qual o nosso trabalho serve como ponto de partida.

Abstract

T-cell acute lymphoblastic leukemia (T-ALL) is an aggressive childhood malignancy in which the transformed clone is arrested during T-cell development. Despite significant improvements in treatment outcome that led to rates of cure close to 80% in children, survivors tend to face long term complications and develop serious secondary health problems. Therefore, the current challenge is to develop more efficient therapeutic strategies that target the leukemia cells in a more specific way, diminishing the toxic effects of the treatment. To achieve this, it is essential to improve our knowledge regarding the causes, pathophysiology and regulation of T-ALL, in particular by identifying the molecules and pathways fundamental for leukemia progression.

TAL1 is an important transcription factor for the maintenance of hematopoietic stem cells and regulation of early hematopoiesis, being rapidly down-regulated upon T-cell lineage commitment. Aberrant expression of TAL1 in committed T-cell precursors is associated with leukemia and TAL1 is a well-established T-ALL oncogene, being over-expressed in more than 60% of T-ALL patients.

MicroRNAs (miRNAs) are small, non-coding RNAs that primarily function as endogenous translational repressors of protein-coding genes. They decrease the expression of numerous genes by preventing translation or promoting mRNA degradation. The involvement of miRNAs in the regulation of cancer progression is well-established, namely by down-regulating the expression of key oncogenes or tumor suppressors and thereby preventing or promoting tumorigenesis, respectively. Similar to other cancers, several miRNA genes have been identified and revealed in the context of T-ALL.

In this thesis, we aimed to characterize the network of interactions between TAL1 and microRNA genes. In doing so, we sought to contribute to the understanding of the mechanisms by which TAL1 promotes T-cell leukemia. By identifying novel genes acting as effectors of TAL1 or involved in the regulation of this major T-cell oncogene, we wish to contribute to the overall knowledge of T-ALL biology and pathophysiology and to identify potential molecular targets for therapeutic intervention in this malignancy.

The mechanisms leading to aberrant activation of *TAL1* in the majority of T-ALL patients who lack chromosomal rearrangements remain essentially unknown. Whether dysregulation of enhancer/promoter interactions, epigenetic alterations or deregulated trans-acting mechanisms lead to a disease-causing regulatory variant is still unsolved. We hypothesized that TAL1 levels decrease during normal T-cell development at least in part due to miRNA-dependent down-regulation, in which case TAL1 over-expression in some T-ALL cases should be the consequence of deregulated miRNA expression. To address this question, we analyzed conditional Dicer knockout mice and found that thymocytes lacking the expression of miRNAs expressed significantly more *Tal1* transcript than controls, suggesting that TAL1 could be regulated post-transcriptionally by miRNAs during normal thymic development. By performing computational prediction of miRNAs that bind to the human *TAL1* mRNA we then compiled a list of miRNAs that are candidates to regulate TAL1 transcript. Using a luciferase reporter system we confirmed the candidate miRNA/TAL1 mRNA interactions and selected miRNAs that significantly lowered the reporter expression in at least 25%: miR-101, miR-520d-5p, miR-140-5p, miR-448 and miR-485-5p. We further validated the interaction of the selected miRNAs with TAL1 by mutating the miRNA recognizing element (MRE) in the 3'UTR and confirmed that miR-101 and miR-140-5p target *TAL1* mRNA in the predicted sites. Next, we evaluated the physiological importance of the miRNA/TAL1 mRNA pairs. Over-expression of miR-520d, miR-101, miR-140, miR-485 and miR-448 in different T-ALL cell lines consistently resulted in the down-regulation of TAL1 transcript and protein. In accordance, inhibition of miR-101 and miR-520d-5p promoted TAL1 protein expression. Importantly, we found that the expression of miR-101, miR-140-5p, miR-448 and miR-485-5p was decreased in T-ALL patient specimens as compared to normal bone marrow samples. These findings favor our hypothesis that aberrant down-modulation of miRNAs that target TAL1 during early T-cell development could be, at least in part, responsible for ectopic expression of TAL1 in some T-ALL cases.

TAL1 appears to be on the top of a transcriptional network that, in transformed thymocytes, drives the expression of genes involved in abnormal proliferation, differentiation and survival. Yet, the pathways downstream of TAL1 that contribute to leukemia

development are still poorly characterized. Having that in mind, miRNA genes are attractive candidates to fulfill the role of TAL1 targets with important consequences for leukemogenesis. Through characterization of a TAL1-dependent microRNA gene expression profile, we identified miRNAs whose expression changed significantly upon TAL1 over-expression in a T-ALL cell line. The microRNA screen profile results were then validated upon enforced expression or silencing of TAL1 in additional T-ALL cell lines, confirming that miR-135a, miR-223 and miR-330-3p were up-regulated by TAL1, whereas miR-146b-5p and miR-545 were down-regulated. Furthermore, we verified that TAL1 binds to a genomic area 3.5kbs upstream of the miR-223 transcription start site, which indicates that miR-223 is a direct target of TAL1 in T-ALL. The fact that TAL1 positively regulates miR-223, which others have shown to down-regulate the tumor suppressor FBXW7, indicates that TAL1 is at the core of a transcriptional network involving (de)stabilization of important proteins with oncogenic impact in T-ALL. Interestingly, functional annotation of validated target genes of the TAL1-regulated miRNAs that we identified revealed that these should regulate biological processes related to inflammation and cancer. Overall, we identified for the first time a set of non-protein coding TAL1 target genes, implicating microRNA genes as part of the transcriptional network downstream of TAL1 whose role may be important in the context of hematopoiesis and T-cell leukemogenesis.

Finally, we focused on miR-146b-5p and evaluated the functional and molecular effects of its deregulation by TAL1 in the context of T-ALL. Our results point to a direct down-regulation of miR-146b-5p by TAL1. Moreover, although no clear differences were found concerning cell viability or proliferation, miR-146-5p reduced expression improved the migration and invasion capacities of T-ALL cells. Therefore, miR-146-5p appears to have a tumor suppressor function *in vitro*. In line with these results, we verified that leukemia cells from T-ALL patients express significantly lower levels of miR-146b-5p than normal control samples. This further suggests that reduced expression of miR-146b-5p might be beneficial for T-ALL cells. In agreement, we showed that over-expression of miR-146b in human T-ALL cells significantly increased the survival of recipient mice in a xenotransplant model of human leukemia. In summary, our results suggest that miR-146b-5p is a functionally relevant TAL1

microRNA target gene, whose down-regulation may contribute to T-ALL by modulating leukemia cell motility and disease aggressiveness.

Overall, the work presented in this thesis establishes the existence of a dual-way talk between TAL1 and microRNA genes in a manner that involves the upstream epigenetic regulation of TAL1 by specific miRNAs and the downstream transcriptional regulation of miRNA genes by TAL1. In other words, our studies contribute to the understanding of the mechanisms involved in TAL1 over-expression in T-ALL and to the clarification of the tumorigenic network downstream from TAL1, namely concerning the identification of miRNA genes transcriptionally regulated by TAL1. Whether the interactions identified here may be explored therapeutically for the benefit of T-ALL patients remains a challenge for the future.

Abbreviations

3'UTR	3'Untranslated Region
Ag	Antigen
ALL	Acute Lymphoblastic Leukemia
AML	Acute Myeloid Leukemia
B-ALL	B-cell Acute Lymphoblastic Leukemia
bHLH	basic Helix-Loop-Helix
BM	Bone Marrow
bp	Base pairs
BRCA1	Breast cancer type 1 susceptibility protein
C/EBP	CCAAT-enhancer-binding protein
cCD3	Cytoplasmic CD3
CD	Cluster of Differentiation
CDK	Cyclin Dependent Kinase
cDNA	Complementary Deoxyribonucleic Acid
ChIP	Chromatin Immunoprecipitation
CK2	Casein Kinase 2
CLL	Chronic lymphocytic leukemia
CLP	Common Lymphoid Precursor
CML	Chronic Myeloid Leukemia
CMP	Common myeloid Precursor
CNS	Central Nervous System
CR	Complete response
DLBCL	Diffuse large B cell lymphoma
DMEM	Dubelco's Modified Eagle Medium
DN	Double Negative
DNA	Deoxyribonucleic Acid
DP	Double Positive
EGFR	Epidermal growth factor receptor
EGIL	European Group for Immunological Characterization of Leukemias
ETP	Early Thymic Progenitors
ETP-ALL	Early T-cell Precursor ALL
FACS	Flow Activated Cell Sorting

FBS	Fetal Bovine Serum
FBW7	F-box and WD repeat domain containing 7
FITC	Fluorescein Isothiocyanate
FOXO1	Forkhead box O1
FSC	Forward Scattered Light
GFP	Green Fluorescent Protein
GPA	Glycophorin A
H3K9Ac	Acetylated Histone H3 at Lysine 9
HDAC	Histone Deacetylase
HDACi(s)	Histone Deacetylase Inhibitor(s)
HOX	Homeobox genes
HSC	Hematopoietic Stem Cells
ICN	Intracellular Notch
IL	Interleukin
ISP	Immature Single Positive
IV	Intravenous
JAK	Janus Kinase/ Just Another Kinase
Kb	Kilobases
Kd	knockdown
KDa	KiloDalton
KO	Knockout
lincRNAs	Long intergenic noncoding RNAs
LMO	LIM-only domain proteins
LN	Lymph Node
LT-HSC	Long-term Hematopoietic Stem Cells
LTR	Long Terminal Repeat
MAPK	Mitogen-Activated Protein Kinase
mCD3	Membrane CD3
Mcl-1	Induced Myeloid leukemia Cell differentiation protein
MEP	Megakaryocyte/Erythroid lineage Progenitor
MHC	Major Histocompatibility Complex
miR, miRNA	microRNA
MLL	Mixed Lineage Leukemia
MPPs	Multipotent Progenitors

MRD	Minimal Residual Disease
MRE	microRNA Recognition Element
mRNA	Messenger Ribonucleic Acid
MSCs	Mesenchymal Stromal Cells
mTOR	Mammalian Target Of Rapamycin
ncRNAs	Non-coding RNAs
NF-κB	Nuclear Factor *kappa*-light-chain-enhancer of activated B cells
NK	Natural Killer cells
NT	Non-targeting
nt	Nucleotides
OE	Over-expression
OR	Overall response
ORFs	Open Reading Frames
OS	Overall Survival
PB	Peripheral Blood
PBMCs	Peripheral Blood Mononuclear Cells
PBS	Phosphate Buffered Saline
PE	Phycoerythrin
PerCP	Peridinin Chlorophyll Protein
PI3K	Phosphatidylinositol 3 kinase
PRC2	Polycomb Repressive Complex 2
pri-miRNA	Primary microRNA transcript
pTα	Pre-TCRα chain
PTEN	Phosphatase and Tensin homolog
PTK	Protein Tyrosine Kinase
qPCR	Real-Time Polimerase Chain Reaction
R0	RPMI culture medium (serum-free, 0% FBS)
R10	RPMI culture medium supplemented with 10% FBS
RALDH2	Retinaldehyde Dehydrogenase 2
RB	Retinoblastoma protein
RBCL	Red Blood Cell Lysis buffer
RFP	Red Fluorescent Protein
RISC	RNA Induced Silencing Complex
RNA	Ribonucleic Acid

RNAi	RNA interference
RPMI	Roswell Park Memorial Institute medium
RT-PCR	Reverse Transcriptase PCR
SCID	Severe Combined Immunodeficiency
SCL	Stem Cell Leukemia (TAL1)
SCR	Scramble
SD	Standard Deviation
SDS-PAGE	SDS Polyacrylamide Gel Electrophoresis
SIL	SCL-Interrupting Locus
siRNA	Small Interfering RNA
SP	Single Positive
Sp1	Specificity Protein 1
SSC	Side Scattered Light
STAT	Signal Transducer and Activator of Transcription
ST-HSC	Short-term Hematopoietic Stem cells
TAL1	T-cell Acute Lymphocytic Leukemia protein 1 (SCL)
T-ALL	T-cell Acute Lymphoblastic Leukemia
TAN1	Translocation Associated Notch1
TCR	T-cell Receptor
TCRA/D	T-cell Receptor α/δ gene
TCRB	T-cell Receptor β gene
TCRG	T-cell Receptor γ gene
TGFb	Transforming growth factor beta
TSS	Transcriptional Starting Site
UCB	Umbilical Cord-Blood
VDJ	Variable-Diversity-Joining
VEGFR	Vascular Endothelial Growth Factor Receptor
VSVG	Vesicular Stomatitis Virus protein G
WT	Wild Type

CONTENTS

Figures and Tables index

Chapter 1

GENERAL INTRODUCTION

Hematologic malignancies

Hematopoiesis (from the greek *Haema* – 'blood' and *poiese* – 'to make') is the developmental process that originates all cells from the blood. The process starts with a pluripotent stem cell (hematopoietic stem cell or HSC) that resides in the bone marrow of adult mammals. HSCs originate progenitor cells that become progressively restricted to several or single lineages and have the capacity to differentiate into all the cell lineages from the hematopoietic system, including erythrocytes, megakaryocytes, basophils, eosinophils, neutrophils, monocytes (commonly known as myeloid cells), and lymphocytes [1, 2].

Hematologic malignancies designate tumors from blood cell origin and they can be subdivided in leukemia and lymphoma. Leukemic cells have origin in cells from the bone marrow and can be found in circulation in the peripheral blood, whereas lymphoma is a tumor from the lymphatic system and tends to produce tumor masses. Leukemias can be classified as lymphoid or myeloid according to the lineage of origin of the malignant cells. Leukemias from lymphoid origin are subdivided into B-cell or T-cell leukemias, depending on the lymphoid cell they are originated from. Leukemias can be further classified into acute or chronic. Acute leukemias presents rapid proliferation of immature blast cells, while chronic leukemias are characterized by a slower and indolent progression of more mature cells [1].

Childhood acute leukemia

Childhood acute leukemia is the most prevalent pediatric malignancy representing 31% of all new cancer cases per year in the US (~ 3.250 cases), with a rate of incidence of 30–45 per one million children per year [3, 4]. Overall, in Europe, leukemia represents about one-third of cancers in children bellow 15 years and the risk of a newborn to develop leukemia is about 0.08% [5]. These statistics represent an annual European incidence of childhood leukemia of 44 cases per million, being 35.9 per million of ALL (acute lymphoblastic

leukemia) and 6.5 per million of AML (acute myeloid leukemia) [5] (data from 1988 to 1997). Of concern is the fact that the incidence of leukemia has risen continuously approximately 1% per year over the past two decades [4]. In Europe specifically, the average annual increase of incidence is 0.6% for leukemia in general and 0.8% for ALL [5]. This is in accordance to the observation that the incidence of childhood cancer — not only leukemia — has been increasing in developed countries.

In the past 50 years the development of more efficient therapies increased the five-year event-free survival rate to around 80% for children (and nearly 40% for adults) [6], but survivors face long term complications and morbidities [7] and develop serious health problems within 30 years of their initial diagnosis [8]. In fact, the risk of chronic health conditions is high, particularly for second cancers (breast, colorectal and skin cancer), cardiovascular disease, renal dysfunction, severe musculoskeletal problems, and endocrine pathologies (premature gonadal failure, thyroid disease, osteoporosis, and hypothalamic and pituitary dysfunction) [7]. Therefore the current challenge in the field is to develop more efficacious therapies, biologically targeted and with decreased toxicity rates. A better comprehension of the pathogenesis of the disease, namely molecular analysis of the common genetic alterations in leukemic cells may be the solution. This could provide a better understanding of why some cases fail to respond to chemotherapy and improve selective targeting of leukemic cells without long-term effects on the normal tissues.

Leukemia is believed to be the product of two or more molecular alterations that give the cell the capacity to proliferate while maintaining an immature state. Because they are originated from precursor blood cells, leukemic cells have an intrinsic ability for intra and extravasation, movement in the bloodstream and capacity to proliferate. The known causes for leukemia origin explain less than 10% of the cases, and include ionizing radiation and congenital genetic syndromes that predispose to the disease such as Down's, neurofibromatosis, Fanconi's anemia, and Bloom's Syndrome [3]. The high degree of cellular division during hematopoiesis creates a scenario whereby perturbation via environmental insults including chemicals may induce mutations. The only validated environmental cause so far for leukemia is ionizing radiation from diagnostic imaging during pregnancy or atomic

bomb exposure during childhood, but other suspected causes include the diet of the mother and child, parental smoking, pesticides and household chemicals, traffic fumes, and immunologic modifiers [3].

Childhood leukemia (2 to 10 years) is dominated by a pre-B-lymphocytic phenotype, called common acute lymphoblastic leukemia or B-cell cALL subtype (characterized by the cell markers CD19 and CD10) [7], with an age peak at 2–3 years [5]. Adult leukemias differ substantially from their childhood equivalents. For example, chronic lymphocytic leukemia (CLL), not found at all in children, is the most prevalent leukemia subtype in adults. In addition, adults present more frequently with myeloid leukemias than children and the molecular phenotypes – chromosomal aberrations, mutations – displayed by the leukemic cells are different between the two age groups [3].

Acute Lymphoblastic Leukemia

Acute lymphoblastic leukemia (ALL) is a hematologic cancer originated from the malignant transformation of a lymphoid precursor that is blocked at an immature stage of the lymphoid differentiation. The rates of cure (absence of disease for at least 10 years) are about 80% for children and 40% for adults, already including the very few relapse cases that respond well to a second round of therapy [6].

ALL can arise in T- or B-cell precursors, hence being classified as T-ALL or B-ALL, respectively, which can be further subdivided according to specific genetic abnormalities [8] (Figure 1.1). The most common genetic events leading to ALL include aberrant expression of proto-oncogenes, translocations leading to fusion transcription factors or activated signaling kinases, hyperdiploidy, deletions in cell cycle checkpoint genes, and mutated genes in pro-proliferative pathways [6]. These genetic alterations allow the hematopoietic stem cells or progenitors to overcome and subvert their intrinsic controls for cellular function, blocking the differentiation process and conferring them unlimited capacity for self-renewal and the ability to escape from death signals.

Currently almost all patients with ALL can be classified according to specific genetic abnormalities. Despite the important prognostic value and therapeutic implications of single primary somatic genetic abnormalities, experimental models favor the concept of cooperative mutations being necessary to promote leukemia. The MLL translocations may constitute an exception, given that they are highly transforming and the leukemia induced has very short latency and lower number of accumulated lesions [8, 9]. However, in general the mutations associated with leukemia appear insufficient to cause disease by themselves and require a second hit event. In fact, some of the common genetic mutations in leukemia where shown to be present already during the fetal period. For instance, the rate incidence of most frequent translocations *in AML* - TEL-AML1 in ALL and AML1-ETO – in the cord blood from normal born children is around 1%, a frequency that is 100-fold greater than the risk of developing the corresponding leukemia [10]. This means that a significant proportion of the population carries pre-leukemic clones that are not necessarily translated to a foreseeable diagnosis of leukemia. In what concerns leukemia relapse, it was shown that, in the majority of the cases, the clone responsible for the relapse retains (some of) the lesions present at diagnosis while acquiring additional genetic aberrations. In other words, the predominant clone at relapse already exists in the majority of the cases as a minor sub-clone within the diagnostic sample before the initiation of therapy. Cases where the relapse clone has distinct genetic lesions unrelated to the primary leukemia constitute a minority [11].

In the case of pediatric B-ALL, the most frequent genetic defects observed in leukemic blasts are gains or losses of whole chromosomes and/or translocations (Figure 1.1). Actually, hyperdiploidy (more than 50 chromosomes) characterizes more than 25% of ALL cases [6, 7]. Additionally, translocations affecting hematopoietic transcription factors (TF) or oncogenes are key events and include the frequent translocation t(12;21) *TEL-AML1* (or *ETV6-RUNX1*), present in almost one quarter of the cases; the translocation t(1;19) *TCF3-PBX1*; the t(9;22) *BCR-ABL1* and rearrangements of the *MLL* gene. These genetic alterations also define the main biological subgroups and have prognostic impact [12]. Moreover, more than two thirds of children with B-ALL have genetic defects that perturb the lymphoid maturation process by affecting several TF responsible for lineage commitment of lymphocytes, differentiation,

repression of alternative lineages, and/or lymphocyte maturation. These include *PAX5*, the most commonly affected (30% of the cases), *IKZF* (Ikaros), *EBF1* (Early B-factor 1), *LEF1* (Lymphoid Enhancer Factor 1), *IKZF3* (AIOLOS) and *E2A* (TCF3) [12, 13].

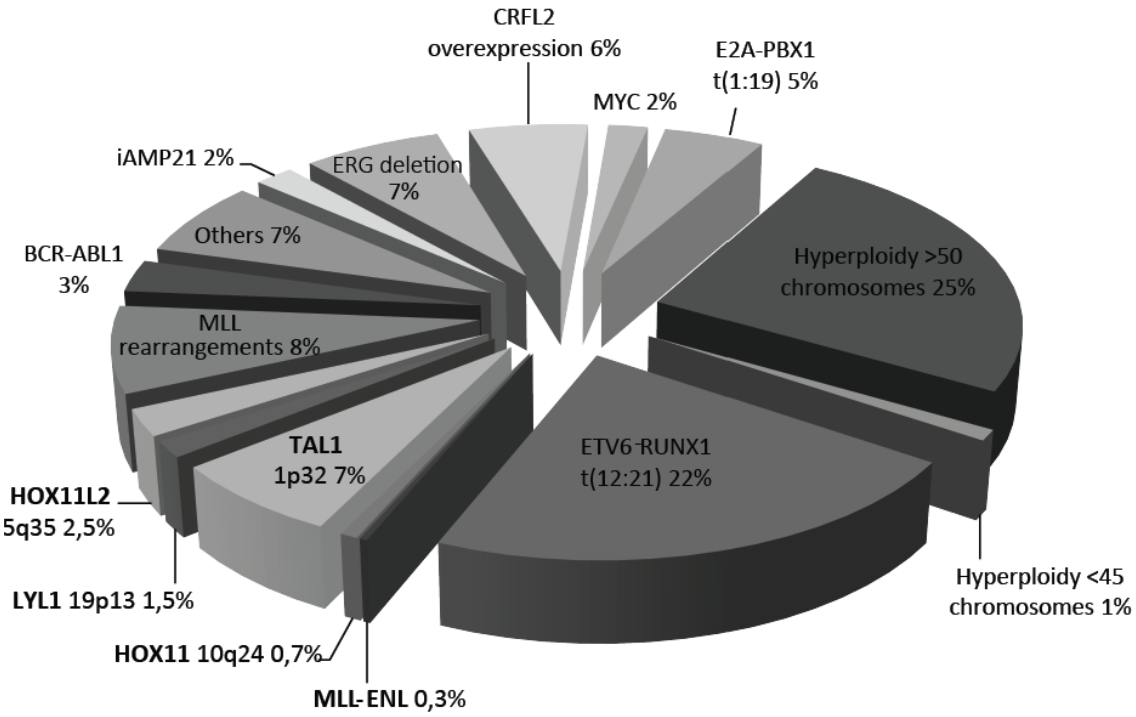


Figure 1.1 **– Estimated frequency of specific genotypes in childhood ALL.**

The genetic lesions that are exclusively seen in cases of T-cell ALL are indicated in bold (TAL1, HOX 11, MLL-ENL, LYL1 and HOX11L2). Adapted from [11].

Diagnosis and Treatment

The symptoms of childhood ALL depend of the degree of disruption of normal hematopoiesis due to bone marrow infiltration by leukemic blasts or infiltration of other extra-medullary sites. Typical symptoms include fever, pallor, fatigue, bruises, enlargement of liver, spleen, and lymph nodes, and bone pain, the majority of them due to anemia (low erythrocyte counts) and thrombocytopenia (low platelet count). The diagnosis is established when at least 25% leukemic blasts are present in the bone marrow [14]. The advances in chemotherapy over the last 50 years allow for an actual cure rate for ALL approaching 90% in

many developed countries. In fact, the developments in the treatment of ALL makes it one of the most successful stories of modern medicine: the 10 year survival rate increased from 11 to 90% in almost 50 years (total therapy studies data from St. Jude Children's Research Hospital between 1962 and 2007) [8]. This improvement was possible by introduction of risk-adjusted intensive multi-agent chemotherapeutic regimens, involving the use of glucocorticoids and other drugs such as vincristine, asparaginase, methotrexate and/or anthracyclines, and additional radiotherapy and/or hematopoietic stem cell transplantation for specific subgroups of patients [14].

The recent enhancement in treatment outcome for BCR/ABL-positive ALL patients [14] that received an ABL tyrosine kinase inhibitor (Imatinib) serves as a remarkable proof of concept of the clinical advantages of the use of oncogenic pathway-directed therapy and more personalized targeted therapies in ALL [8]. Patient differences in drug metabolism and pharmacokinetics have also been incorporated in recent trials. The idea is to personalize the amount of therapeutic agent administered in accordance to the ability of each patient to clear the drugs or to the genotype of their leukemic cells, without compromising efficacy of ALL treatment [8].

T-cell Acute Lymphoblastic Leukemia

T-ALL is much less frequent than ALL from B-Cell origin, representing only 15% of childhood and up to 25% of adult ALL [6]. Perhaps for this reason it has historically deserved less attention from researchers and is frequently included several times as a small part of large studies in general ALL, not being treated as an independent entity.

T-ALL patients belong to the high-risk group and tend to associate with poor prognosis. Nonetheless, the overall survival rate is about 80% in childhood and 40-60% in adult T-ALL cases [15]. Although adjustment of the intensity of chemotherapy to risk category improved outcome of T-ALL patients, these still have increased risk for induction failure, early relapse, and isolated CNS relapse [16]. Contrary to B-ALL, a big proportion of T-ALL cases (50%) present with a normal karyotype. In fact, translocations leading to the activation of a

small number of oncogenes occur in 25–50% of T-ALL cases. Other genetic abnormalities include frequent micro-deletions leading to the loss of tumor suppressor genes. The events driving a full malignant phenotype in T-ALL include also over-expression of oncogenes, defects in the cell cycle control, aberrant activation of protein kinases, and activating mutations of *NOTCH1* [17].

In terms of immunophenotype, cytogenetic and molecular genetic abnormalities, T-ALL is a very heterogeneous disease, which is in agreement with the very distinct clinical outcomes observed in T-ALL patients treated with the same intensive multidrug regiments. The organization of T-ALL subgroups is mainly based on the immunophenotype, being the classification proposed by the European Group for the Immunological Characterization of Leukemia (EGIL) the most commonly used in Europe [18] (Table 1.1).

Table 1.1 **- Immunophenotypic classification of T-ALL.**

	Immature T-ALL		**Cortical T-ALL** EGIL T-III		**Mature**
Markers	pro-T-ALL EGIL T-I	Immature T-ALL EGIL T-II	smCD3-	smCD3+	T-ALL EGIL T-IV
TdT	++	++	++	++	++
HLA-DR	+	-	-	-	-
CD34	+	-	-	-	-
CD1	-	-	++	++	-
CD2	+	++	++	++	++
cCD3	++	++	++	++	++
CD5	-	++	++	++	++
CD7	++	++	++	++	++
CD4-/CD8-	++	+	-	-	-
CD4+/CD8-	-	±	±	±	±
CD4-/CD8+	-	±	±	±	±
CD4+/CD8+	-	-	+	+	±
smCD3	-	-	-	++	++
TCRαβ	-	-	-	+	+
TCRγδ	-	-	-	+	+

Based on [17, 19]. **-**: <10% of leukemias are positive; **±**: 10-25% of leukemias are positive; **+**: 25-75% of leukemias are positive; **++**: >75% of leukemias are positive. In the case of TCRαβ, 60-70% of SmCD3+ EGIL T-III and EGIL T-IV leukemias are positive. In the case of TCRγδ, 30-40% of SmCD3+ EGIL T-III and EGIL T-IV leukemias are positive.

The T-ALL cases are identified by the cytoplasmic expression of CD3 (cCD3). Based on the expression of CD2, CD7, CD1 and surface membrane-bound CD3 (smCD3), T-ALL can be subdivided into major groups. The immature group, EGIL-TI or pro-T-ALL expresses only CD7 (cCD3+/CD7+). The pre-T-ALL or EGIL-TII subgroup is further characterized by the expression of CD2 and/or CD5 and/or CD8; a significant proportion of these cells do not express cytoplasmic T-cell receptor beta chains (cTCRβ), and frequently co-express precursor markers such as CD13, CD33, CD56 or CD34. Intermediate or cortical T-ALL or EGIL-TIII cases express CD1a, and are frequently double-positive for CD4/CD8, and partly positive for smCD3. Lastly, mature T-ALL or EGIL IV is characterized by the presence of surface CD3 and absence of CD1a. The surface CD3 expressing cases can be further sub-divided into group a or b, depending on whether they express TCRαβ or γδ, respectively [17, 19].

Normal T-cell development

T-ALL arises from the clonal expansion of a lymphoid precursor that is transformed and blocked at a certain stage of differentiation. The genetic alterations acquired by the leukemic blasts also confer them self-renewal capacity, enhanced proliferation and survival. Thus, to better understand the molecular mechanisms behind T-cell malignancy one should have a clear comprehension of normal T-cell development (Figure 1.2).

Lymphopoiesis is the process by which the mature populations of T, B and NK (natural killer) cells are originated from multipotent hematopoietic stem cells (HSCs) [20]. T-cell development starts in the bone marrow where the HSCs reside in a favorable niche. These cells have the capacity of self-renewal and differentiation into any hematopoietic lineage [21]. The long-term repopulating HSCs (LT-HSCs) give rise to short-term (ST)-HSCs, characterized by limited self-renewal activity, and to multipotent progenitors (MPPs). MPPs can commit to either common lymphoid progenitors (CLP), restricted to give rise to all lymphoid cells, or to common myeloid progenitors (CMP), which will originate all cells from the myeloid lineage [2]. It is not yet clear which BM hematopoietic progenitor is the bona fide T-lymphocyte progenitor, but this will migrate from the bone marrow into the thymus as

early T-cell lineage progenitor (ETP) and further commit to T-cell differentiation. Some studies showed that when CLPs enter into the thymus, they are still able to differentiate into B, T, NK or dendritic cells [20, 22].

This input of progenitor cells from the bone marrow to the thymus is essential to maintain T-cell development [23]. Briefly, the ETPs in the thymus represent the first stage (DN1) of the double negative fraction, thus called because of the undetectable expression of the surface markers CD4 and CD8. DN1 cells (CD3- CD4- CD8-CD34+) progress in development by down-regulation of the stem marker CD34, up-regulation of CD5, CD1a, CD4 and by the rearrangement of the T-cell receptor (TCR). The TCR is a heterodimer composed by two transmembrane chains: either the combination of $\alpha\beta$ or $\gamma\delta$ chains is possible. Each chain has a variable and a constant domain [17, 22]. The $\gamma\delta$ T-cells originate at the end of DN2 stage [20, 22, 24] following rearrangement of the δ (in the DN1 stage) and γ (DN2 stage) genes. However, most T-cell precursors do not originate this lineage, but rather $\alpha\beta$ T-cells. During the transition from DN3 to DN4, the cells start to rearrange the β locus of the TCR. The genomic locus encoding each of the TCR possible chains is composed of several gene clusters corresponding to the Variable (V), Diversity (D), Joining (J) and Constant (C) regions of the chain. Thus, a random choice between the various V-, D-, J segments and through the deletion or addition of extra nucleotides at the junctions of the segments, mediates the vast repertoire of antigens that can be recognized by each individual TCR. This process follows a strict order of gene rearrangements and the recombinase-activating genes *RAG-1* and *RAG-2* are essential for the rearrangement of variable region genes and the development from the DN3 stage onwards [17, 22]. Despite being highly regulated, subversion of these recombination events is associated with translocations involving the TCR loci in T-ALL, as will be latter discussed.

At the immature single positive (ISP) stage the pre-T cells acquire the surface expression of the pre-TCR (constituted by a productively rearranged TCRβ chain and an invariant form of the TCR called the pre-Tα -encoded by the gene *PTCRA*) and expression of CD4 co-receptor (CD8 in the mouse). The assembly of the pre-TCR is essential to progression to the double positive (DP; CD4+CD8+) stage and the pre-TCR-derived signals allow intense

cellular expansion and suppression of cell death [25]. As a result DP cells comprise the majority (~90%) of the cells in the adult thymus. The subsequent rearrangement of the TCRA genes allows the expression of a mature αβ-TCR. Is through the variable region of the mature αβ-TCR that T-cell recognize small peptides bound to antigen-presenting molecules, the MHC class I and MHC class II molecules, expressed in antigen-presenting cells [26].

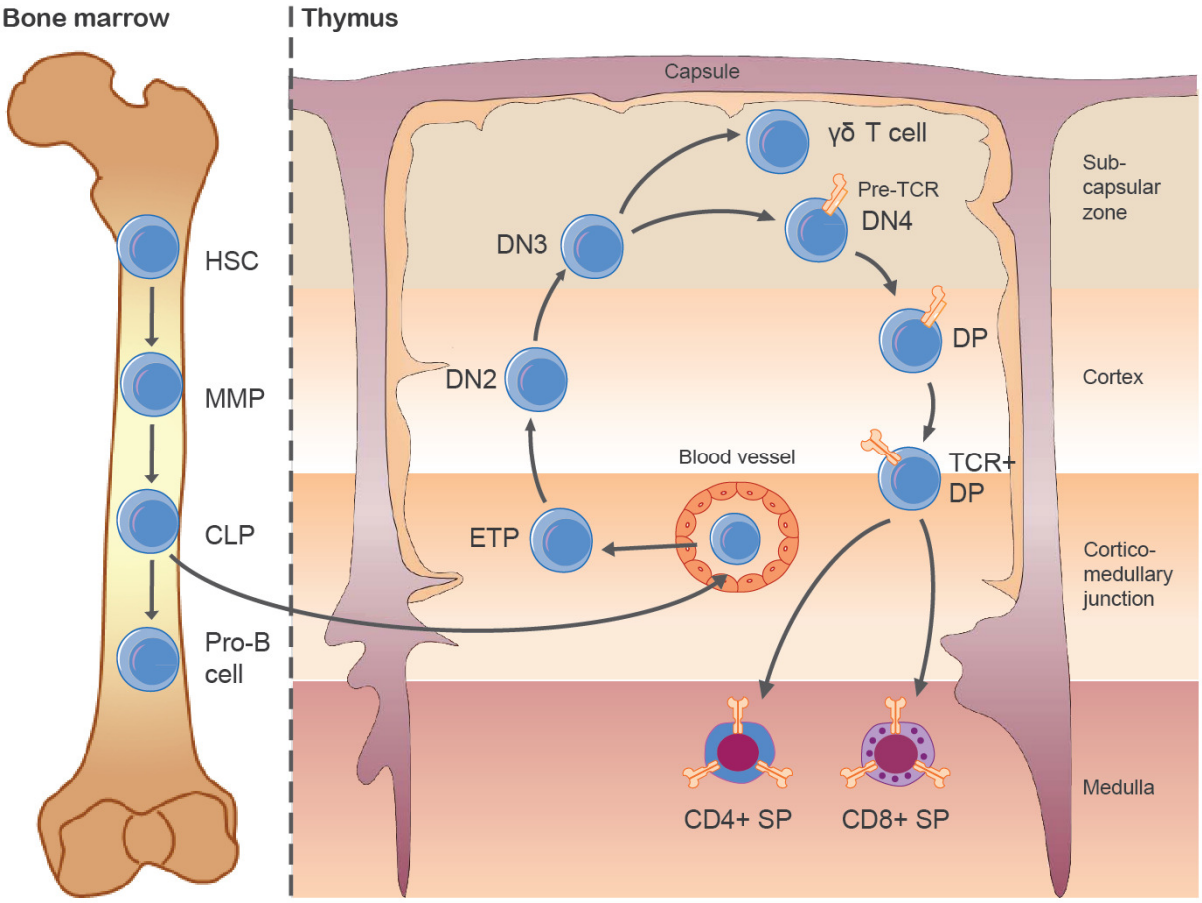


Figure 1.2 - **Stages of hematopoiesis and T-cell development**

Bone-marrow HSCs exit the quiescent 'niche' and differentiate to become multipotent progenitors (MPPs). MPPs further commit to the lymphoid lineage generating CLPs. These subsets migrate to the thymus (as ETPs) and commit to the T‑cell lineage, progressing through the double negative (DN; CD4–CD8–) stages, DN2, DN3 and DN4. Upon successful recombination at the T‑cell receptor β (TCRB) locus, pre‑T cells acquire surface expression of the pre-TCR that promotes massive proliferation and differentiation. Pre-TCR-selected cells reach the double positive (DP; CD4+CD8+) stage, at which point they are subjected to positive and negative selection. Selected cells then exit the thymus as single positive (SP) CD4+ or CD8+ T cells. Adapted from [22].

The DP thymocytes are programed to undergo apoptosis unless they are rescued by the survival signals mediated by TCR recognition with moderate affinity of self-MHC/self-peptide molecules on thymic epithelial cells. This process is called positive selection, and the cells that do not detect any antigen die by neglect. On the other hand, DP cells with high TCR affinity to self-MHC complex are eliminated by apoptosis (negative selection) to avoid autoreactivity. This double selection process assures the generation non-self-reactive thymocytes that recognize self-MHC molecules and the production of antigen-specific single positive (SP) mature T-cells with high CD3/TCR$\alpha\beta$ expression [17, 26]. Mature SP T-cells are then divided in two classes that are distinguished by the CD4 or CD8 membrane co-receptor expression, which translates in different functions in the immune system. In a simplistic view, CD4+ or helper T-cells recognize MHC II-bound peptides, derived from extracellular sources and regulate the immune response of B cells, other T cells and the activity of cells of the innate immune system. The CD8+ or cytotoxic T-cells recognize MHC I-bound peptides, derived from proteins synthesized within the cell, and are involved in the lysis of infected or transformed cells [26] (Figure 1.2).

The commitment of progenitor cells to the T-cell lineage depends on signals from the microenvironment and contact with thymic epithelial stromal cells. These cells produce essential growth factors, such as interleukin 7 (IL7) [27], which are required for survival and proliferation of human T-cell precursors. Moreover, the contact with the NOTCH ligands expressed by stromal cells is vital to the early developmental program [28, 29]. Later on, differentiation-inducing signals are generated by the TCR, upon pre-TCR expression. The TCR signaling is very important for thymocyte survival and is mediated by downstream effectors that include the tyrosine kinases LCK, FYN and ABL1, the kinases ZAP-70 and PI3K, the RAS-MAP kinase pathway, the anti-apoptotic transcription factor NFκB and cyclin D3. The activated signaling cascades induce the transcription of various genes promoting clonal expansion of the antigen-reactive T-cell [26].

Genetic abnormalities in T-ALL

T-ALL and the subversion of the normal T-Cell development

The aberrant expression of T-cell oncogenes (mainly transcription factors) is a very common event in T-ALL (80%), without a detectable causal cytogenetic abnormality in the chromosomal locus of these genes [30]. Moreover, T-ALL gene expression studies revealed that the signatures associated with activation of oncogene expression in leukemic cells can be interpreted in the light of the block at particular stages of normal thymopoiesis and that the oncogene-driven gene expression profiles can be associated with the immunophenotypic classification of T-ALL discussed above: the LYL1 signature associates with immature thymocytes (pro-T-ALL); activation of TLX1, TLX3, NKX2.1 and NKX2.2 [31] homeobox genes are characteristic of CD1a/CD4/CD8+ early cortical thymocyte T-ALL; and high TAL1 expression to late cortical CD4/CD8/CD3+ T-ALL [30]. Recently, a new subtype of T-ALL has been defined by a specific gene-expression signature and distinct immunophenotype, termed early T-cell precursor (ETP-ALL), encompassing up to 15% of T-ALL patients. This subtype is characterized by diverse genetic alterations, absence of surface CD1a and CD8 and weak expression of CD5 and expression of stem-cell or myeloid markers. Moreover, ETP-ALL is associated with an increased risk of therapy failure [32, 33].

Expression analysis of TCR, pre-TCR and RAG-1 in T-ALL samples confirms that it reflects the stages of normal T-lymphoid maturation [34]. In fact, 50% of TCR$\alpha\beta$ lineage T-ALLs that express pre-TCR, RAG-1, pTα, and cytoplasmatic TCRβ are CD1a+ CD4/8+ double-positive, without TCRD deletion, similar to a physiological population undergoing or having just undergone β selection. On the other hand, the majority of T-ALLs from the TCR$\gamma\delta$ lineage have only TCRβ DJ rearrangement, are negative for pTα, TdT and RAG-1. Nonetheless, 40% of TCR$\gamma\delta$ T-ALLs express pTα, TdT, and RAG-1, and are CD4 CD8 DP with TCR β V(D)J rearrangements. Around 30% of T-ALL cases express surface CD3 without TCRβ rearrangement, and non-T-cell restricted markers such as CD34, CD13, CD33, and CD56 corresponding to non-restricted thymic precursors. Not all features of T-ALL can be explained on this basis, nevertheless T-ALLs largely represent physiological T-lymphoid maturation, and

understanding the differences allow the distinction of physiological and leukemogenic profiles [34].

Besides the mutually exclusive oncogenic lesions involved in T-ALL ontogenesis, a number of recurrent cytogenetic and molecular alterations, common to all molecular subtypes, are observed. These are responsible for the deregulation of several critical cellular processes, such as cell cycle signaling, cell growth and proliferation, chromatin remodeling, T cell differentiation and self-renewal (such as those affecting *NOTCH1, CDKN2A/B*, *FBXW7*, *PTEN* and others discussed below).

Cytogenetic abnormalities

The analysis of cytogenetic abnormalities has led to a better understanding of the molecular causes of T-ALL. These genomic relocations typically juxtapose the strong promoter and enhancer elements of TCR genes with developmentally important transcription factor genes, such as the basic helix–loop–helix family members *TAL1*, *TAL2*, *LYL1* and BHLHB1; the LIM-only domain genes LMO1 and LMO2; the homeobox genes *HOXA-HOXD*, *HOX11*, *HOX11L2* and NKX family; and oncogenes MYC, MYB, and NOTCH1, leading to the up-regulated expression of these genes [17, 35-37]. Fusion events involving TCR genes can also be found involving potential oncogenes other than TF, such as IL7R and PLAG1 [36].

In fact, 35% of translocations found in T-ALL cases are associated with the TCR loci, involving rearrangement of the human chromosome 7 (*TCRB* and *TCRG*) or chromosome 14 (*TCRA* and *TCRD*) with the genes mentioned and other unknown partner genes (5–10% of the cases) [35]. This results most probably because during TCR rearrangement other regions in the genome have an open chromatin configuration and become susceptible to the activity of the RAG1 and RAG2 enzymes. Noteworthy, some of these events only have an oncogenic phenotype when associated with other genomic aberrations, suggesting that other pathways have to be subverted for leukemogenesis to occur [22]. Translocation events not affecting TCR genes include the translocations involving the *MLL* gene with different partner genes and accounts for 8% of T-ALL cases [38]. More frequent, the translocation t(10;11)(p13;q14) encoding CALM-AF10 and restricted to the TCRγδ lineage is present in 10% of T-ALL cases

[39]. Although more uncommon, translocations involving the *ABL1* gene, such as the t(*NUP214-ABL1)* are present on amplified episomes of about 6% of T-ALL cases [40].

T-cell transcription factor oncogenes can be activated as result of other genetic rearrangements, such as generation of fusion genes that encode a chimeric protein with oncogenic properties, mostly due to translocations but also to micro-deletions [17]. These include the *SIL-TAL1* fusion gene found in 9–30% of childhood T-ALLs as a result of a micro-deletion in the *TAL1* locus that fuses the *TAL1* coding region to the *SIL* regulatory elements. This is one of the most recurrent fusion events in T-ALL [36]. In addition, cryptic deletions in chromosome 11p13 lead to the activation of the proximal *LMO2* promoter in approximately 4% of pediatric patients [41]. With the advent of whole transcriptome sequencing techniques, new fusion genes were detected (SSBP2-FER and TPM3-JAK2) that encode for typical tyrosine-kinase fusions that join the tyrosine-kinase domain of JAK2 or FER to the dimerization units of TPM3 or SSBP2, respectively [36]. Although rare, additional fusions not involving the TCR genes but also leading to the ectopic expression of oncogenic transcription factors can be found in T-ALL and include PLAG1, MEF2C, ZNF219, and BMI1 genes [36].

Another genomic event leading to oncogene over-expression is the phenomenon of gene duplication that was found for the MYB oncogene in around 8% of T-ALL individuals [36, 42]. Additionally, deletions involving the loss of tumor suppressor genes are also found in this hematological malignancy, being the most frequent the loss of the *INK4/ARF* locus at chromosome 9p2, a region coding for important proteins in the cell cycle regulation [43]. In addition, small deletions and/or insertions such as those involved in NOTCH1 activating mutations [44] and gain of function mutations in cytokine receptors and tyrosine kinases, namely IL7R [45] and JAK3 [46], also occur in T-ALL with oncogenic driving force.

Recently, the use of techniques of whole transcriptome sequencing to comprehend the mutational spectrum of T-ALL patients and cell lines has further confirmed that leukemia is caused by a mixture of gene fusions, over-expression of transcription factors and cooperative point mutations in oncogenes and tumor suppressor genes [36, 47]. Hence, over-expression of TF, such as TLX1/3 and TAL1, combined with NOTCH1 activating mutations, additional mutations affecting chromatin modifiers [48], signaling factors specially those

involved JAK-STAT signaling pathway [46], tumor suppressor genes (like TP53, PTEN, WT1) and mutations in ribosomal genes like RPL5 and RPL10 [47] define a more updated view in the field of the pediatric T-ALL genome [36].

Cell cycle defects

Overcoming cell cycle progression controls and checkpoints, with consequent propagation of genomic instability and uncontrolled cell division, is one of the hallmarks of cancer [49]. Thus, mutations affecting proteins controlling cell cycle transitions and checkpoints, and DNA damage mediated responses are common events in cancer, and leukemia is not an exception.

The negative regulator of the cell cycle p16INK4A (*CDKN2A*) inhibits the cyclin D-dependent kinases (CDKs) CDK4 and CDK6. These kinases contribute to the release from a quiescent cell state (G0) into G1 and subsequent commitment to the cycle by inactivating the retinoblastoma protein (RB), which in turn regulates the entry into the S phase of the cell cycle. Interestingly, the *p14ARF* (*CDKN2A*) gene is transcribed from the same locus as *p16INK4A* but originates the tumor suppressor ARF by an alternative reading frame. ARF is involved in cell cycle regulation by inhibiting MDM2, a negative regulator of the major tumor suppressor p53 that targets p53 to degradation. Association of ARF with MDM2 results in p53 up-regulation and p21 (CDKN1A) activation. In turn, p21 which is a CDK inhibitor, mediates cell cycle arrest in G1 allowing either DNA repair or programed cell death [50].

In T-ALL, the most frequent genomic anomaly detected is a deletion affecting the *INK4/ARF* locus on chromosome 9p21. The inactivation of p16/p14 is found in 70% of the patients due to deletion, mutation or hypermethylation [43, 51] and also by transcriptional and post-transcriptional inhibition [17]. The vast majority of pediatric T-ALL cases display functional inactivation of these cell cycle regulators, making the RB and p53 pathways very attractive to be therapeutically targeted [51]. This assumes an even higher importance by the fact that homozygous *INK4/ARF* locus deletion in T-ALL has unfavorable prognostic value, with a significantly lower 5-year disease-free survival [52] and by the fact that the majority of TAL1+ and HOX11+ T-ALL cases present with homozygous deletion of p16 [51]. Furthermore,

a chromosomal translocation targeting the *CCND2* (Cyclin D2) locus at 12p13 and TCRB or TCRA/D loci leads to a strong up-regulation of the protein in T-ALL. Cyclin D2 activity is required for G1/S cell cycle transition, and its high expression was associated with TAL1+, TLX1+ or TLX3+ molecular subtypes, *NOTCH1* mutation and *INK4A* deletion [53].

Finally, transcription factor modulation can also disrupt cell cycle control in T-ALL, as will be detailed later. Briefly, CALM-AF10 T-ALLs over-express BMI1, which controls cellular proliferation by suppressing p16; enforced TAL1 expression negatively regulates *p16* gene through interfering with the E-box sequences of the promoter; and HOX11 interacts with the protein phosphatases PP2A and PP1 disrupting a G2/M cell-cycle checkpoint [17].

Aberrant signaling

It is important to note that the genetic alterations found in T-ALL and responsible for the irregular transcriptional programs frequently lead to the abnormal activation of signaling pathways crucial to proliferation and death escape of leukemic cells [54]. Besides cell intrinsic lesions, also environmental stimuli are important in triggering oncogenic signals that promote T-ALL expansion [55]. In this regard, signaling transduction pathways are key effectors in regulation of T-cell viability, proliferation and immune response, namely as important downstream effectors of the TCR signaling and membrane receptors that transmit environmental cues. Therefore it is not surprising that certain signaling pathways are aberrantly activated in T-ALL due to chromosomal translocations, creation of chimeric fusion genes, point mutations and deletion of negative regulators [17]. Below, we will briefly mention some of the most important deregulated signaling pathways in T-ALL.

The ubiquitously expressed tyrosine kinase ABL1 can be part of the highly transformative BCR-ABL1 fusion kinase, characteristic of chronic myeloid leukemia, and also found in 25% of precursor B-ALL, but very infrequent in T-ALL (less than 1% of the cases). More frequent in T-ALL is the *NUP214-ABL1* fusion, found in up to 6% of T-ALL cases, also characterized by up-regulation of *TLX1* or *TLX3* and deletion of *INK4A* locus [40]. Other fusions involving ABL1 occur very rarely in T-ALL, like the *ETV6-ABL1* [56] and the *EML1-ABL1* gene fusion [57]. All these fusions result in constitutive activation of ABL1, leading to

enhanced survival and proliferation. Another member of the signaling cascade downstream of the TCR and upstream of ABL1 is the protein tyrosine kinase LCK. In rare cases of T-ALL, the translocation t(1;7)(p34;q34) that juxtaposes *LCK* to the *TCRB* locus was described, leading to overexpression of this kinase [58].

The Janus kinase (JAK) family of tyrosine kinases plays an important role in transmitting signals from cytokine receptors to downstream effectors, being therefore very important for core biological processes such as apoptosis, differentiation, proliferation and also immune responses [54]. In childhood T-ALL, a fusion of the ETS-variant gene 6 (ETV6) to *JAK2* originates a chimeric fusion protein with constitutive tyrosine kinase activity and transformative capacity in transgenic mice [59]. Moreover, mutations in *JAK1* were described in 18%-27% of adult and 2% of pediatric patients with T-ALL and were associated with poor response to therapy and overall poor prognosis [60, 61]. Recently, activating JAK3 mutations were described in 7% of T-ALL patients [46]. In addition, activating mutations in JAK1 and JAK3 have also been reported in ETP-ALL [32]. Recently, we and others discovered gain-of-function mutations in the gene coding for the α chain of the IL7 receptor. These mutations result in constitutive activation of the JAK/STAT signaling in approximately 10% of T-ALLs, reinforcing the importance of this signal transducing pathway in T-ALL [32, 45, 62]. IL7 receptor mutations are sufficient to induce cell transformation and tumor formation and one possible way for therapeutic targeting of IL-7R-mediated signaling in T-ALL involves pharmacological inhibition of JAK1 activation [45].

The small GTPases of the RAS family (N-RAS, K-RAS and H-RAS) are anchored at the cellular membrane. They have a pivotal role in transmitting survival signals coming from cytokines and growth factors that bind to membrane receptors, to downstream transduction pathways via MAP kinases and other effector proteins [63]. In T-ALL, activating mutations of *N-RAS* have been detected in 10% of pediatric patients [64]. Moreover, in 50% of T-ALL patients, the level of RAS activation in leukemic blasts exceeds those of the normal cell counterpart, supporting a role for RAS activation in this malignancy [63]. Recently, it was found that the immature/ETP subtype of T-ALL is specially characterized by activating

mutations in genes regulating cytokine receptor and RAS signaling (in 67% of the cases), involving NRAS, KRAS but also FLT3,IL7R, JAK3, JAK1, SH2B3 and BRAF [32].

NOTCH1 encodes a transmembrane receptor that regulates normal T cell development and hematopoietic stem cell maintenance [65]. Despite the fact that translocations involving *NOTCH1* are very rare in human T-ALL, more than 50% of the patients possess activating mutations involving either the extracellular heterodimerization domain (44%), resulting in ligand independent intracellular Notch domain (ICN), or mutation in the C-terminal PEST domain of NOTCH1 (30%), resulting in extended half-life of ICN, or even both events (17%) [44]. Patients harboring NOTCH1 activating mutations belong to all major molecular oncogenic subtypes, *TAL1+*; *LYL1+*; *HOX11+*; *MLL-ENL+*; *CALM-AF10+*, which may indicate that they occur very early in T-cell differentiation [44]. Given that NOTCH proteins (NOTCH 1—4) are indispensable for the commitment of hematopoietic progenitors to the T-cell lineage, the fact that NOTCH is so broadly activated in T-ALL establishes another parallelism between T-cell development and the induction of T-ALL.

Several groups identified the oncogenic transcription factor MYC as a direct transcriptional target of NOTCH1 in T-ALL through the NOTCH1-binding sites present in the *MYC* promoter [66-68]. Furthermore, NOTCH1 activates NF-κB signaling pathway [69], and more than 40% of the potential NOTCH-responsive genes are regulators of cell metabolism and protein biosynthesis [66]. One example is the PI3K–AKT pathway, controlled by NOTCH signaling through the down-regulation of the phosphatase PTEN (phosphatase and tensin homologue) that in turn negatively regulates PI3K signaling [70]. Moreover it was shown that active NOTCH1 directly regulates IL-7Rα gene (IL7R) transcription in thymocytes and T-ALL cells and that IL-7R signaling (through PI3K-AKT and JAK-STAT pathways) is important for proliferation of NOTCH-dependent T-ALL cells [71]. Furthermore, the E3 ubiquitin ligase FBW7 (F-box and WD repeat domain containing 7) responsible for targeting important cell-cycle regulators (MYC, JUN and cyclin E) to degradation, can bind to NOTCH1 PEST domain, regulating its stability [72, 73]. Actually, around 20% of T-ALL patients have inactivating mutations in the FBW7, revealing its tumor suppressor role in this disease [73].

Finally, one of the most important signaling pathways subverted in T-ALL includes the PI3K-AKT-mTOR axis and the tumor suppressor gene *PTEN*. PI3K (Phosphatidylinositol-4.5-bisphosphate 3-kinase) is a key effector kinase downstream of the TCR signaling but also of many other surface receptors. Together with mTOR (mammalian target of Rapamycin), PI3K regulates essential cellular processes such as viability, proliferation, and differentiation [54]. When activated at the cell membrane, PI3K allows the anchoring of the Ser/Thr kinase AKT [74]. AKT mediates pro-survival and pro-proliferative signals by inhibiting through phosphorylation Bad, GSK3, FOXOs and caspase-9, by activating through phosphorylation mTOR, and by the release of transcription factors such as NF-κB [55]. The negative regulators of this pathway, PTEN and SHIP, inhibit the activation of AKT [74]. Furthermore, genetic abnormalities in PTEN, PI3K, and AKT genes are found in roughly 50% of T-ALL patients [75], strengthening the importance of this oncogenic pathway in this malignancy, and inactivation of PTEN without genomic alteration of its locus, contributes to the hyperactivation of PI3K/Akt in primary T-ALL samples [76]. Thus, therapeutic inactivation of PI3K pathway is currently under consideration for T-ALL treatment. In fact, the use of Rapamycin, an mTOR inhibitor, promotes apoptosis of T-ALL blasts [77] and pharmacological inhibition of PI3K increases apoptosis and arrests the cell cycle in T-ALL cells [78] and can prevent chemoresistance to drugs currently used [70].

Epigenetic alterations

Mutations involving epigenetic modifying genes are common events in AML, but less common in ALL. Nevertheless, the most immature subtype of T-ALL, the ETP-ALL, which is characterized by a mutational spectrum more similar to myeloid leukemia and hematopoietic stem cells, has a prevalence of inactivating mutations (in 48% of the cases) in genes involved in histone modification, such as EZH2, EED, SUZ12, SETD2 and EP300 [32].

The polycomb repressive complex 2 (PRC2) is responsible for a major repressive chromatin modification, H3K27me3, and consists of SUZ12, EED, EZH2 and RbAp48. Deleterious mutations in PRC2 genes were found mainly associated with the immature type of T-ALL but not with more differentiated cases [32]. In fact, loss-of-function mutations and

deletions of *EZH2* and *SUZ12* genes were found in 25% of T-ALL patients from an adult cohort, consistent with a tumor suppressor role for PRC2 in T-cell transformation [48].

Another factor, PHF6, a PHD-containing factor with a proposed role in epigenetic regulation, was found mutated and deleted in 16% of pediatric T-ALL cases, almost exclusively in male patients. In childhood T-ALL, PHF6 mutations correlate with the aberrant expression of the homeobox oncogenes TLX1 and TLX3. Despite the uncertain function of PHF6, proteins with PHD domains are associated with readout of methylation marks present at histone residues, thus having a role in the chromatin structure [79].

T-ALL specific transcription factor oncogenes

T-ALL is characterized by the recurrent over-expression of a specific set of transcription factors, namely TLX1, TLX3, TAL1, LMO, HOXA, LYL1 and NKX family members, each defining a distinctive gene expression signature that identifies molecular subtypes in T-ALL [30]. Besides gene rearrangements, disruption of regulatory pathways that normally tightly restrict the expression of transcription factors during normal thymopoiesis, can also be involved in the pathogenesis of T-ALL. In fact, different T cell oncogenes (HOX11, TAL1, LYL1, LMO1, and LMO2) are often aberrantly expressed in the absence of chromosomal abnormalities [30]. Therefore, it has been proposed that the disruption of gene silencing mechanisms that normally down-regulate them during normal T-cell development may explain these cases [80].

The basic helix–loop–helix (bHLH) family of transcription factors is found commonly deregulated in T-ALL. The bHLH motif enables these TF to homo or hetero-dimerize through the HLH domain and to bind to DNA through the basic regions of the dimerized proteins. The bHLH family is divided in two classes: class A, that includes E2A and HEB (named the E proteins), that are widely expressed as homodimers or heterodimers with other bHLH proteins; and class B, that includes LYL1, TAL1, TAL2 and bHLHB1, that only heterodimerize with class A bHLH proteins and are expressed in a tissue-specific manner [81].

The E proteins are very important transcription factors in thymocyte development. They bind to the DNA specifically in E-Box sequences (CANNTG), many of which are located in

enhancers of T-cell lineage restricted genes, like *CD4* and *PTCRA* [82]. The *E2A* codes for two bHLH products, E12 and E47, by alternative splicing. These are critical for proper early B-cell commitment and development but also for the earliest stages of T-lineage development. Loss of *E2A* results in block in early T-cell development and in development of murine T-cell malignancies of immature phenotype, consistent with a tumor suppressor role [83]. Besides pTα expression, E2A proteins regulate also V(D)J recombination and the expression of *RAG* and, in this manner, regulate the thymic development before the formation of pre-TCR [83]. They act by inhibiting cell cycle progression in thymic precursors prior to TCRβ expression [84]. Moreover, reintroduction of E47 or E12 in lymphoma cells originated by the loss of the E2A proteins induces programmed cell death [85], reinforcing their tumor suppressor role in leukemia. Furthermore, enforced expression of the inhibitor of E proteins Id1 in T-cells induces massive apoptosis of differentiating T-cells leading to an accumulation of CD4 and CD8 double negative thymocytes with multipotent progenitor cell markers. Id1 transgenic mice frequently develop T-cell lymphoma with long latency, reinforcing the role of E proteins as regulators for normal T-cell differentiation and tumor suppression [86].

On the contrary, class B bHLH proteins LYL1, TAL1, TAL2 and bHLHB1 are not expressed in the normal thymic development, but are found commonly ectopically expressed in T-ALL, acting as oncogenes [87]. These proteins bind to E proteins and recognize the same E box elements. For this reason it has been proposed that ectopic expression of class B bHLH proteins interferes with the normal activity of the class A proteins during thymic development, and this would in part explain their oncogenic properties. For instance, in mouse cells the E2A/HEB transcriptional activity is repressed by heterodimerization with TAL1, which should be down-regulated before pTα expression, impacting on differentiation of the immature T-cells [82].

LYL1 (lymphoblastic leukemia-derived sequence 1) is not expressed in normal T-cell precursors. However, it is present in T-ALL as the result of a rare translocation with the *TCRB* gene driving its constitutive expression [88]. In addition, ectopic expression of LYL1 is also found in a subset of T-ALL patients without genetic abnormalities, defining an immature DN stage of the disease, with expression of stem and myeloid markers. It has been proposed that

the unfavorable clinical outcome associated with this subtype of T-ALL may be the consequence of the up-regulation of anti-apoptotic genes (e.g. *BCL2*) characteristic of early precursors [30]. Transgenic mice over-expressing LYL1 develop T-cell and B-cell leukemia over one year, characterized mainly by DP T-cells and mature B-cells, suggesting a role for LYL1 in the induction of T- and B-cell leukemia. It has also been proposed that excess of LYL1 blocks the dimerization of E2A, inhibiting the regulatory activity of E2A on the CD4 promoter and other E2A/HEB target genes that were found down-regulated [89].

The most commonly aberrantly expressed bHLH protein in childhood T-ALL is TAL1, a subject that will be further detailed below. Nevertheless, *TAL2* and *BHLH1* are also found up-regulated in T-ALL due to translocations that juxtapose their coding sequence with *TCRB* [90] or *TCRA* [91] loci, respectively. The homology of their bHLH domain with TAL1 domain suggests a common mechanism in T-ALL promotion. In this regard, ectopic expression of *TAL2* or *BHLH1* inhibited E2A-mediated transcription activation [91], suggesting that binding to E proteins might also be the way these proteins exert a leukemogenic effect.

The LIM-only domain proteins LMO1 and LMO2 do not interact directly with the DNA, but instead form stable transcriptional complexes with bHLH proteins like TAL1 and LYL1 [92]. This is consistent with the fact that the ectopic expression of these TF is frequently associated with deregulated expression of TAL1 (LMO1 and 2) and/or LYL1 (LMO2) [31]. Although translocations involving juxtaposition of LMO1/LMO2 to the TCRA/D locus occur in 9% of pediatric T-ALL, and cryptic deletions leading to loss of a negative regulatory region upstream of *LMO2* occur in only 4% of pediatric patients [41], the aberrant expression of LIM domain proteins can be found in up to 45% of T-ALL cases [30, 80]. *LMO1* and *LMO2* are expressed in hematopoietic stem cells and are normally down-regulated as hematopoietic precursors commit and progress through to the lymphoid lineage. In contrast, LMO2 is indispensable for erythroid development in mice, where it takes part on an oligomeric DNA-binding complex that includes GATA1, TAL1, LDB1 and E2A [93]. Notably, ectopic *LMO2* expression in CD34+ progenitor cells causes an incomplete inhibition at DN stages of thymocyte development and a severe block at the ISP stage, disrupting normal T-cell differentiation [94]. In mice, ectopic expression of LMO proteins gives rise to an immature

DN leukemia but with a long latency, suggesting that LMO proteins are necessary but not sufficient to cause leukemia in mouse models [95, 96]. In fact, other studies have shown that other aberrantly expressed oncogenes, such as TAL1, are needed to accelerate the leukemia [92, 97]. Moreover, TAL1 and LMO2 are commonly found co-expressed in human leukemia and the cases with ectopic expression of LMO proteins or TAL1 cluster together in what regards to gene expression profiles [30].

The involvement of LMO2 in the pathogenesis of T-ALL had important insights from an unfortunate event associated with retrovirus-based gene therapy trials for X-linked Severe Combined Immunodeficiency (SCID-X1). In several cases, the retroviral construction carrying the *IL2RG* gene integrated near the *LMO2* locus leading to aberrant activation of LMO2 in the retrovirally-transduced hematopoietic precursors transplanted into the patients. Consequently, four of the 11 patients receiving gene therapy developed a T-ALL-like disease [98], suggesting that, in agreement with its frequent ectopic expression in T-ALL patients, LMO2 is in fact able to drive leukemogenesis in humans.

Homeobox (HOX) genes are crucial regulators of body patterning and organogenesis during embryonic development. There are two classes of homeobox genes: Class I, which includes HOXA–D members; and Class II, which includes TLX1 and TLX3. Class I HOX genes encompass 39 genes dispersed in four *HOX* clusters (A–D). *HOXA* genes (A7, A9, A10, and A11) are expressed during the early stages of human T-cell development [99]. In T-ALL, HOXA genes, HOXA10 and HOXA9 can be found up-regulated due to translocations involving the *TCRB* enhancer in about 3% of patients. These cases are arrested in an immature DP stage before β selection [100]. Altogether, a subgroup of HOXA expressing T-ALL comprehends the TCR-HOXA, the MLL-translocated and CALM-AF10 translocated T-ALL, and a few cases without these rearrangements, suggesting a more general pathogenic role of HOXA deregulation in the genesis of T-ALL [30, 100].

Increased expression of the class II orphan homeobox HOX11 (TLX1) can result from the juxtaposition with promoter elements of *TCRA* and *B*, loss of negative regulatory elements upstream of the promoter of *TLX1* or other activating mechanisms in the absence of genetic rearrangements [80]. Overall, TLX1 is ectopically expressed in 5-10% of pediatric T-

ALL patients. These cases are characterized by an early cortical phenotype. Genes expressed in association with this cluster are related with increased cell proliferation and absence of ectopic expression of anti-apoptotic genes, making these cases more responsive to drug-induced programmed cell death, resulting therefore with a better clinical outcome [22, 30].

The HOX11L2 (TLX3) transcription factor is highly expressed in 20–25% of pediatric T-ALLs due to the translocation juxtaposing *TLX3* to the distal region of *BCL11B*, which is strongly expressed during T-cell differentiation [101]. Several other type of rearrangements with rare incidence involving *TLX3* have been reported [102]. The *in vivo* role of aberrant TLX3 expression in T-ALL remains to be clarified, but it is known that TLX3+ T-ALL cases cluster together with TLX1+, with analogous gene expression signatures and associated additional genetic events, suggesting similar mechanisms of action [30, 79, 100].

In an attempt to reveal underlying oncogenic alterations that are still unknown for 40% of pediatric T-ALL cases, a recent study identified two T-ALL clusters characterized by high expression of the homeodomain transcription factors NKX2-1/NKX2-2 or high expression of the MADS-box transcription factor MEF2C, both representing 20% of all T-ALL cases [31].

The T-ALL cases with rearrangements involving NKX2-1/NKX2-2 cluster together with TLX1-rearranged cases to form the ‘proliferative cluster’. This cluster is associated with high expression of proliferation genes, expression of CD1a, ectopic expression of NKX2-1 or NKX2-2 and cortical arrest [31]. Interestingly, the NK-like homeobox transcription factor NKX3-1 was previously shown to have an oncogenic role in T-ALL as a direct TAL1 target gene [103]. The second cluster identified includes cases with a very immature immunophenotype, with prevalent CD34 expression, and CD13 and/or CD33 myeloid markers [31], which appears to largely coincide with the previous description of ETP-ALL, and is characterized by several rearrangements that directly or indirectly activate MEF2C [33]. Moreover, this ‘immature cluster’ has a high expression of the transcription factors LYL1 and LMO2 that were shown to be regulated by MEF2C, thus concordant with previously described LYL1 signature (pro-T-ALL) [30]. MEF2C is a key regulator of normal lymphoid development activated by PU.1. It is expressed in normal pre-DN1 and DN1 human thymocytes and absent past the DN2 stage.

Therefore, MEF2C represents a newly found oncogenic transcription factor in T-ALL and down-regulation of MEF2C in a T-ALL cell line reactivates differentiation [31].

The TAL1/SCL oncogene

The *TAL1* gene product is a serine phosphoprotein and basic helix-loop-helix (bHLH) transcription factor known to regulate embryonic hematopoiesis. TAL1, also named SCL (stem cell leukemia), is the most commonly aberrantly expressed transcription factor in childhood T-ALL with increased transcript levels found in more than 60% of the patients, depending on the cohort considered [30, 104]. The blasts of the majority of these patients are devoid of obvious genetic alterations in TAL1 locus that could explain its up-regulation.

TAL1 was for the first time related to T-ALL with the discovery of recurrent non-random translocations juxtaposing the *TAL1* gene to the strong regulatory elements that drive the expression of the *TCRA/D* gene in the 14q11 locus. This translocation t(1;14)(p34;q11) is found only in 3% of children with T-ALL [105]. Other rare translocations targeting *TAL1* have also been described [106] but only in individual cases. The most common chromosomal alteration involving the *TAL1* locus is the already mentioned SIL-TAL1 gene fusion, also called Tal1d, which is found in 10-25% of the patients [107, 108]. Interestingly, only 25% of the patients harbor DNA rearrangements that activate TAL1 transcription [104]. Thus, the majority of the patients with high TAL1 levels lack cytogenetic or molecular evidence of gene variants affecting *TAL1.*

Gene expression signature studies revealed that high *TAL1* expression is associated with an arrest of the leukemic blasts in a late cortical double positive stage of thymocyte differentiation. This is evidenced by up-regulation of *LCK*, *TCRA*, *TCRB*, *CD2*, *CD6*, and *CD3E*. Moreover, *TAL1* over-expression was also associated with proto-oncogenes such as *CBFA2 (AML1)* and the *MYB*-related gene *MYBL2*, receptor genes such as *IL8R* and *CSFR1*, the anti-apoptotic gene *BCL2A1*, and frequent co-expression of the transcription factors LMO1/LMO2 [30]. The fact that TAL1+ blasts up-regulate anti-apoptotic proteins normally induced by TCR signaling in the late cortical stage offers a plausible explanation for the less favorable

outcome observed in the TAL1+ subgroup in some studies. In fact, some of these patients demonstrate also an increased risk of failure to achieve complete remission, hinting for a possible role of TAL1 mediated genetic program in primary drug resistance [30]. However, in latter reports, high TAL1 levels were again related with immunophenotypically advanced T-ALL cases, but were instead associated with a good outcome [109, 110].

Despite the strong evidence for TAL1 involvement in T-ALL ontogeny, the first mouse models generated to access *Tal1* oncogenic potential failed to develop disease, either by using mouse bone marrow reconstitution models with early hematopoietic progenitors expressing *Tal1* [111] or transgenic mice models in which *Tal1* expression was directed to the T cell lineage using the *CD2* enhancer [92, 112]. In these early reports it was suggested that additional genetic abnormalities are required besides *Tal1* ectopic expression to originate leukemogenesis. In fact, there are several reports of additional abnormalities, frequent in leukemia and other cancers, that accelerate the onset of Tal1 induced murine leukemia, such as casein kinase IIα (CKIIα) over-expression [113], p53 deficiency [114], N-ras over-expression [115], or loss of the *Ink4* locus [116]. Additionally, mice with *Tal1* expression regulated by the *CD2* enhancer did not develop leukemia unless they were crossed with CD2-lmo2 transgenic mice resulting in late onset T-cell leukemia [92]. Despite the strong evidence for a collaborative effect of aberrant expression of TAL1 and LMO proteins, other mouse models show that Tal1 is able to transform thymocytes without ectopic activation of Lmo2 or Lmo1, for instance by interfering with E47/HEB function [117]. This is also true in the human disease, given that even if the majority of TAL1+ T-ALL cases also have high expression of LMO2, not all human T-ALL patients who express TAL1 express LMO1 or LMO2 [30]. Moreover, it was shown that over-expression of *Tal1* alone can be transforming in transgenic mice were *Tal1* expression is under the control of the *Lck* promoter that directs high level expression of the transgene to thymocytes [113, 114]. In fact these mice died of clonal T lymphoblastic lymphoma or leukemia, evocative of the pathology of the human disease, but with a low penetrance, since only 28% of the animals developed disease within one year [113, 117]. Importantly, *in vitro* primary lymphocyte cultures showed partial independence from exogenous growth stimuli and increased resistance to apoptosis [114]. The thymocytes

of the *Lck-Tal1* transgenic mice presented a developmental block at the DN2 stage. The reduction in *Rag2* and *pre-Tα* observed in this model might explain, at least in part, the developmental abnormalities observed [117].

The combination of the *Lck-Tal1* transgenic mice with a heterozygous background for E2A or HEB, provokes a drastic decrease in the absolute numbers of DP and SP4 thymocytes, with an increase in ISP population. These thymocytes present decreased expression of TCRβ chain, CD4, CD5, pre-Tα and Rag2 [117], all known targets of E-proteins in mice [83, 118]. This general decrease in important developmental genes is higher in the context of E-protein heterozygosity in comparison with Tal1 expression alone, suggesting that Tal1 interference with the E47/HEB heterodimer may be responsible for a more severe phenotype. This is translated in a faster disease development with increased penetrance (80% in the case of *HEB* background), providing evidence that E2A or HEB haploinsufficiency collaborates with Tal1 to induce leukemia [117]. As in *Lck-Tal1* transgenic mice [113], the tumors developed in the Tal1/E-protein mice were blocked at various stages of thymocyte development, but with increased incidence of tumors with an immature DN phenotype [117].

How does TAL1 predispose thymocytes to leukemia? By one hand, aberrant expression of TAL1 might lead to the formation of complexes with E-box proteins that are not normally present in T-cells, leading to abnormal activation of genes in the thymocytes, such as those associated with stemness. Alternatively or concomitantly, the formation of those complexes might decrease the availability of the E-proteins to transcribe their normal target genes, mainly involved in differentiation [113, 117, 118]. Such dominant-negative mechanism of action of TAL1 is supported by the fact that a Tal1 construct lacking the transactivation domain is still able to collaborate with Lmo1 to cause aggressive T-cell malignancies [97]. Also, transgenic mice expressing a DNA binding mutant of Tal1 but capable of forming stable complexes with E2A proteins, still develop disease [119]. Also supporting the model of TAL1-mediated reduction of transcriptional activity is the fact that histone deacetylase inhibitors induce apoptosis of TAL1+ tumors [117, 120]. Moreover, TAL1 is also capable of repressing transcription of target genes by association with co-repressor factors such as mSin3A, HDAC1 [117], HP1α, Suv39h1 [121] and LSD1 [122], proteins with deacetylase, methyltransferase

and heterochromatin activities. Alternatively, several reports demonstrate that TAL1 can also function as an activator of gene expression, further detailed below. Furthermore, TAL1 can be acetylated by the co-activator p300 and the p300/CBP-associated factor P/CAF, increasing DNA binding capacity of TAL1 while preventing its interaction with the transcriptional co-repressor mSin3A [123]. Moreover, ChIP (Chromatin immunoprecipitation) studies for direct targeting confirmed that TAL1 may function both as repressor and activator of transcription, with the majority of promoters occupied by TAL1 being bound also by E2A and HEB [124]. Thus, the current view is that TAL1 is a bifunctional transcriptional regulator, which is at the onset of a complex transcriptional network on T-cell progenitors that disrupts normal T-cell homeostasis and contributes to leukemogenesis.

Physiological roles of TAL1

TAL1 is necessary for hematopoietic fate commitment during embryogenesis and is indispensable for specification of HSCs from mesoderm and maintenance of immature progenitors [125]. Mice knockout for *Tal1* die during embryonic development, between day 8.5 and 10.5, due to absence of hematopoiesis. Loss of Tal1 in the embryo mimics that of the loss of Lmo2 or Gata1 in the erythroid lineage, but is additionally characterized by a reduced capacity in the myeloid differentiation [126]. Importantly, the DNA binding activity of Tal1 is dispensable for hematopoietic commitment, but essential for maturation of erythroid and megakaryocytic precursors [127]. Also, Tal1 is essential for embryonic angiogenesis and thus indispensable for both cell types (blood and vascular cells) descending from the hemangioblast [128].

In the adult mouse, Tal1 is expressed in HSCs and multipotent progenitors, and in erythroid and megakaryocytic lineages. Despite being essential for the genesis of HSCs in the embryo, Tal1 is not necessary for HSCs engraftment, self-renewal and differentiation into myeloid and lymphoid lineages, being only indispensable for erythroid and megakaryocytic differentiation in adult animals [129]. Surprisingly, in the absence of *Tal1* in the adult mice, HSCs are still capable of surviving, self-renewing and repopulating the hematopoietic system. Thus, Tal1 belongs to the category of transcription factors necessary for HSCs genesis, but

not required for later long-term repopulating activity and multipotency of HSCs [129]. Inside the adult HSC compartment, *Tal1* is more expressed in cells with an LT-HSC phenotype compared with ST-HSCs and progenitors [130]. Tal1 is important in maintaining dormant adult HSCs but also to preserve reconstitution potential of HSCs, mediated by regulation of p21 and Id1 expression levels [130]. A different study showed that mouse cells defective in Tal1 present severe multilineage defects in repopulation capacity by ST-HSCs without their capacity self-renewal being affected [131].

In addition to its role in embryonic angiogenesis, TAL1 also takes part in the adult formation of new vessels, including tumor vasculature. Despite being undetectable in the adult quiescent endothelium, in an *in vitro* angiogenesis model, TAL1 ectopic expression enhanced migration of primary endothelial cells, formation of capillary-like structures and vascularization, and enlargement of capillary lumens. Thus, this demonstrated a previously unknown role for TAL1 in regulation of postnatal vascular remodeling [132].

Analysis of *TAL1* mRNA expression during human adult hematopoietic differentiation shows that *TAL1* expression is silenced earlier [94, 133] than in the mouse [80], where *Tal1* is still detected in the thymic DN3 population (CD44-CD25+). In humans, *TAL1* expression is detected in all hematopoietic populations with erythroid potential (HSCs, multipotential progenitors, CMP, megakaryocyte/erythrocyte progenitors, and nucleated erythroid cells), but also lymphoid and myeloid-restricted progenitors (Figure 1.3). TAL1 is found in $CD34^{+}CD38^{-}$ HSCs from cord blood and bone marrow, and also in the $CD34^{+}CD38^{+}$ population, which includes a mixture of lymphoid, myeloid, and erythroid progenitors. On the contrary, on pro-B cells, early thymic progenitors, myeloid precursors and all mature cells of the non-erythroid lineages, *TAL1* is undetectable [133]. From mouse studies, it was suggested that *Tal1* could also influence lymphoid and myeloid lineage commitment [129, 131]. If that is the case in humans, it makes sense that TAL1 is expressed in CLP and GMP, although down-regulated compared with HSCs [133]. In fact, decreased expression of *TAL1* in human CD34+ cells compromised the erythroid and myeloid lineage, whereas lymphoid B-cell development was normal. Notably, *in vivo* repopulation studies showed dramatically low

levels of human cells of all lineages including the B-lymphoid lineage, thus implicating TAL1 in early commitment of adult hematopoietic stem cells in humans [134].

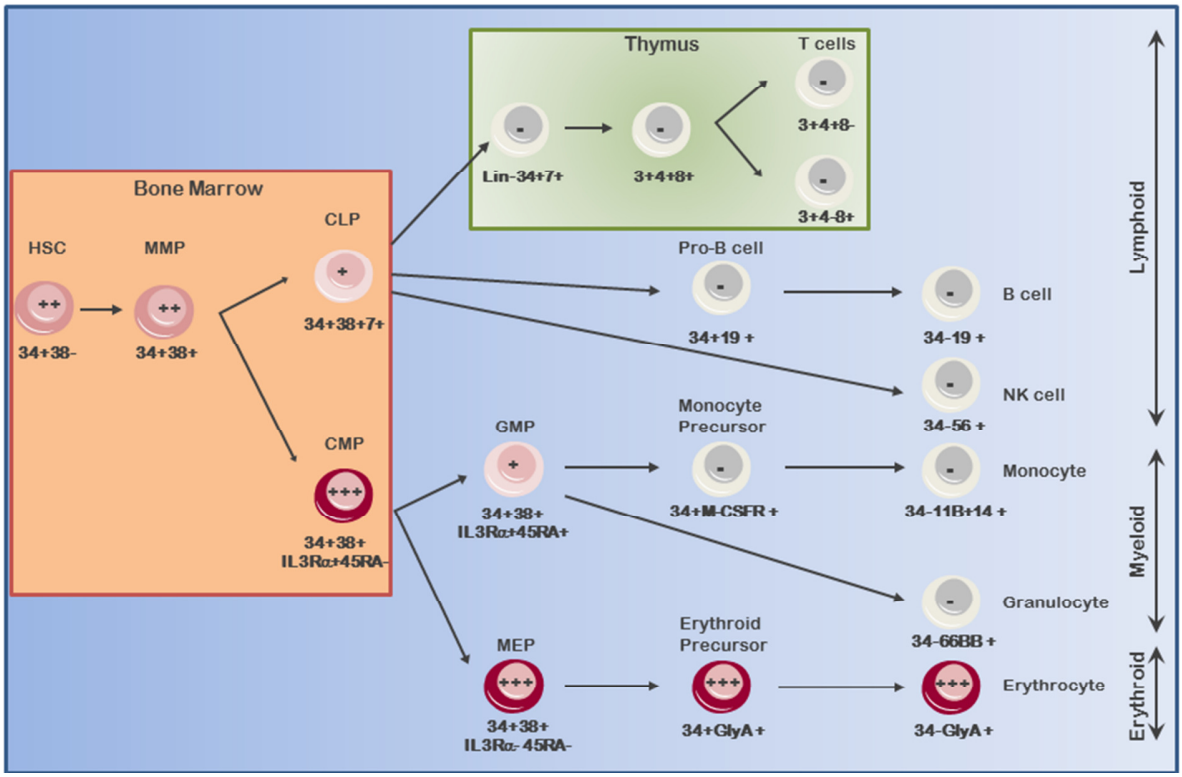


Figure 1.3 **– TAL1 expression during human hematopoiesis.**

The relative mRNA expression levels of TAL1 are indicated, based on [133], by increasing color intensity and by symbols (–, +, ++, +++). Abbreviations: CLP, common lymphoid progenitor; CMP, common myeloid progenitor; GMP, granulocyte/monocyte progenitor; HSC, hematopoietic stem cell; IL, interleukin; M-CSFR, monocyte colony stimulating factor receptor; MEP, megakaryocyte/erythroid progenitor; NK, natural killer.

The importance of TAL1 in erythroid differentiation is very well established. During mouse erythropoiesis, Tal1 regulates expression of components of the red blood cell membrane, such as Glycophorin A (*GPA*) [135] and protein 4.2 (P4.2) [136], and ubiquitinylation machinery components, such as the E2-ubiquitin conjugase UBE2H [137]. Their promoter activation is dependent on the assembly of a multifactorial complex containing TAL1, E47, Sp1, Ldb1, LMO2 and GATA1 transcription factors [135, 136].

TAL1 is also expressed in the CNS of the embryonic and adult mouse, in neurons of the lateral and caudal thalamic region, midbrain and hindbrain. Conditional deletion of *Tal1* in neuronal precursor cells leads to premature death, growth retardation and altered motor capacities, resulting from an affected brain morphology and abnormal neuronal development [137]. These observations underline the critical role of TAL1 in both neural and hematopoietic cell development.

TAL1 target genes

The understanding of the mechanism by which TAL1 recognizes binding motifs in its direct target genes may help comprehending the differential outcomes of TAL1 activity depending on the cell lineage: e.g. differentiation in erythrocytes versus transformation in T-lymphocytes. As stated above, TAL1 binds to the DNA in E-box motifs (CANNTG) through heterodimerization with an E-protein. The binding occurs preferentially to the E-box CAGATG (represented by the consensus sequence AACAGATGGT), rather than the E-box CAGGTG that is favored by E-protein homodimers [138]. Moreover, the E-box constitution might not always be the most important determinant in TAL1 recognition, as it may be attracted to other regions by other DNA-binding transcription factors. These includes GATA3 in leukemic T cells [139] and SP1 [140] or GATA1 [93] in erythroid cells.

In erythroid development TAL1 forms a pentameric DNA-binding complex between the heterodimer of TAL1 and E2A and the zinc finger proteins GATA1, that are bridged by LMO2 and the LIM-binding protein LDB1. This multiprotein complex binds to a bipartite DNA motif comprising an E-box, CAGGTG, followed ~9 bp downstream by a GATA site, establishing a transcriptional activating complex [93]. GATA1 and TAL1 occupancy is mainly related with the transcriptional activation of the genes, but also associated to repression [141, 142]. This composite E-box/GATA motif, found in erythroid cells and frequently associated with TAL1 occupancy, was not initially detected in T-ALL cells [124] as consequence of a different preference for composite DNA motifs in T-ALL [143]. Thus, TAL1 oncogenic role can be in part rendered by genomic binding selectivity in different cellular environments. In T-ALL cells, TAL1 binding regions frequently overlap with the LMO1/2-, GATA3-, and RUNX1-enriched

regions [144]. In fact, in T-ALL cells TAL1 preferentially binds a combination of DNA motifs that includes the E-box variant (CAGGTG) favored by E-protein homodimers [138], and also RUNX- and ETS-binding sites. Nevertheless, the E-box/GATA composite motif is still over-represented in TAL1 binding sites in both erythroid and T-ALL cells, reinforcing the importance of GATA factors in targeting TAL1 binding to specific sites. The alternative binding motive found in T-ALL allows the access of TAL1 to RUNX and ETS transcriptional regulatory networks in these cells, providing a possible molecular framework for TAL1-mediated leukemogenesis [143]. In addition, it was found that RUNX1/3 and ETS1 are required for TAL1 binding to genomic loci in T-ALL cells. Interestingly, TAL1 directly activates the expression of *RUNX1*, *ETS1* [143] and *GATA3* [144] genes in blasts from T-ALL patients and RUNX1 is part of the TAL1+ gene signature described for T-ALL patients [30]. Importantly, these transcription factors are essential for the maturation of SP thymocytes, and may contribute to the DP stage differentiation block associated with TAL1. Curiously, GATA3 and RUNX1 seem to also regulate positively *TAL1* expression, which suggests the existence of a possible positive feedback loop sustaining an aberrant gene expression profile in T-ALL [144].

Discovering TAL1 direct transcriptional targets has been complicated by the fact that TAL1-binding sites are located mainly away from promoter regions of known genes, mostly within introns and intergenic regions, in distal regulatory elements such as enhancers [143]. The TAL1 target genes identified in ChIP-seq studies code for proteins mainly involved in T-cell differentiation, proliferation, morphology, activation, apoptosis and also recognized tumor suppressor genes that are repressed by TAL1 aberrant expression in T-ALL [124, 143, 144]. In fact, in T-ALL cells TAL1 binding peaks were found in the following genes: *TRAF3*, *RAB40B* and *EPHB1* that are activated by TAL1, and *PTPRU*, *TTC3* and *RPS3A* that are repressed [124]. In another study, TAL1 was shown to require RUNX1/3 and/or ETS1 to bind to the promoter of T-cell marker genes (e.g. AIOLOS/IKZF3, TOX, CCR9, CD69, TCRBV, CDK6) and pro- and anti-apoptotic genes (e.g. PMAIP1/NOXA,PLK3 AND CD226) [143].

The studies based on TAL1 genome binding information and/or on TAL1 knockdown consequences on the gene expression aim to define possible TAL1 target genes. Most of the

candidate genes still need to be validated in different cellular and organism contexts. We will mention the few TAL1 target genes identified and clearly validated in human cells (Table 1.2).

Table 1.2 - **TAL1 target genes identified and validated in human cells, the context of regulation and the functional effect of TAL1 in the expression of the gene.**

Gene	Context	TAL1 effect
c-KIT	Hematopoiesis	Up-regulation
RUNX1	T-ALL	Up-regulation
ERG	T-ALL	Up-regulation
RALDH2	T-ALL	Up-regulation
NKX3.1	T-ALL	Up-regulation
TALLA1	T-ALL	Up-regulation
NFKB	T-ALL	Down-regulation
MYB	T-ALL	Up-regulation
TRIB2	T-ALL	Up-regulation
CDK6	T-cell differentiation	Up-regulation
CD69	T-cell differentiation	Down-regulation
TCRA/G	T-cell differentiation	Down-regulation
pTα	T-cell differentiation	Down-regulation
P16	T-cell differentiation	Down-regulation

In hematopoietic precursors, a multimeric complex containing TAL1, LMO2, GATA1/GATA2, E2A and LDB1 is attracted to DNA via a specificity protein 1 (Sp1) motif to the *c-KIT* promoter. KIT is a tyrosine kinase receptor essential for normal hematopoietic development and is activated by TAL1 [140]. In the context of genes important for T-cell differentiation, TAL1 was shown to bind to E47 and HEB to block activation of the TCRα enhancer [145]. Moreover, enforced expression of TAL1 and LMO1 in mouse and human thymic progenitors represses pre-Tα chain expression [118, 146]. Furthermore, TAL1 can bind the E boxes in the *p16* promoter, and functionally suppress its activity, linking again TAL1 to a role in cell proliferation [146]. TAL1 can also bind to the intronic region of *CDK6*, a regulator of T-cell differentiation, which is up-regulated by TAL1 in T-ALL cells. In the case of the *CD69* gene, one of the earliest inducible cell surface glycoproteins acquired during lymphoid activation, TAL1 binds to a promoter region in erythroid and lymphoid cells, but induces down-regulation of its expression only in T-ALL cells [143].

In T-ALL, TAL1, LMO and GATA3 act as cofactors for *RALDH2* transcriptional activation, a gene that codes the enzyme that synthesizes retinoic acid (RA). A GATA site present in the second intron is essential for the transcriptional activation, which is independent of the DNA binding capacity of TAL1 [139]. The same complex of the three transcription factors also activates the expression of *TALLA1*, a highly specific marker of T-ALL [147]. A TAL1–LMO–LDB1 complex is recruited by GATA3 to activate the NKX3.1 gene. The recruitment of the complex is associated with suppression of HP1-α binding and opening of chromatin, suppressing *NKX3.1* gene repression. In T-ALL context, NKX3.1, a member of the NKX family of homeobox genes, promotes proliferation and can partially restore proliferation capacities upon TAL1 knockdown. Moreover, NKX3.1 directly regulates miR-17-92 [103], an oncomiR in T-ALL [148]. As in the case of the RALDH2 gene, the regulatory sequences responsible for TAL1 recruitment on the NKX3.1 gene are also not conserved between human and mouse [103]. Furthermore, as previously found in mouse embryonic cells, TAL1 directly activates the expression RUNX1 and ETS1 genes in T-ALL blasts [148]. In primary human T-ALL cells the expression of ERG, another ETS family member, is mediated by the binding of TAL1, LMO2, LYL1, FLI1, GATA3, but also ERG, to an enhancer that is active in stem cells [149]. Recently the TRIB2 gene has been identified as a target gene up-regulated by TAL1 in T-ALL, but negatively regulated by the E2A/HEB dimer. This encodes for an adaptor protein with a possible role as negative regulator of signaling pathways and appears to be required for survival of T-ALL cells [144]. Moreover, the oncogenic transcription factor, MYB, with a known involvement in malignant hematopoiesis [42] was identified as a direct target of TAL1 that forms a transcriptional complex together with GATA3 and RUNX1. Many of downstream targets of MYB are also controlled by TAL1. Therefore, these two transcription factors activate an overlapping oncogenic network to work in concert to maintain a deregulated gene expression program in T-ALL [144].

Contrary to the genes already mentioned, *NFKB1* gene expression was reported to be repressed by TAL1. In T-ALL cell lines, the expression of p50, encoded by *NFKB1* gene, is reduced and that allows the activation of the atypical p65:cRel complex rather than the classic p50:p65 dimer. TAL1 and LMO1 bind to the promoter and recruit HDAC1 to repress

the transcription of the *NFKB1* gene, shutting down NF-κB-dependent transcriptional program in T-ALL [150].

In summary, TAL1 appears to be on the top of a transcriptional network that drives the expression of genes involved in abnormal proliferation, differentiation and survival in transformed thymocytes.

Gene structure and regulation

In the human genome, *TAL1* gene is located in the small arm of chromosome 1 (1p32). It is composed of eight exons distributed over 16 kb in-between the *SIL* and the *PDZK1IP1* loci [151]. Expression of *TAL1* is tightly temporal and spatially regulated in different lineages and stages by distinct promoters and distal enhancer elements.

The promoter Ia and Ib are utilized in a lineage-restricted manner. The promoter Ia is active in erythroid, megakaryocytic and mast cells. The promoter Ib is active in primitive myeloid and mast cells [152] and it is controlled by GATA1 and Sp1. On the contrary, the promoter IV, located within the fourth exon, is specifically active in the T-cell lineage, being responsible for transcript production in T-ALL cases [153] and cell lines [154].

The distal enhancers also display lineage specificity in the *Tal1* locus [155]. The expression of TAL1 during mesoderm differentiation is associated with activation of the 5' proximal enhancer (−10 Kb in human) and the stem cell enhancer (19/+20/+21) [156, 157]. This last enhancer is active in hematopoietic stem and progenitor cells and endothelium. Additionally, the proximal enhancer also drives *Tal1* expression in hematopoietic progenitors and endothelium, and it is bound by Ets family transcription factors, including Fli-1 and Elf-1 [157]. On the other hand, the distal enhancer (+51 Kb in human) is transcriptionally active in erythroid cells, and thus called erythroid enhancer [158]. In erythroid cells an intrachromatin loop, which is not found in T-ALL cells, mediates the proximity of the +51 enhancer with promoter I [154]. This distinction is apparently due to the CCCTC-binding factor (CTCF), which mediates differently chromatin loops facilitating enhancer/promoter interaction of the *TAL1* locus in erythroid cells but preventing the same interaction in human T-cell leukemia [154]. On the other hand, in human T-ALL cells, another chromosomal interaction brings into close

proximity a T-cell-specific DNA regulatory element with enhancer characteristics on chromosome 16 (*TIL16* element) and *TAL1* promoter I. This interaction is mediated by the proto-oncoprotein c-Maf, being very important for maintenance of aberrant *TAL1* gene expression [154]. Other enhancers have been described in the *TAL1* locus [159], but apparently none of them is transcriptionally active in T-ALL cells [154].

Additionally, there are two known repressor elements in the *TAL1* locus. The first was found in the 3' UTR of the gene [160] and it is responsible for restricting promoter IV usage in erythroid cells. PU.1 is necessary but not sufficient to maintain activity of this repressor element [161]. An additional region (−13) showed repressor activity on both erythroid and T-ALL cell lines, possibly mediated by the ETV6/7 proteins [162]. Besides promoters and enhancers, the genomic region containing the entire *TAL1* regulatory elements span ~88 kb in the human genome and it is flanked by CTCF-bound elements (at +57 and -31) (Figure 1.4) [163].

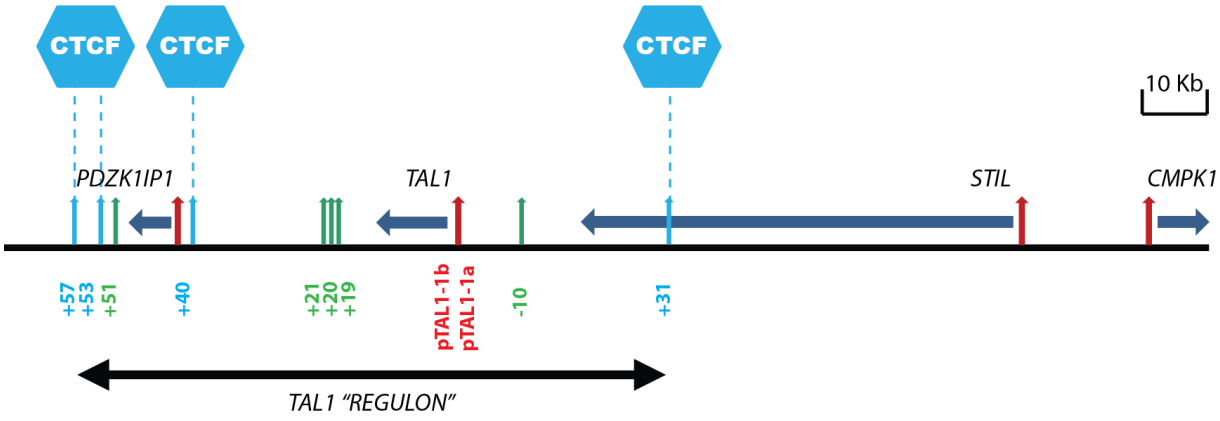


Figure 1.4 **– Diagram of the organization of the human TAL1 regulon.**

The cis-regulatory elements of the human *TAL1* locus span ~88 kbs in the small arm of chromosome 1, the so-called TAL1 'regulon' (horizontal black line with Arrowheads); adapted from [163]. TAL1 and its flanking genes PDZK1IP1, STIL and CMPK1 transcriptional orientation is depicted by the horizontal blue arrows; locations of the promoters are shown with vertical red arrows. The principal characterized enhancers are also depicted, with vertical green arrows (the +19/+20/+21 stem cell enhancer, the +51 erythroid enhancer and the -10 enhancer). CTCF-bound elements (+57, +53, +40 and -31) [162] are shown with vertical blue arrows and labeled with CTCF in the blue hexagon.

Recently a non-coding RNA named ncRNA-a3, belonging to the class of activating RNAs that activate their neighboring genes using a cis-mediated mechanism, was found downstream of the *TAL1* locus in the opposite strand. Depletion of ncRNA-a3 in breast cancer cells resulted in a potent reduction of *TAL1* expression [164]. A specific and robust DNA looping is formed between ncRNA-a3 and TAL1 [165]. It remains to be clarified if this mechanism of TAL1 regulation occurs during normal development and/or transformation.

Importantly, the mechanism of aberrant activation of *TAL1* in the majority of T-ALL patients who lack chromosomal rearrangements remains unknown. Whether deregulation of enhancer/promoter interactions, epigenetic alterations or aberrant trans-acting mechanisms lead to a disease-causing regulatory variant is still a mystery.

Non-coding RNAs

Non-coding RNAs (ncRNAs) are RNA transcripts that are not translated into proteins. Nonetheless, they have the capacity to regulate the transcription, stability or translation of protein-coding genes. Their occurrence in the metazoan genome increases with developmental complexity [166].

So far, the most studied ncRNAs are microRNAs but several other classes of ncRNAs with variable lengths and characteristics have been experimentally identified [167] (reviewed in [167-169]). Based on recent annotation by the GENCODE project [169] (version 20, April 2014 - Ensembl 76) only 34% of the 58.688 annotated genes are protein-coding, encompassing 1.2% of the genome. Long non-coding RNA, small non-coding RNA genes and pseudogenes represent 24.7%, 16.2% and 24.5% of the total number of annotated human genes. This indicates that the number of ncRNAs is potentially much higher than that of protein-coding genes. In fact, the human genome is extensively transcribed, being 75% of bases represented in at least one primary transcript, whereas only 3% is transcribed into protein-coding mRNAs [169, 170].

MicroRNAs: definition and biogenesis

The most studied class of ncRNAs comprehends small, 19-22 nt long, non-coding single stranded RNAs that provide post-transcriptional control of gene expression, named microRNAs (miRNAs). In mammals, miRNAs function mostly as endogenous translational repressors of protein-coding genes through sequence-specific binding to the 3'-untranslated region (3'UTR) of a target messenger RNA [171]. At the beginning of this project the number of miRNA loci known in humans was 533 [172]. Today, about 1500 human miRNAs have been reported (http://microrna.sanger.ac.uk/sequences/) and are thought to regulate more than 30% of protein-coding genes [173, 174].

MicroRNAs were first described in 1993 in *C. elegans* by Rosalin Lee and colleagues [175]. The authors found two small *lin-4* transcripts with 22 and 62 nt that enclosed sequences complementary to the 3'UTR of Lin-14 mRNA. Importantly, they showed that Lin-4 negatively regulated the level of LIN-14 protein despite the fact that its sequence did not suggest to be protein-coding. This suggested to the authors that they were revealing an antisense RNA-RNA interaction mechanism of translation regulation [175]. This class of transcripts was initially named small temporal RNAs (stRNAs) due to the roles in the developmental transitions in *C. elegans*, but after reports of several other genes for small noncoding RNAs in other metazoan classes, invertebrates, mammals and even humans, the term 'microRNA' was born to name the stRNAs and all the other small RNAs with analogous characteristics but, at the time, unknown functions [176, 177].

MicroRNA genes are generally transcribed by RNA pol II as long primary 5' capped and polyadenylated transcripts (pri-miRNA), with one or more hairpins (Figure 1.5). These hairpin structures are cleaved at the base by a nuclear enzyme complex, which includes the RNase III Drosha and the dsRNA-binding protein (dsRBPs) DGCR8, releasing a 60–110-nt hairpin precursor (pre-miRNA). The pre-miRNA is transported to the cytoplasm by Exportin 5 where it is further processed by another RNase III – DICER – assisted by transactivation-responsive (TAR) RNA-binding protein (TRBP) in mammals [178]. The result is a 19–22-nt double-stranded miRNA product. One of the strands is then incorporated as mature miRNA in the effector RNA Induced Silencing Complex (RISC). MicroRNAs can next exert their posttranscriptional activity, mostly as repressors of translation or, under some circumstances, alteration of mRNA stability. The final output is typically a reduction in the protein levels of the target gene [179]. The RISC is a protein complex that includes the highly conserved Argonaut proteins. The core component of this complex consists of an Argonaut protein together with a single-stranded small RNA. In humans, Argonaut family can be divided into two subfamilies (Ago and Piwi) based on sequence similarities [180]. The Ago subfamily in mammals is composed by four ubiquitously expressed members, Ago1-4, which participate in miRNA-mediated repression. Only Ago2 (or 'Slicer') is involved in the RNA interference (RNAi) mechanism by endonucleolytically cleaving the mRNA target [180].

MiRNA loci can be found encoded in introns or exons of protein coding genes or within either the introns or exons of noncoding RNA genes, as well in intergenic regions. More than half of mammalian miRNAs genes are intronic, i.e. found within introns of either protein-coding or noncoding transcripts. The majority are encoded within introns of protein-coding genes (~40% of all miRNA loci), and only 10% are located within introns of long ncRNA transcripts. Around 10% of miRNA genes are found in exons of long non-protein-coding transcripts [181]. Interestingly, more than 50% of mammalian miRNA loci are clustered, therefore found in close proximity (<50 kb) to other miRNAs and transcribed generally from a single polycistronic transcription unit. In humans, 42% of miRNA genes are found in clusters that are less than 3000 nt apart and are regulated from a common promoter, transcribed into polycistronic units with multiple discrete loops further processed into mature miRNAs [182]. Moreover, the majority of human miRNA loci are located within intronic regions and their expression is coordinated with their host gene mRNA, implying that they also generally derive from a common transcript [183]. In these cases, expression levels of the coding gene are directly linked to expression levels of the miRNA, and vice versa. For example, the miR-106b-25 cluster, found aberrantly over-expressed in prostate cancer along with its host gene *MCM7,* cooperate in cellular transformation both *in vitro* and *in vivo* [184]. On the other hand, we can also find examples of miRNA expression antagonizing, rather than supporting, the function of their host transcripts. For instance, miR-218 is encoded within a *Slit* gene intron, and it was found to negatively regulate the expression of Robo1 and Robo2 that are receptors for the SLIT ligand, thus antagonizing its host gene functions [185].

Our knowledge of miRNA biology has been recently challenged by evidences of unexpected pathways of miRNA biogenesis that differ from the canonical described above. The first identified non-canonical pathway for miRNA biogenesis comprehends a bypass of the Drosha–DGCR8 processing step and occurs during processing of very short introns (mirtrons) as a result of splicing and debranching (Figure 1.5). The pre-miRNA is instead generated through splicing of a host mRNA that releases a lariat. The lariat refolds into a short stem–loop structure that resembles a pre-miRNA [178, 186]. In addition, the generation of miRNAs from snoRNAs, which are small (60-300 nt) nucleolar RNAs encoded in

the introns of proteins involved in ribosome synthesis or translation, has also been described. This was discovered when the analysis of the RNA fraction immunoprecipitated with human Ago complexes revealed a number snoRNA fragments of miRNA size [187]. Similarly, it was found that tRNAs produce a huge variety of small RNA fragments *in vivo* and that one tRNA fragment, cloned from mature human B cells, results from DICER1 cleavage and can function as a miRNA in an Ago-dependent way [188].

MicroRNA target recognition

Target recognition is one of the major challenges in the miRNA field, leaving elusive the functions of most mammalian miRNAs. In fact, the region of complementarity between a miRNA and its target mRNA, referred to as the 'seed' region, encompasses only the nucleotides 2–7 from the 5'-end of the miRNA. The rest of the miRNA sequence has several levels of imperfect complementarity, depending on their target mRNAs, making accurate computational prediction of target sites very difficult [189]. Many, if not most, protein-coding transcripts are potential targets for miRNA regulation [173].

Site directed mutagenesis experiments [171] have shown that there are two types of miRNA target sites in animals [190], one showing perfect complementarity to the 5'end of the miRNA without requirement of significant further base pairings and other showing imperfect 5' matches compensated via extended base pairings with the 3' end of the miRNA. In mammals, contrary to flies, many predicted targets and the corresponding miRNAs are manifestly expressed in the same tissue [189, 191], but the mRNA of the targets is expressed at significantly lower levels compared with most other tissues and the endogenous expression of the miRNAs is negatively correlated with the mRNA levels of their targets [171, 190].

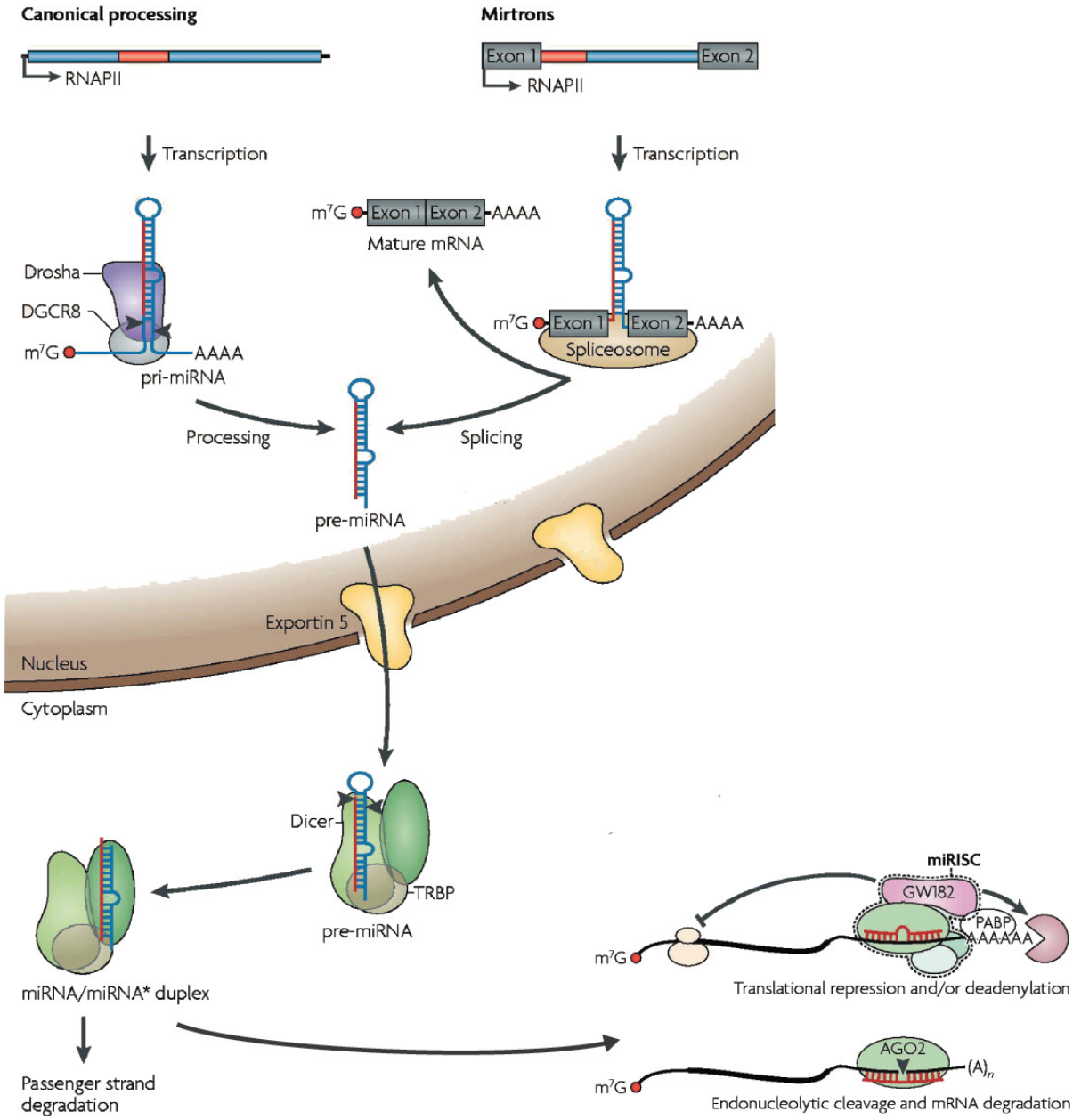


Figure 1.5 – **MicroRNA biogenesis pathways.**

MicroRNAs are processed by RNA polymerase II from transcripts of independent genes or from introns of protein-coding genes (Figure adapted from [178]). In the canonical pathway, primary precursor (pri-miRNA) is processed by the Drosha–DGCR8 into a ~70-nucleotide precursor hairpin (pre-miRNA), which is exported to the cytoplasm. Some pre-miRNAs are produced from very short introns (mirtrons) as a result of splicing and debranching, thereby bypassing this step. In either case, cleavage by Dicer, assisted by TRBP, in the cytoplasm yields a ~20-bp miRNA/miRNA* duplex. Following processing, one strand of the duplex (the guide strand) is preferentially incorporated into the miRNA-induced silencing complex (miRISC), whereas the other strand is released and degraded. MiRNAs exert post-transcriptional control of gene expression either by promoting the transcript degradation or by repressing translation.

There are four main types of seed-matched sites found in the transcripts modulated upon miRNA introduction [173], one 6mer, two 7mers, and one 8mer, all selectively conserved (Figure 1.6). These follow a hierarchy of site efficacy: 8mer > 7mer-m8 > 7mer-A1 > 6mer [189]. The same is true when examining protein levels [192, 193]. Computational studies predict a widespread targeting of mammalian miRNAs: on average 300 targets per miRNA family conserved through vertebrates, if considered conserved 7-8 mer sites, a number that can increase to more than 400 if 6-mer sites are also considered [191].

The canonical 7-8mer sites are clearly important but there are other determinants besides seed pairing with important roles for target recognition [189]. In fact, proximity of target sites to co-expressed miRNAs is an important determinant since it was shown that closely spaced target sites contribute more to repression than the independent contributions of two single sites [189]. This cooperative effect is true for target sites of the same miRNA or different miRNA targeting the same transcript, and repression reaches stronger levels when spacing between the miRNA target sites is between 8 to 40 nt. In addition, the so called 3'-compensatory sites (additional paring with the 3' region of the miRNA) are effective determinants for targeting, in particularly for 7mer-m8 sites. Like seed pairing, 3' matching is sensitive to position, with pairing at the 3' core (positions 13–16) being more important for efficacy than pairing to other positions [189]. Despite allowing prediction of target sites with great specificity, mammals extensive 3' pairing is atypical and only slightly more effective [194]. One rare example of biological targeting with extensive 3' pairing in mammals is the miR-196 site in *HoxB8* [195]. It is also important to note that functional target sites are embedded in a highly enriched A and U context of nucleotides immediately flanking the site. The 5'UTRs and open reading frames (ORFs) can have functional miRNA targeting sites but these are less functional than the ones in the 3'UTR, both for mRNA destabilization [189] and translational inhibition [193]. Moreover, position of the target sequences within the 3'UTR is not innocent, as UTR quartiles near the ORF and near the poly(A) tail harbor more effective targeting than the two central quartiles. This effect is more important for UTRs higher than 1300 nt and importantly, the distance to the Stop Codon has to be higher than 15 nt [189].

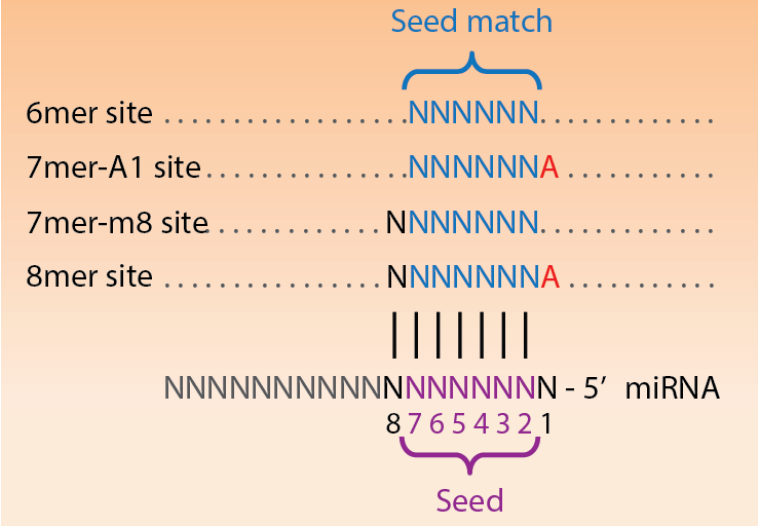


Figure 1.6 **– Main types of seed-matched sites found in the mammalian transcripts.**

The 6mer site perfectly matches the 6-nt miRNA seed (miRNA nucleotides 2–7); the 7mer-m8 site comprises the seed plus a match to miRNA nucleotide 8 (miRNA nucleotides 2–8); the 7mer-A1 site comprises the seed supplemented by an Adenine at target position 1; finally the 8mer site comprises a match to miRNA nucleotide 8, the miRNA seed and the A at position 1 [189, 191].

The current different algorithms for target prediction consider stringent seed matching, thus having a high degree of overlap between them. One common point appears to be the conservation of the target sites during evolution as many of the potential miRNA binding sites are in evolutionarily conserved UTR regions [171, 193]. In fact, more than 60% of human protein-coding genes have been under selective pressure to maintain pairing to miRNAs in their 3'UTRs, although mammalian-specific miRNAs have far fewer conserved targets than do the more broadly conserved miRNAs [173, 191]. Nonetheless, the computational predictions are not 100% identical, not only because they use slightly different UTR databases but also because they attribute different importance and scores to the characteristics of site-matching discussed above [194]. It is important to note that the 3'UTR length of highly expressed genes can be tissue-dependent. Moreover, when genes have alternative 3'UTR isoforms most predicted target sites fall into the alternative part, indicating that 3'UTR isoform regulation is very important to define miRNA targeting [190].

It is not yet fully understood how microRNAs 'choose' between the two alternative mechanisms by which they mediate silencing of their targets: either translational inhibition or mRNA stability decay [196]. It may be species-, tissue- [197] or microRNA specific, target-dependent or reliant on the RNA binding proteins involved and it may be under the control of signaling pathways [198].

Although it may occur less frequently in mammals, there is evidence that regulation at the level of mRNA stabilization (via mRNA degradation or deadenylation) may serve as a common mechanism for miRNA function [171, 190, 199]. In this regard, animal miRNAs have been found to mediate mRNA degradation even when the target sites have incomplete complementarity to them [200, 201]. Also, microarray experiments show that the over-expression of miRNAs in cells causes mostly mild down-regulation of several transcripts (less than two fold) [171]. As an example, over-expression in HeLa cells of miR-1, preferentially expressed in heart, or miR-124, preferentially expressed in the brain, causes down-regulation of about 100 mRNAs, shifting the expression profile towards muscle-like or brain–like, respectively [171]. Moreover, it was noticed that sequence motifs over-represented in these down-regulated mRNA are enriched for MRE (MicroRNA Recognition Elements) of the respective microRNA, matching to the seed sequence in the miRNA. This demonstrates that the knockdown of those transcripts is caused mainly by direct binding of the transfected miRNAs to the 3'UTRs. Other experiments [171], using site directed mutagenesis have shown that pairing to the miRNA seed region contributes directly to mRNA decrease. In conclusion, miRNAs have a mostly mild effect (less than two fold) of down-regulation of hundreds of transcripts. Interestingly, deadenylation mediated by miRNA action can promote not only mRNA decay but also translational repression, by attenuation of the stimulatory role of the poly(A) tail in mRNA translation [202].

The regulation mediated by translational repression is a very attractive strategy since it allows for potential translational reactivation of the repressed mRNAs, instead of irreversibly removing them through mRNA turnover. When addressing this issue, it is also important to take into account the mRNA turnover rate. If a mRNA has fast turnover rate it may appear to be solely translationally repressed while long-lived mRNAs may be more

susceptible to an increase in decay rates by miRNA repression [203]. The translational repression involves recruitment of the ribosome anti-association factor eIF6 or binding of Ago2 to the mRNA m(7)G cap. This prevents recruitment of eIF4E and translation initiation [198].

Similar to what happens at the mRNA level, recent studies showed that the repressive effect of miRNA over-expression on individual proteins is overall very mild [193], rarely exceeding fourfold [192], while having a broad impact in tuning the synthesis levels of a large fraction of the proteome [192, 193]. Most targets are in fact repressed at both the mRNA and the translational level and the amount of each process contribution to the down-regulation seems to depend on the individual miRNA–mRNA pair and its localization [192], since translational repression appears stronger for mRNAs translated at endoplasmic-reticulum-associated ribosomes compared to free cytosolic ones. For those targets with robust protein down-regulation (>one third) mRNA destabilization usually was responsible for the major component of repression [193]. Translational repression is responsible for a substantial amount of miRNA-mediated repression, but targets that are repressed only at the level of translation are down-regulated quite modestly (<33%) [193]. In conclusion, although some detected proteins are repressed by 50%–80%, the general effects are more modest, even if the targets are conserved, with individual sites usually reducing protein output by one third to half.

As in the case of mRNA decay effect, evolutionarily conserved target sites cause stronger effects in translational repression than non-conserved sites [192, 193] and the use of the conservation criteria enriches for sites with possible functional roles and also for those that are more effective. Nevertheless, non-conserved targeting is more widespread than conserved targeting and predictive tools have incorporated the possibility of not using conservation cutoffs (like PITA, TargetScan and Miranda). Though ignoring conservation diminish overall performance of prediction, the highest ranked targets perform as well as the respective highest ranked conserved predictions [193]. The majority of the down-regulated genes found could be explained by seed matching sites but there have been found repressed proteins without seeds represented in their genes and these are most probably direct targets

of the respective miRNAs. Again, the presence of an A in the first miRNA nucleotide favors miRNA-mediated protein down-regulation, which can explain the preferential conservation of an A at this position, even when there is a mismatch to the mRNA [173]. Besides that, again the most influential component was local AU composition.

Recent evidences indicate that miRNAs can also undergo post-transcriptional modifications that impact on their regulation and functional effects. Just to mention a few examples, a process called RNA editing, involves nuclear enzymes called ADAR (adenosine deaminase acting on RNA) and occurs at the level of the pri-miRNA, altering miRNA processing or specificity. ADAR enzymes are nuclear enzymes that catalyze adenosine-to-inosine (A-to-I) transitions in dsRNA substrates, and given that inosine pairs preferentially with cytidine, this base editing can alter base-pair specificity [204]. When specific adenosine-to-inosine editing occurs in the seed region of a miRNA it provokes a change in their target mRNAs. Hence, a different set of genes can be reassigned to an edited miRNA isoform in, for instance, a tissue-specific way [205]. RNA editing can also have a function in miRNA biogenesis, altering the pri-miR processing. For instance, editing of pri-miR-142 can inhibit Drosha, decreasing the mature miRNA-142 expression [206].

MicroRNAs and lymphopoiesis: focus on T-cell lymphopoiesis

MiRNAs have been implicated in practically all biological processes, being crucial for the development of multiple organisms. Not surprisingly, knockouts of proteins involved in miRNA biogenesis, such as Dicer or DGCR8, results in early embryonic lethality or developmental defects [207-209]. This and the fact that some microRNAs are located at sites of genomic alterations linked to human leukemia [210], led to the speculation they may be involved in a tightly orchestrated process of regulation of mammalian hematopoiesis. Several studies confirmed that miRNAs have important roles as ‘fine tuners’ of normal hematopoiesis [211]. Interesting five miRNAs were shown to be highly specific for hematopoietic cells as compared to other tissues: miR-142, miR-144, miR-150, miR-155, and miR-223 [212].

Moreover, studies in mice have shown that some miRNAs are preferentially expressed in hematopoietic tissues and others are differentially expressed between hematopoietic lineages [213]. For example, miR-181a was found to be strongly expressed in the thymus, besides brain and lungs; miR-223 was almost exclusively expressed in the murine bone-marrow; and miR-142 preferentially expressed in murine hematopoietic tissues (bone marrow, spleen and thymus). Within each hematopoietic organ, each one of these miRNAs had a lineage specific expression pointing to a possible function in hematopoietic lineage differentiation. In fact, ectopic expression of these miRNAs in murine bone marrow hematopoietic progenitor cells substantially altered lineage differentiation: expression of miR-181 lead to an increase of the B-lineage fraction both *in vitro* and *in vivo*, while expression of miR-142s or miR-223 lead to a 30 to 40% increase in the T-lymphoid lineage *in vitro*, with little or no effect in the B-lymphoid lineage [213]. Latter studies in human hematopoietic lineages showed that hematopoietic expressed-miRNAs may be present and act differently in human versus mouse hematopoiesis given that the same types of human and mouse hematopoietic cells show large differences in miRNA expression [214]. In humans, miRNAs are differentially expressed during normal lymphoid commitment of the multipotent progenitor (MPP) cell. For instance, miR-128a, miR-181a and miR-146, were predicted to inhibit the differentiation of MPPs into common lymphoid progenitors. On the other arm of the hematopoietic differentiation, miR-155 was shown to negatively regulate human myelopoiesis and erythropoiesis [214]. Moreover, Li et al. [215] demonstrated that miR-181a regulates T-lymphocyte receptor sensitivity during development and that increasing miR-181a expression in mature T-lymphocytes augments the TCR sensitivity to peptide antigens, whereas the inhibition of miR-181a in immature T-lymphocytes reduces the TCR sensitivity and impairs both positive and negative selection [215]. Thus, miRNAs exert a pivotal role not only in lymphocyte differentiation but also in the immune response.

A broader view of how relevant microRNAs are in regulating the T-lymphocyte development has been addressed using Dicer knockout mice in which conditional deletion of Dicer alleles was used to avoid embryonic lethality [216-218]. Dicer deletion in early T-cell development resulted in a sharp reduction of miRNAs at the DP stage. Thymus of knockout

mice had 10 times less TCRαβ thymocytes than control counterparts, but maintained the normal DN cell numbers, explaining increased DN frequency in these thymi. The reduction in cellularity could be explained by increased apoptosis of αβ lineage cells, whereas the numbers of γδ-expressing thymocytes were not affected in mice lacking Dicer. CD4/8 lineage commitment was unaffected with maintenance of the normal mature CD4/CD8 lineage proportions, up-regulation of lineage-specific genes and the stable silencing of *Tdt* [218]. In a later T cell development stage, deletion of Dicer in mice using a Cre transgene under the control of the *Cd4* enhancer/promoter/silencer resulted in normal total numbers and percentages of thymocytes [217]. Nonetheless, in the spleen, lymph nodes and blood there was a moderate reduction of CD4 T cells and marked reduction of CD8 T cells due to decreased proliferation and increased apoptosis of Dicer-deficient T cells. In these conditions, helper T cells preferentially differentiated into Th1-like, IFN-γ-expressing cells even when cultured under strong Th2-polarizing cues. This shows that Dicer is essential for the maturation and/or maintenance of peripheral T lymphocytes, especially cytotoxic CD8 T cells, and also for the cytokine production during helper T cell differentiation [217]. Moreover, differentiation of natural regulatory T (Treg) cells in the thymus and Foxp3 expression were also compromised in the absence of Dicer and mature miRNAs [216].

Interestingly, general microRNA expression up-regulation was observed in the maturation of human thymocytes from the DP to the SP stage [219]. Among the miRNAs that are highly modulated, some had previously been reported to play an important role in the differentiation of the lymphoid (miR-150, miR-155, miR-23a) or myeloid (miR-146, miR-223) hematopoietic lineages and in the regulation of cellular proliferation or apoptosis (miR-150, miR-23a, miR-27a, miR-24). Some miRNAs decrease their expression in the transition from DP to SP stage, but the level of up-regulation of miRs generally exceeded that of down-regulation. This result is consistent with previous studies that correlated miRNA levels to cell differentiation and suggests that miRNA up-regulation in this context could be important for down-regulating genes associated with an immature phenotype [219].

MicroRNAs and cancer

MicroRNAs have been implicated in practically all cellular pathways [179]. So it is not surprising that their deregulation has been associated with tumorigenesis. The complex interaction between microRNAs and protein-coding genes is found deregulated in many diseases, including human cancer.

The initial evidence for the involvement of miRNAs in cancer came from the molecular characterization of the 30kb deletion 13q14 in human chronic lymphocytic leukemia (CLL). This genomic region harbors the *miR15* and *miR16* genes that were found to be deleted or down-regulated in the majority of CLL patients [210]. Both microRNAs from this cluster negatively regulate BCL2, inducing apoptosis in leukemic cells [220]. Further evidence came from the realization that miR-196 and miR-10a are located in *HOX* clusters that encode for the HOX transcription factors, which have a crucial role in development and oncogenesis [221, 222]. Many miRNA genes are located at fragile sites, regions of minimal loss of heterozygosity or amplification, or common breakpoints in human cancers, highlighting the important role that miRNAs may have in the pathogenesis of human cancer [196, 221]. The genomic abnormalities found so far that influence the activity of miRNAs are the same as those described for protein-coding genes, such as chromosomal rearrangements and genomic deletions. On the contrary, given the small size of microRNA genes, the accumulation of point mutations in these are rare events. Therefore, homozygous deletions or gene amplification, promoter methylation or relocalization of a miRNA close to a regulatory element seem to be the main mechanisms of inactivation or activation of microRNAs [221].

Several miRNA profiling studies in multiple human cancers revealed that miRNA expression signatures can be surprisingly informative, reflecting the developmental lineage and differentiation state of the tumors, and can contain diagnostic information [223, 224]. Contrary to mRNA profiles, miRNA expression profiles can be successfully used to classify even poorly differentiated tumors [223], and a small number of microRNAs is sufficient for the classification of human cancers [223, 224]. Almost all miRNAs have a differential

expression across several cancer types, reflecting the developmental origin of the tissues and the cell lineage. Interestingly, irrespectively of tumor type, the majority (59.4%) of the miRNAs have been found down-regulated in tumors when compared with normal tissues, implicating that down-regulation of some miRNAs might play a causal role in the generation and/or maintenance of tumors [223]. Besides that, cancer cell lines tend to express mRNAs with shorter 3'UTR than non-transformed cell lines, due to alternative cleavage and polyadenylation (APA). This leads to loss of miRNA target sites and it explains, at least partially, the increased mRNA stability and translation of mRNA shorter isoforms [225]. Nevertheless, both increased and decreased miRNA abundance has been observed in cancer cells in comparison with normal tissue. In fact, miRNAs can act as tumor suppressors [171, 210] by negatively regulating proto-oncogenes. Conversely, by inhibiting tumor suppressors, miRNAs can function as oncogenes [203, 221, 226].

Oncomirs

Many miRNAs have been found over-expressed in different tumors, functioning as oncogenes, and for that reason called oncomirs, but only some have been well characterized. Oncomirs generally promote tumor development by negatively inhibiting tumor suppressor genes and/or genes that control cell differentiation or apoptosis. The observation that over-expression of miR-155 alone is sufficient to cause lymphoblastic leukemia or high-grade lymphoma in transgenic mice [227] fully demonstrated that microRNAs can act as an oncogenic driving force. Moreover, deregulation of miRNA with oncogenic functions can be part of the tumorigenic program of oncogenes, as in the case of the transcription factor Myc, frequently deregulated in human malignancies [228]. Myc directly activates the transcription of the miR-17–92 cluster that in turn negatively regulates the expression of E2F1**.** Thus, c-Myc tightly controls proliferative signals by activating its transcriptional target E2F1 and simultaneously limiting E2F1 translation through miRNA regulation [229].

Oncomirs can also participate in post-transcriptional regulation of tumor suppressors, such as the bona-fide tumor suppressor PTEN [230]. As an example, miR-21, one of the most frequently deregulated oncogenic microRNAs, directly targets PTEN in hepatocellular

carcinoma, interfering with cell proliferation and apoptosis, and promoting cell migration and invasion [231]. MiR-21 was shown to directly modulate PTEN expression in other cancers, such as epithelial ovarian cancer [232], non-small cell lung cancer [233], and squamous cell carcinoma [234]. Down-regulation of PTEN is frequently involved in the development of cancers, and it can be regulated direct or indirectly by other miRNAs in different cellular contexts: miR-214 in human ovarian cancer [235]; miR-26a in a murine glioma [236] and lung cancer [237]; miR-221 and miR-222 in non-small cell lung cancer and hepatocarcinoma [238] and gastric cancer [239]; miR-29b [240] and miR-301 [241] in breast cancer; and miR-153 in prostate cancer [242]. Furthermore, different oncomiRs can act on the same oncogenic pathway. For example, all three components of the miR-106b-25 cluster cooperate in decreasing PTEN protein abundance, contributing to prostate tumorigenesis [184].

Tumor suppressor miRs

The fact that a general down-regulation of microRNAs [223] is observed in human cancers and the fact that there is a tendency for shortening of UTRs in malignant cells [225], suggest that microRNAs may also have an intrinsic tumor suppressive role. Tumor suppressor miRNAs generally prevent tumor development by negatively inhibiting oncogenes and/or genes that control cell differentiation or apoptosis. In fact, miRNA control can be part of the tumor suppressive programs of bona fide tumor suppressor genes, as exemplified by p53 and miR-34 [221, 243]. p53 directly regulates the transcription of miR-34a–c cluster increasing cell cycle arrest and cellular senescence, effects partially attributed to direct repression of CDK4 and CCNE2 by miR-34 [243].

Repression of tumor suppressor miRNAs can also be part of the oncogenic program of oncogenic transcription factors. In fact, it was shown that Myc induction leads to an extensive repression of microRNAs in the context of human tumorigenesis [228]. Myc binds at upstream conserved regulatory regions of several miRNAs down-regulated in human cells. For instance, it was shown that miR-34a, miR-150, miR- 195/miR-497 and miR-15a/miR-16-1 have tumor-suppressing properties in the context of Myc-mediated transformation [228].

One miRNA may target several components of a single pathway but also several miRNAs can have an additive impact on several genes in the same functional pathway. This kind of 'social' networking or combinatorial effort of miRNAs to control the same critical pathway [244] has been demonstrated to be important in cancer. In fact, several miRNAs concomitantly regulate the EGFR pathway [197], a very important signaling pathway in cancer. In breast cancer, three miRNAs (miR-124, -147 and -193a-3p) were shown to inhibit G1/S transition and cell proliferation by targeting effectors of EGFR signaling (e.g., AKT2, STAT3, p38 and JNK1) and/or cell-cycle proteins (e.g. Cyclin D1) simultaneously, pinpointing these miRNAs as potential tumor suppressors [197].

Noteworthy, unraveling the involvement of miRNAs in cancer can be complicated by the fact that miRNAs may function as oncogenes in one cell type and as tumor suppressors in another. For example, miR-221 and miR-222 target the tumor suppressor PTEN in lung and liver cancers [238], but the same miRNAs inhibit erythroleukemic cell growth via reduction of *KIT* oncogene expression [245].

Competing endogenous RNAs

Recently, another layer of complexity involving microRNAs and mRNA interaction was exposed, with importance in malignancy [246]. By quenching microRNAs through their MRE, a transcript can have a non-coding regulatory function. In fact, it was shown that the *PTEN* pseudogene1 (*PTENP1*) transcript is biologically active and acts as a 'sponge' for microRNAs that target *PTEN*, through a region homologous to *PTEN* 3'UTR enriched for known miRNA target sites. By regulating PTEN cellular levels *PTENP1* exerts a growth-suppressive role [247]. This phenomenon is transversal to other RNA transcripts. The *PTEN* transcript itself can function as a decoy for microRNAs, regulating their availability for transcription repression of other targets [247]. A similar role can be attributed to *CNOT6L* and *VAPA* transcripts, which impair miRNA-mediated regulation of PTEN by acting as miRNA decoys [248]. This notion disrupts the canonical idea of a protein-coding gene transcript working solely as an information molecule for the protein coding machinery. Thus, RNAs that actively regulate other transcripts through direct competition for microRNA binding are been named

competing endogenous RNAs (ceRNAs). In fact, this mechanism reveals a communication path between several types of transcripts, such as pseudogenes, lncRNAs and mRNAs, that is mediated by microRNA response elements [246].

MicroRNAs and Leukemia

The differential expression of miRNAs in hematopoiesis suggested early on that deregulated miRNAs could have a role in leukemogenesis [213]. Initially, reports of miRNA expression profiles of several types of solid cancers and leukemia (including pediatric ALL patients) showed that those profiles can be used to distinguish cancer types [223]. Moreover, each profile can be interpreted in light of the microRNA variations during normal hematologic ontogenesis [249]. Currently, the involvement of miRNAs in leukemogenesis has been established and several miRNAs have been identified as oncogenes or tumor suppressors in human leukemia.

Several studies have aimed to define a signature that differentiates leukemia cells from their normal counterparts. In what regards acute lymphoblastic leukemia, comparison of several pediatric ALL samples with normal CD34+ hematopoietic cells has shown that leukemia cells can be distinguished by the up- or down-regulation of several microRNAs (Table 1.3) [250]. Part of this leukemic microRNA signature was confirmed in another study by comparison of ALL cells to normal bone marrow cells (Table 1.3) [251]. In the latter study, miR-196b expression was found to be higher in T-ALL patients [251]. Another work, set out to reveal new microRNAs by small RNAseq, acknowledged 16 novel, 170 candidate to novel and 153 known mature miRNAs/miRNA-star strands expressed only in childhood ALL [252].

MiRNA expression signatures can be used not only to classify human cancers but also to predict prognosis. The relation of miRNAs to risk categories in childhood ALL has been investigated (Table 1.4). For instance, patients who develop CNS relapse at one year follow-up showed significant differential miRNA expression as compared to non-CNS relapsed ALL cases [253]. An analysis of risk not restricted to CNS involvement determined a microRNA signature distinct between relapsed and complete remission cases [253]. Furthermore, some

studies suggest that miRNAs can be used to predict the risk of relapse before patients undergo therapy. The analysis of miRNA expression profiles and relapse-associated miRNA patterns in a panel of matched diagnosis–relapse or diagnosis–complete remission (CR) childhood ALL samples, has shown that three miRNAs were enough to predict relapse-free survival (RFS) in a representative cohort of ALL patients with a three-year follow-up (Table 1.4) [254]. In another study, an integrated analysis of 14 miRNAs was highly predictive of clinical outcome and able to differentiate a group of patients with a favorable expression profile and a 5-year DFS of around 90% from those with a less favorable miRNA profile and a 5-year DFS rate of around 61% [255]. The signature associated with an unfavorable prognosis included the miR-33 which is significantly up-regulated in T-ALL in comparison to normal thymocytes [255].

Table 1.3 **– Comparison of miRNAs expression between childhood ALL and normal samples**

	Most discriminative miRNAs	**Normal Samples**
High expression in ALL	miR-128a, miR-142, miR-150, miR-181, miR-30e-5p, miR-193, miR-34b, miR-365, miR-582, miR-708	CD34+ progenitors form Peripheral Blood [250]
	miR-128 and miR-181	Bone marrow cells [251]
Low expression in ALL	miR-100, miR-125b, miR-99a, miR-196b, miR-let-7e	CD34+ progenitors form Peripheral Blood [250]
	miR-100, miR-196b, let-7e	Bone marrow cells [251]
High expression in T-ALL	miR-5194, -5193, -5192, -5191, -5188, -3151 miR-5197*, -5196*, -3942*, -3136*	Thymocytes [252]
Low expression in TALL	miR-3183, miR-3190, sol-miR-16	Thymocytes [252]

Note: candidate novel miRNAs from Solexa sequencing have the prefix sol [252]

In respect to miRNAs associated with resistance to commonly used drugs in childhood ALL treatment, a unique miRNA expression signature composed of eight (out of 576) miRNA genes (miR-18a, -532, -218, -625, -193a, -638, -550, and miR-633) is able to differentiate

between a good or poor prednisone response in pediatric ALL cases [253]. Resistance to vincristine, daunorubicin and l-asparaginase in precursor B-ALL patients is also distinguished by differential expression of miRNAs (Table 1.4) [255]. The fact that specific drug-resistant cases have unique miRNA expression profiles helps to comprehend the biology underlying drug resistance.

Table 1.4 **– miRNA expression and risk categories, prognosis and treatment response in ALL**

		microRNAs
up-regulation	Patients with CNS relapse versus non-CNS relapsed cases [253]	miR-7, miR-198, miR-633
down-regulation	Patients with CNS relapse versus non-CNS relapsed cases [253]	miR-126, miR-345, miR-222, miR-551a
up-regulation	Relapsed patients versus complete remission cases [253, 254]	miR-7, miR-216, miR-100 [253], miR-708 [254]
down-regulation	Relapsed patients versus complete remission cases [253, 254]	miR-486, miR-191, miR-150, miR-487, miR-342 [253], miR-223, miR-27a [254]
High expression	Associated with an unfavorable prognosis [98]	miR-33, miR-215, miR-369-5p,miR -496, miR-518d, miR-599
High expression	Associated with a favorable prognosis [98]	miR-10a, miR-134, miR-214, miR-484, miR -572, miR-580, miR-624, miR-627
up-regulation	Associated with resistance to vincristine and daunorubicin [255]	miR-125b, miR-99a, miR-100
down-regulation	Associated with resistance to l-asparaginase [255]	miR-454

Furthermore, microRNA expression profiles have also been shown to discriminate leukemia cells with origin in different hematopoietic lineages. In fact, a specific miRNA signature is able to discriminate ALL from acute myeloid leukemia (AML) with high accuracy [256]. Moreover, the miRNA-expression profiling of ALL patients allowed the discrimination of different ALL subtypes and the clustering of the samples into three segregated hierarchical branches: B-ALL, T-ALL and mixed lineage leukemia samples (MLL). This was the first

evidence that miRNA signatures may reflect the developmental history of human cancers [223]. Furthermore, the characterization of microRNA expression is able to differentiate the major subtypes of ALL, such as T-ALL (Table 1.5), *MLL*-rearranged, *TEL-AML1*-positive, *E2A-PBX1*-positive, and hyperdiploid cells, with the exception of BCR-ABL-positive and 'B-other' ALL [255].

In fact, current data show that miRNA expression profiles are more likely representative of ALL subtypes than of the differentiation stage of each subtype [250]. For instance, T-ALL and MLL-rearranged cells have common miRNA expression that is distinct from other ALL subtypes (TEL-AML1, BCR-ABL, E2A-PBX1, hyperdiploid, and B-other) [250]. Furthermore, in agreement with previous findings showing that T-ALL cases are clearly distinct from B-ALL at the gene expression level, the expression of miR-148a, miR-151, and miR-424 have been found to discriminate T- cell from B-cell lineage ALL [257].

Table 1.5 **- The miRNAs most discriminative of T-ALL as opposed to other ALL subtypes**

	Most discriminative miRNAs	Ref
up-regulation	miR-132, miR-151, miR-191, miR-222*, miR-425-5p, miR-425-3p (*), miR-708 (*#), miR-148a, miR-424 ALL (&)	(*) [250] (#) [255] (&) [257]
down-regulation	miR-196b (*),miR-190, miR-342-3p, miR-542-5p (#), miR-151(&)	

Overall, these studies suggest the potential role of specific microRNAs in specific pediatric ALL subtypes and in the development of different phenotypes, namely those associated with drug resistance and risk of relapse. The functional importance of each specific miRNA by itself or as part of a network of microRNAs needs to be further addressed to determine specific roles in the biology of acute leukemia.

Regarding hematopoietic malignancies, it is important to mention the polycistron encoding the miR-17-92 microRNA cluster – which in humans is located at the chromosome 13q31, a genomic region that is recurrently amplified in lymphomas and other cancers. The

human miR-17−92 transcript can be processed into seven mature miRNAs (miR-17−5p, -17−3p, -18a, -19a, -20a, -19b, and miR-92). The miRNAs encoded from this cluster are highly expressed in murine lymphocytes, embryonic stem cells and precursors. These miRNAs expression levels decrease upon maturation during lymphocyte development [258]. The genomic region encoding the miR-17-92 cluster is often amplified in human B-cell lymphomas and cooperates with up-regulated c-Myc expression in fetal liver cells to accelerate the formation of B cell lymphomas in mice [226]. Furthermore, increased expression of miR-17−92 in mouse lymphocytes results in the development of a lymphoproliferative disease and autoimmunity, causing the premature death of the mice [258]. These phenotypes are due to increased proliferation and reduced activation-induced cell death of T- and B-cells upon activation in the periphery, rather than a global inflammatory response. miR-17−5p and miR-19 were shown to cooperatively suppress Pten protein expression and miR-92 to down-modulate the pro-apoptotic Bim protein, partially explaining the transgenic phenotype [258]. Moreover, miR-17-5p and miR-20a target the transcription factor E2F1, important for cell cycle progression, and its reduction promotes the development of hematopoietic malignancy [229]. Increased miR-17−92 expression might also cooperate with pre-existing oncogenic activation, such as c-Myc over-expression [228]. The miR-17-92 cluster was also shown to promote the activation of the PI3K pathway by directly mediating the inhibition of PHLPP2, a negative regulator of the PI3K pathway [259], thus strengthening the importance of this oncogenic polycistron in lymphomagenesis.

MicroRNAs and T-ALL

The participation of miRNA genes, individually or as part of a network, has been implicated in T-ALL pathogenesis. Mouse transplantation with Lin- hematopoietic fetal liver cells over-expressing mature miR-125b was shown to cause the malignant transformation of different hematopoietic lineages, leading to B-cell ALL, T-cell ALL, or a myeloproliferative neoplasm. These results suggest a role for miR-125b in the differentiation of lymphoid and myeloid lineages and identified a microRNA involved in the genesis of T-ALL [260].

It is important to notice that, contrary to B-ALL subtypes [257], hierarchical clustering and principal component analysis of the expression levels of 430 miRNAs in 50 clinical T-ALL specimens does not distinguish between the major cytogenetic groups (HOXA, TAL or LMO and TLX1 or TLX3), which differ by few miRNAs [261]. Nevertheless, in the high-risk subgroup of ETP-ALL, the microRNAs miR-221 and miR-222 were found significantly up-regulated when compared to non-ETP-ALL [262]. Moreover, it has been proposed that miR-222 may, to some extent, contribute to the myeloid character of ETP-ALL by down modulating *ETS1* expression. Through ETS1 down-regulation, miR-222 significantly inhibits proliferation and causes cell cycle arrest and apoptosis in leukemic cells [262]. In addition, miR-221 associates with poor prognosis: increased expression correlates significantly with lower 5-year OS rates [263].

The already mentioned oncogenic miR-17-92 microRNA cluster has also been implicated in T-ALL. MiR-19, the cluster component with higher expression in T-ALL, enhances lymphocyte survival and cooperates to promote leukemogenesis in a mouse model of Notch1-induced T-ALL [264]. Moreover, the miR-17-92 cluster is involved in a genomic rearrangement in T-ALL, the translocation t(13;14)(q32;q11) with the TCRA/D locus [264]. In the murine model of T-ALL, miR-19 targets the pro-apoptotic protein *Bim*, AMP-activated kinase (*Prkaa1*), and the tumor suppressors *Pten* and *PP2A*, resulting in the overall activation of PI3K signaling. In this way, miR-19 directs a coordinated action to control the PI3K signaling to affect lymphocyte survival and leukemogenesis [264]. The mechanism of pri-miR-17-92 activation in T-ALL remains to be fully understood, but it was proposed that NK-like homeodomain proteins could stimulate the expression of this polycistron in T-ALL [265].

Furthermore, it has been shown that a small set of miRNAs (miR-19b, -20a/93, -26a, -92 and miR-223) is responsible for the cooperative suppression of several tumor suppressor genes in T-ALL, namely *PTEN, BIM, NF1, FBXW7, IKZF1* and *PHF6* [261]. Combining the analysis of miRNA expression data of T-ALL primary cells and cell lines the authors defined the most highly expressed miRNAs (miR-223, -19b, -20a, -92, -142-3p, -150, -93, -26a, -16 and miR-342). These were further tested in a mouse model of Notch1-induced T-ALL. The assessment of individual and combined contribution of each miRNA functional effects revealed that highly expressed miRNAs cooperate in regulating key tumor suppressor genes

in human T-ALL cells. In fact, mir-223, reported as a 'myeloid' gene, was notable for its differential up-regulation in T-ALL. It was shown to promote Notch1-driven leukemia at least in part by controlling the E3 ligase FBXW7 [261]. A follow up study identified miR-128-3p as a novel candidate oncomiR in T-ALL by targeting PHF6 tumor suppressor gene. Over-expression of miR-128-3p accelerated leukemia onset in a Notch1-induced T-ALL mouse model [266].

The use of microRNAs to inhibit oncogenic signals is an attractive alternative to the targeting of oncogenes themselves, which often have essential functions and therefore are more difficult to inhibit without substantial harmful effects in normal tissues. One of such instances is the case of Notch-induced oncogenesis in T-ALL, whose pharmacological inhibition has been associated with gut toxicity [267]. Importantly, it has been shown that *mir-181ab1* gene deletion significantly delays T-ALL development induced by Notch oncogenic signals, without significant impact on normal development [268]. In T-cell lineage, *mir-181ab1* deletion influenced early thymocyte development, resulting in a modest decrease in ETP, DN3 and DP cell populations and in an increase in CD4 SP thymocytes. Mir181ab1 role in the oncogenic program appears to be more detrimental, as deletion of *mir-181ab1* reverts ICN1-controlled gene set (*Dtx1*, *Notch1*, *Hes1*, *Hey* and *Nrarp*) and pre-TCR (*Ptcra*) back to the levels of normal DP cells. Also, miR-181a, a component of the *mir-181ab1* gene cluster, regulates T-cell receptor sensitivity during development by suppressing the expression of multiple phosphatases (*Dusp5*, *Dusp6*, *Shp* and *Ptpn22*) [215]. Thus, miR-181a may contribute to the maintenance of oncogenic signals by diminishing negative feedbacks and potentiating NOTCH and pre-TCR signals [268].

Still in the context of Notch1-induced oncogenesis, it was found that repression of miR-451 and miR-709 is a necessary event in murine T-ALL [269]. In mice, Notch-1 indirectly down-regulates miR-451 and miR-709 by inducing the degradation of their transcription activator E2a. Inhibition of these miRNAs is probably necessary since both miRNAs directly repress Myc expression. In addition, miR-709 also directly represses the oncogenes Akt and Ras-GRF1. Human T-ALL cells with *NOTCH1* activating mutations have decreased miR-451 (miR-709 is not conserved in humans) and increased MYC levels, compared to T-ALLs without

NOTCH1 mutations. This is an example of miRs, behaving as tumor suppressors, whose expression must be down-regulated during ICN1-induced T-ALL [269].

Another member of the Notch receptor family, NOTCH3, has also been validated as a target of microRNAs, namely of miR-150 in T-ALL cell lines [219]. Moreover, another study determined a NOTCH3-induced microRNA signature in T-ALL [270]. Overall, seven microRNAs were found coherently modulated by both NOTCH3 over-expression and silencing, in mouse and human cells (miR-223, miR-183, let-7d, miR-425, miR-25, miR-139-5p and miR-103). In particular, miR-223 is a direct target of NOTCH in human T-ALL cells. The binding of NOTCH to mir-223 promoter region requests NF-kB activation. Surprisingly, in primary human samples the authors were unable to show a direct correlation between miR-223 expression levels and the very common T-ALL-related events of up-regulation of NOTCH and NF-kB. This result suggests that miR-223 expression may be maintained by other pathways aberrantly activated in T-ALL [270].

The miR-142-3p stands out as another example of the essential role that an individual miRNA can play in leukemia progression, chemotherapeutic resistance and prognosis assessment. In T-ALL, miR-142-3p was shown to promote leukemic cell growth and to induce resistance to glucocorticoid (GC) treatment. [271]. Originally, miR-142-3p was identified as highly specific microRNA for hematopoietic cells [212, 213]. Besides that, miR-142-3p is also highly expressed in pediatric ALL samples, particularly in T-ALL cells as compared with healthy donor T-cells [250] and especially in T-ALL cells from patients with poor prognosis [250]. In addition, ectopic expression of miR-142-3p results in the increased proliferation of T-ALL cells without affecting apoptosis. This miRNA specifically targets the cAMP/PKA pathway and glucocorticoid receptor alpha (GRa) and, more importantly, T-ALL cells from patients with a poor response to prednisolone experience increased cell death induced by dexamethasone upon down-modulation of miR-142-3p [271].

The role of some miRNAs in leukemia is still controversial since it may depend on the cellular context. For instance, miR-196b is highly expressed in T-ALL when compared to B-cell patients [250]. The miR-196b gene is positioned between HOXA9 and HOXA10. In pediatric ALL miR-196b is highly co-expressed with genes from to the HOXA cluster, namely in T-ALL

cases characterized by the activation of HOXA genes, suggesting co-transcriptional activation [222]. The over-expression of this miRNA in mouse bone marrow cells leads to increased proliferative capacity and survival, pointing to a role in leukemogenesis [272]. In contrast, it was reported that miR-196b can down-regulate the oncogenic transcription factor ERG in adult AML and T-ALL patients. As a result, instead of promoting leukemogenesis, miR-196b may also inhibit this process. Moreover, no significant differences in clinical outcome between high versus low miR-196b expression levels were observed [272]. Additionally, miR-196b was found down-modulated in T-ALL patients with respect to normal cells, which suggests a possible tumor suppressor function for this miRNA in T-ALL [273]. In contrast to what was previously found in B-ALL, c-Myc gene expression is not down-regulated by miR-196b in T-ALL. Interestingly, miR-196b loses its ability to down-regulate *c-Myc* gene expression in T-ALL as a result of mutations in target 3'UTR of the *c-myc* gene, pointing again to a possible tumor suppressive role in this disease. This microRNA may therefore have a dual role in leukemia, depending on the genetic context [273].

Apart from genomic aberrations or aberrant activity of regulatory factors (for example, c-Myc), miRNA deregulation may also be caused by aberrant expression of neighboring protein-coding genes (for example, miR-196b and HOXA genes). However, these mechanisms can only explain a minority of deregulated miRNAs, whereas an explanation is still lacking for the majority of aberrantly expressed miRNAs in T-cell acute lymphoblastic leukemia [274].

Objective

Regardless of improved therapy regimens, acute lymphoblastic leukemia from T-cell origin is still not a curable disease [15]. More intensive regimens are used for T-ALL in most clinical trials [8] and current specific treatment protocols for this subtype of ALL allow a 5-year event-free survival rate of 85% [9]. Nevertheless survivors face long term complications [7] and develop serious health problems within 30 years of their initial diagnosis. This issue is even more critical considering the age range of the patients affected by leukemia.

Therefore, the field currently faces the challenge of creating more efficacious therapies, rationally-designed and less toxic. A better understanding of the pathogenesis of the disease, namely molecular analysis of the common genetic alterations in leukemic cells, may be the solution to understand why some cases fail to respond to chemotherapy and to improve selective targeting of leukemic cells without long-term effects on the normal tissues.

With this in mind, this work was conducted aiming to contribute to a better comprehension of the biology of the disease and cell-intrinsic characteristics, namely in relation to one of the most frequently up-regulated oncogenes in childhood leukemia: TAL1.

Surprisingly, the current understanding of the molecular mechanisms that lead to ectopic TAL1 activation is still relatively scarce, specifically for the majority of the T-ALL cases that lack chromosomal rearrangements in the *TAL1* locus. During the past decade, several studies evaluated the miRNA gene expression signatures associated specifically with T-ALL, revealing that some miRNAs can contribute singularly or in combination to the pathogenesis of the disease. Moreover, TAL1 is a putative target for several miRNAs that are differentially expressed in hematopoietic lineages [275, 276], suggesting that miRNAs might regulate TAL1 at different stages of hematopoietic development. The interplay between TAL1 and miRNAs in T-ALL has not been explored thus far. Moreover, miRNA regulatory networks regulating T-ALL oncogenes have not been so far extensively studied.

Thus, the work presented in the current thesis aimed to explore the hypothesis that TAL1 ectopic expression in T-ALL from currently unknown reasons could, in some cases,

result from, or be amplified by, abnormal down-regulation of particular miRNAs targeting TAL1.

Furthermore, TAL1 is an oncogenic transcription factor able to either activate or repress the expression of downstream targets as part of a transcriptional complex. In the current view, TAL1 is at the edge of a complex transcriptional network that is aberrantly expressed in T-cell progenitors and disrupts normal T-cell homeostasis, contributing to the onset of leukemia. At the start of the project reported herein, the relatively small list of known TAL1 target genes included exclusively protein-encoding genes. We therefore aimed to explore the possibility that miRNA genes are transcriptionally regulated by TAL1 and partake in the development of T-ALL.

By addressing these hypotheses we aimed to contribute significantly to the current knowledge of the mechanisms leading to aberrant expression of TAL1 in T-ALL and to the identification of TAL1-regulated miRNA genes. Ultimately, as miRNA-based therapeutics are in the order of the day towards clinical application, we expect our studies may reveal novel molecular targets for improved therapeutic intervention in T-ALL.

Chapter 2

MATERIAL AND METHODS

Material and Methods

MicroRNA nomenclature and annotation. Official nomenclature and sequence annotation based on the miRBase database [277] has changed during the course of this study. In this manuscript the nomenclature reports to the annotated in miRBase release 13.0 (2009). Here we list the microRNA sequences relevant for this study (Table 2.1) and the new annotation according to the last annotation (miRBase release 21.0).

Table 2.1 **– List of microRNAs relevant for this study and their sequence**

The miRNA that changed their names appear in bold on the right column.

miRNA name (v13.0)	**Sequence**	**New annotation (v21.0)**
hsa-miR-101	UACAGUACUGUGAUAACUGAA	**hsa-miR-101-3p**
hsa-miR-140-5p	CAGUGGUUUUACCCUAUGGUAG	hsa-miR-140-5p
hsa-miR-140-3p	UACCACAGGGUAGAACCACGG	hsa-miR-140-3p
hsa-miR-520-5p	CUACAAAGGGAAGCCCUUUC	hsa-miR-520-5p
hsa-miR-520-3p	AAAGUGCUUCUCUUUGGUGGGU	hsa-miR-520-3p
hsa-miR-448	UUGCAUAUGUAGGAUGUCCCAU	hsa-miR-448
hsa-miR-485-5p	AGAGGCUGGCCGUGAUGAAUUC	hsa-miR-485-5p
hsa-miR-135a	UAUGGCUUUUUAUUCCUAUGUGA	**hsa-miR-135a-5p**
hsa-miR-223	UGUCAGUUUGUCAAAUACCCCA	**hsa-miR-223-3p**
hsa-miR-330-3p	GCAAAGCACACGGCCUGCAGAGA	hsa-miR-330-3p
hsa-miR-146b-5p	UGAGAACUGAAUUCCAUAGGCU	hsa-miR-146b-5p
hsa-miR-545	UCAGCAAACAUUUAUUGUGUGC	**hsa-miR-545-3p**

Dicer knockout mice. The transgenic mice with a conditional Dicer allele were obtained from Merkenschlager lab. In these mice the deletion of Dicer in early T-cell development is assured by a CRE transgene under the control of the *Lck* promoter. These mice were generated by crossing Dicer lox/lox mice with LckCre transgenic mice. In the Dicer knockout (ko) mice there

is a substantial deletion of Dicer at the DN3 thymocyte stage and no undeleted alleles are detectable from DN4 stage onwards [218]. The Dicer lox/lox mice were used as controls. Adult age-matched females were sacrificed and their thymuses were extracted and smashed to obtain a thymocyte single-cell suspension. At least two thymuses from mice of each group (control group and Dicer ko group) were pooled together to obtain a dry pellet for RNA extraction. These dry pellets were obtained twice for each group and RNA extraction of each pool was performed in two independent days.

RNA extraction, RT-PCR and quantitative-PCR. RNA was extracted using TRIZOL (Life Technologies Corporation) according to the manufacturer's instructions. When subsequent expression analysis intended to quantify microRNA expression, an additional step of ice incubation for two hours was added after the addition of isopropanol, to ensure an efficient precipitation of small size RNA species. For the RT-PCR to detect protein-coding genes and primary miRNA transcripts, up to 1µg of total RNA was reverse transcribed using SuperScript II (Invitrogen) and random hexamers, according to the manufacturer's instructions. Expression of each gene was normalized to the expression level of the ribosomal RNA *18S* (in the case of human samples) or *Hprt* (in the case of mouse samples) using the dCt method. To evaluate the fold difference of the gene of interest between samples, the ddCt method was used. Primers used for the qPCR are indicated in Table 2.2. The transcripts were amplified in 10µl volume reactions, using 4µl of cDNA, 5µl Power SYBR Green (Applied Biosystems) and 100pM of each primer, according to the manufacturer's instructions. For the detection of mature human microRNA expression, amounts ranging from 100-1000ng of total RNA were reverse transcribed using miRCURY LNA™ Universal RT kit (Exiqon). Real time PCR was performed with commercially available LNA-based primers (Exiqon) for mature microRNA detection with SYBERGreen (Exiqon). The transcripts were amplified in 10µl volume reactions, using 4µl of cDNA, 5µl Power SYBR Green (Applied Biosystems) and 0.8ul of the primer mix prepared according to manufacturer's instructions. Relative expression of the microRNAs was normalized to *SNORD38B* expression using the dCt method. The expression

levels of miR-485-5p were not accessed due to the lack of LNA primers commercially available at this time.

All the amplifications were performed in a ViiA7 Real-Time PCR System (Life Technologies) with the following program: 2 min at 50°C, 10 min at 95°C, followed by 45 cycles of 15 sec at 95°C, and one min at 60°C. This run protocol was always followed by a melting curve protocol to verify the primers' specificity.

Table 2.2 **– List of primers used in quantitative-PCR**

Gene	Forward primer	Reverse primer
TAL1	AACAATCGAGTGAAGAGGAG	CTTTGGTGTGGGGACCAT
18S	GGAGAGGGAGCCTGAGAAACG	CGCGGCTGCTGGCACCAGACTT
pri-miR-146b	CTGGGAACGGGAGACGATTC	AAGTTGGGAGCCCAAACCAT
Tal1 (mouse)	CACTAGGCAGTGGGTTCTTTG	GGTGTGAGGACCATCAGAAATCT
Hprt (mouse)	AGTCCCAGCGTCGTGATTAG	TTTCCAAATCCTCGGCATAATGA

Computational prediction of TAL1 3'UTR targeting by microRNAs. Several web-based bioinformatics tools (PicTar (4-way) [278], TargetScanS release 4.2 [173], miRBase [277], microRNA.org [279-281], DIANA-microT algorithm V3.0 [282], miRDB [283] and StarBase [284]) were used to perform the identification of putative regulators of TAL1. The putative microRNAs binding type, when defined, were listed according to the TargetScanS [173] or DianaMicroT [282] (for 9mer) as following: 7mer-m8 site comprises the seed plus a match to miRNA nucleotide 8 (miRNA nucleotides 2–8); 7mer-A1 site comprises the seed supplemented by an Adenine at target position 1; 8mer site that comprises a match to miRNA nucleotide 8, the miRNA seed and the A at position 1; and 9mer defines an exact match on the positions 1-10. The UTR position of the putative binding may vary slightly according to the program used, therefore we listed the ones predicted by TargetScanS [173]. The conservation score was listed according to DianaMicroT [282] program prediction, when available: it depicts the number of species in which the binding nucleotides of this target site

are conserved. We also listed the target sites of conserved microRNAs with good mirSVR score according to microRNA.org program [279]. The mirSVR scores are based on a regression method for predicting the likelihood of target mRNA down-regulation from sequence and structure features in microRNA/mRNA predicted target sites [279]. Additionally we also listed the microRNAs that are also predicted to target LMO2. Finally, we used the DIANA-miRPath tool [285] to predict the main biological pathways where the putative targets of the microRNA are involved.

A list of miRNAs that potentially regulate TAL1 transcript was compiled at the beginning of this thesis. In the meantime more microRNA genes were discovered and annotated and the annotation of some of the previously found miRNAs suffered alterations. Furthermore, computational algorithms to predict miRNA targeting have also evolved during the recent years to incorporate new criteria for the prediction and also to modify the ranking given to the pairing characteristics. For these reasons, there might be some differences if one was to do a new compilation of the miRNAs that may regulate TAL1.

Luciferase activity assays. A commercially available reporter plasmid with *TAL1* 3'UTR (GeneCopoeia Inc) immediately downstream of the luciferase open reading frame (pLuc-TAL1-3'UTR) was used. This plasmid was co-transfected together with the candidate miRNA expressing vector into 293T cells. Both firefly and renilla luciferases are coded in the reporter plasmid, avoiding the transfection of an additional plasmid for luminescence normalization. Marcos Malumbres kindly provided a vector library- the miRVec library - for expression of several microRNA species [286, 287]. The miRVec vectors express the stem loop sequence (pre-miR) that is processed in the miRNA mature forms. All the vectors used in this study were sequenced and the pre-miR sequences verified. Briefly, $1.5x10^5$ 293T cells were plated in 12-well plates, two wells per miRNA, in DMEM-10% FBS medium. After 24h, cells were transfected with a mixture of 2µl Lipofectamine 2000 (Life technologies), 200µl of OPTIMEM medium (GIBCO), 100ng of the reporter vector and 500ng of the miRNA expressing vector. After 24h, the medium was collected and the plate was frozen until luciferase expression reading. For plate reading, cells were lysed and 20µl of the lysate was processed according to

the Dual-Luciferase Reporter Assay System (Promega), with the volumes of the reagents LARII and STOP&GLO reduced to 50µl each. The firefly luciferase and renilla luciferase activity was measured in an Infinite M200 plate reader (Tecan). For the firefly activity calculations, the values were normalized to the renilla luminescence reading and the average of the two technical replicates was calculated. The values of at least three independent transfection experiments were normalized to the measurements of the correspondent scramble transfection.

Site directed mutagenesis. To point-mutate the TAL1 3'UTR in the pLuc-TAL1-3'UTR vector, a PCR-based commercial kit was used according to the manufacturer instructions - QuikChange II **XL** site-directed mutagenesis (Agilent). When clones bearing the desired mutations were difficult to obtain, the 3'UTR sequence from this vector was cloned in a small vector pGEM-T (promega) and the mutations were performed with the QuikChange II site-directed mutagenesis (Agilent). After mutagenesis, the mutated 3'UTR was cloned back in the original vector. All mutations were confirmed by sanger-sequencing. When more than one miRNA target site was possible for a given miRNA, the mutations were sequentially performed and named from the most upstream to the most downstream one. The primers used in the site-directed mutagenesis are depicted in the Table 2.3.

Cell lines. The human T-ALL cell lines SUP-T1, CCRF-CEM, PF-382, HPB-ALL, P12-ICHICAWA, TALL-1, RPMI-8402, LOUCY, DND-41 and JURKAT have already been described and extensively studied. These cell lines were maintained in RPMI medium (GIBCO) supplemented with 10% FBS (RPMI-10), unless stated otherwise. The cells were split every 2-3 days. 293T cells were maintained in DMEM medium (GIBCO) supplemented with 10% FBS (DMEM-10) and split every 2 days. Cells were cultured at 37°C with 5% CO2. At the indicated time points, the cells were harvested and processed as indicated for assessment of cell viability, and RNA and protein extraction.

Table 2.3 **– List of primers used in site-directed mutagenesis.**

The area of the miss-matching to the MRE element in the 3'UTR is shadowed in the forward primer.

Mutation	Forward primer	Reverse primer
520-5p mutI	GTGAAGAATCCTTGTTTCGAATGAACCACTGCCCCTTCATTGATTTCCTG	CAGGAAATCAATGAAGGGGCAGTGGTTCATTCGAAACAAGGATTCTTCAC
520-5p mutII	GGGCAACATTGTTCACCTGTTTCGCACTCAGGCTCTCC	GGAGAGCCTGAGTGCGAAACAGGTGAACAATGTTGCCC
520-5p mutIII	GGGCAAGTCTTTAGGTCTGTTTCAGAACTAAAGAAGATCTG	CAGATCTTCTTTAGTTCTTACAAAGACCTAAAGACTTGCCCTTTCCTACC
520-5p mutIV	CAGGTACCTTGACCTGTTTCCAGCCCAGAGGCCAACAC	GTGTTGGCCTCTGGGCTGGAAACAGGTCAAGGTACCTG
520-3p mutI	CTGTGGGCGGGCCCAGAAATCTCCGTCAACGTTGTAC	GTACAACGTTGACGGAGATTTCTGGGCCCGCCCACAG
101 mut	GGCCCAGCACTTTCCGTCAACGTTGGAATTTATGTGATGAATTGCG	CGCAATTCATCACATAAATTCCAACGTTGACGGAAAGTGCTGGGCC
140-5p mutI	CCTTATCCTTCATCTTTTAAAGAAATACCAAATGCAAGTCCTTTTGTAAAGTG	CACTTTACAAAAGGACTTGCATTTGGTATTTCTTTAAAAGATGAAGGATAAGG
140-5p mutII	GAAGAATCCTTTTGTAGAATGACCAAATGCCCCTTCATTGATTTCCTG	CAGGAAATCAATGAAGGGGCATTTGGTCATTCTACAAAAGGATTCTTC
140-5p mutIII	GAGAACAAAGATGACCATACCAAATGAAGGGAATCACATCTTTTAAGAC	GTCTTAAAAGATGTGATTCCCTTCATTTGGTATGGTCATCTTTGTTCTC
140-3p mutI	CAATCCAGATGGTGGGATTTTGGTTTCTTAAGGTGAGGCCTGTC	GACAGGCCTCACCTTAAGAAACCAAAATCCCACCATCTGGATTG
140-3p mutII	GTGACTCTTTAGCAAAAAAAACCCATTTTGGGATGATGTGTATATATATG	CATATATATACACATCATCCCAAAATGGGTTTTTTTTGCTAAAGAGTCAC

Electroporation of miRVec vectors in T-ALL cell lines. In order to over-express miR-520d, 101, 140, 485 and 448, T-ALL cell lines were transiently transfected with the corresponding miRVec vectors and the scramble control (miRVec-SCR). Given that miRVec vectors do not have a reporter gene, we co-transfected the cells with a GFP expressing vector (pMAX, Lonza). Hence, a total of 30µg of DNA (9µg of pMax and 21µg of miRVec) were added to 10^7 T-ALL cell suspension in the appropriate volume of pre-warmed RPMI-10 medium (without antibiotics). The samples were placed in 4 cm–gap cuvettes (Bio-Rad) and electroporation

was performed using the Gene Pulser II (Bio-Rad), with the parameters depicted in Table 2.4. After electroporation, cells were washed and cultured in RPMI-10 medium. After 24h, cells were sorted to obtain the GFP-expressing cells using a FACSAria III cell sorter (BD Biosciences). After 48h, cells were collected for RNA and/or protein extraction.

Table 2.4 **– Transfection conditions of T-ALL cell lines.**

Cell	Cell Nr	Volume	Volts	uF	pMAX	miRVec
SUPT-1	10M	350ul	350	750	9ug	21ug
JURKAT	10M	350ul	250	950	9ug	21ug
PF-382	10M	250ul	350	500	9ug	21ug

Nucleofection of T-ALL cells. Nucleofection of JURKAT, CCRF-CEM and PF-382 cells was performed using the Amaxa Nucleofector II (Lonza) according to the manufacturer's instructions. For TAL1 knockdown, cells (2 x 10^6) were washed in RPMI-10 medium and ressuspended in 100µl of solution V with 2µM of a non-targeting pool of small interfering RNAs (siRNAs) or a pool of siRNAs against TAL1 (Dharmacon). For miR-101 or miR-520d-5p knockdown, 2uM of miRCURY LNA™ microRNA Inhibitors (Exiqon) and non-targeting control were used. JURKAT and CCRF-CEM cells were nucleofected using the X-001 program and PF-382 cells using the O-017 program. After nucleofection, the cells were cultured for 48h in RPMI-10 medium.

Immunoblot. After the indicated time intervals of culture, cell lysates were prepared as described [288]. Equal amounts of protein were analyzed by 10% SDS-PAGE, transferred onto nitrocellulose membranes, and immunoblotted with the following antibodies: Tubulin (Sigma, Clone DM 1A), Actin (Santa-Cruz Biotecnology clone I-19), and TAL1 (Milipore, clone BTL73). After immunoblotting with primary antibodies, immunodetection was performed using HRP-conjugated anti-mouse IgG (Promega), anti-rabbit IgG (Promega) or anti-goat IgG (Santa-Cruz Biotecnology) as indicated by the host origin of the primary antibody and

developed by chemiluminescence (Thermo Scientific). Where indicated, densiometry analysis was performed using Adobe Photoshop CS5 Extended software. Each band was analyzed with a constant frame and normalized to the respective loading control. Densiometry values are expressed in arbitrary units.

Production of VSVG-pseudotyped lentiviruses. Vesicular-Stomatitis-Virus-pseudotyped third-generation lentiviruses were produced by transient three-plasmid co-transfection into 293T cells. Briefly, a total of 4 x 10^6 293T cells were seeded in 10 mL of DMEM-10 in 6 cm-diameter dish. The transfection occurred when the cells reached 70-80% of confluence, generally 24h latter. A total of 18µg of plasmid DNA was used for each transfection: 3µg of the envelope plasmid pMD2.VSVG, 6µg of packaging plasmid psPAX2 and 9µg of transfer vector plasmid where the genes of interest were cloned. The plasmids were added to 600µl of pre-warmed OPTIMEM medium (GIBCO) and then added dropwise to 600µl OPTIMEM plus 24µl of Lipofectamine 2000 (Life technologies). After 20minutes of incubation at room-temperature, the mixture was added dropwise to the cells. The medium was replaced after 14 to 16h with 6ml of medium containing 20mM HEPES buffer (pH=7.9). The conditioned medium (Lentiviral Supernatants) was collected after another 24h and 48h. The medium was filtered through 0.45mm pore-size cellulose acetate filters, flash frozen in liquid nitrogen and stored at -80°C until use.

Transduction of T-ALL cells for TAL1 over-expression. T-ALL cell lines were transduced with VSVG-pseudotyped bicistronic lentivirus driving the concomitant expression of TAL1 and GFP (pHR-SIN-TAL1) or with the control mock virus (pHR-SIN-Empty). The resulting cell lines expressing TAL1 or the Empty vector were sorted for an equivalent GFP expression. Briefly, 2.5 x 10^5 cells were incubated in 500µl RPMI-10 in a 24-well-plate with 500µl of lentiviral supernatant and 8ng/µl of polybrene (Sigma). Cells were spun down for 120min at 32°C at 2300 rpm and then incubated over-night at 37°C. In the following day, cells were washed, ressuspended in fresh media and the reporter expression (GFP) confirmed by flow cytometry.

The pHR-SIN vectors were described previously [289] and were kindly provided by Prof. Maria Toribio. We cloned *TAL1* and *LMO2* in the pSIN-BX-IR/EMW vector (also provided by Prof. Maria Toribio) in the BamHI and XhoI restriction sites upstream of the IRES-Emerald sequence. Next we removed the *TAL1*-IRES-Emerald and *LMO2*-IRES-Emerald fragments from the pSIN-BX-IR/EMW vector and replaced the GFP on the pHR-SIN-CSGW vector.

MicroRNA expression analysis. Gene expression analysis of 372 human miRNA genes plus six reference genes was performed in three independent samples of P12 mock transduced (pHR-SIN-EMPTY) and P12 transduced with a vector driving the expression of TAL1 (pHR-SIN-TAL1). We used a qRT-PCR-based array (microRNA Ready-to-Use PCR, Human Panel I, V2.M, Exiqon). Total RNA was extracted, in three independent occasions, from P12-Empty and P12-TAL1 cells. The RNA quality was assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies), assuring the presence of low molecular weight RNA species and RNA integrity (RIN>9). The results from the screening were analyzed with DataAssist v2.0 software (Applied biosystems). A cutoff cycle threshold value of 35 was assigned and mean expression value normalization [290] used as normalization method. The p-value was calculated using a two-tailed Student's t-test. Fold changes relative to mock transduced cells and p-values were determined by the Comparative Marker Selection suite [291]. Cutoffs for statistical significance were a p-value <0.05 and a fold change > 1.5.

Heat Map Illustration. Heat map illustration of differentially expressed microRNAs upon TAL1 over-expression was generated with the GENE-E software (http://www.broadinstitute.org/cancer/software/GENE-E/). MicroRNAs were hierarchically clustered (rows, miRNAs; columns, experiments). Relative expression levels were normalized across the samples as described. The levels greater than or less than the mean are shown in shades of red or blue, respectively.

Confirmation of ChIP-seq enrichment by qPCR. Publicly available ChIP-seq data (GEO accession number GSE29181) was analyzed with the Integrative Genomics Viewer tool. To

confirm the ChIP-seq results we performed ChIP of TAL1 in JURKAT and CCRF-CEM cells followed by qPCR for the selected genomic regions. ChIP protocol was based on the literature [292] and further optimized for T-ALL cell lines and TAL1 immunoprecipitation in F. Speleman Lab. Briefly, 100 x 10^6 T-ALL cells were spun down and the crosslinking was performed with formaldehyde (freshly prepared 37% HCHO solution) at room temperature for ten minutes. The crosslinking reaction was quenched with addition of Glycine (final concentration 125mM). The cells were then washed in cold PBS and the final pellet was flash frozen in liquid nitrogen and kept at -80°C until needed. The cell pellet was thawed and lysed with freshly prepared lysis buffer (Buffer 1: 50mM HEPES-KOH, pH 7.5; 140mM NaCl; 1mM EDTA; 10% Glycerol; 0.5% NP-40; 0.25% Triton X-100; and protease inhibitor cocktail -Roche. Buffer 2 - 10mM Tris-HCl, pH 8.0; 200mM NaCl; 1mM EDTA; 0.5mM EGTA; and protease inhibitor cocktail). The crosslinked chromatin material was then suspended in 3 ml of shearing buffer (1% SDS; 10mM EDTA and 50mM Tris-HCl, pH 8.0) and sheared by sonication in a S220 Focused-ultrasonicator (Covaris). Each sample was placed in TC16 tubes (16mm X 100mm Covaris) and suffered 6 cycles of 5 minutes of shearing at 4-8°C. Per five minutes cycle, the shearing was divided in rounds of 30 seconds of shearing followed by 30 seconds of rest. The sheared chromatin material was stored at -80°C until immunoprecipitation (IP). Some material was analyzed by agarose gel electrophoresis to verify the desired shearing efficiency (200-500 bp fragments). For IP, the 3ml of the chromatin material was ressuspended in 27ml of RIPA lysis buffer. An aliquot was taken for the input sample. Each sample was then incubated with the proper antibody for 4 hours at 4°C with rotation. We used 10µg of anti-TAL1 antibody (ab75739, Abcam) and 10µg of anti-Fibrillarin antibody (ab5821, Abcam). After antibody incubation, 100µl of protein A-beads (Pierce) in the proportion of beads:RIPA of 1:1 were added to each sample. The samples with the beads were incubated over night at 4°C with rotation. In the following day the samples were centrifuged and the supernatant discarded. The beads pellet was washed five times with RIPA buffer. The immunoprecipitated material was recovered from the beads with 50µl of elution buffer (1% SDS; 10mM EDTA and 50mM Tris-HCl, pH 8.0). From this volume, 10µl were eluted for immunoblotting check-up of the IP. The remaining material was diluted in 200µl of elution buffer and eluted for DNA

recovery by placing the samples at 65°C for 22 minutes. At each two minutes of incubation the samples were briefly vortexed. After the incubation, the sample was centrifuged and the supernatant was recovered. Both samples that suffered IP and the input sample stored previously suffered reverse crosslink by incubation at 65°C overnight on an oven (maximum 15 hours). For DNA isolation, 200µl of the input and the samples were added 200µl of TE buffer (10mM Tris-HCl pH 8 and 1mM EDTA) and 0.2g/ml of RNAse (Roche) and incubated for two hours at 37°C. This was followed by 0.2 µg/ml Proteinase K (Sigma) addition and two hours incubation at 55°C. The DNA material was separated with 400µl of phenol:chloroform:isoamylalcohol mix by recovering of the aqueous phase (+-800µl). The precipitation was accomplished by addition of NaCl (200 mM), glycogen (30 µg) and 800µl of absolute ethanol, followed by 30 minutes incubation at -20°C. The DNA was pellet down by centrifugation at 4°C at 20.000g for ten minutes. It was further washed with 80% Ethanol and ressuspended in 30µl of water. The DNA concentration was measured in a Nanodrop (Thermo Scientific) and the samples frozen at -80°C until PCR analysis. For PCR analysis, the ChIP DNA samples were diluted ten times in water and the input samples were diluted 20 times, and 2µl were used per reaction.

The occupancy by TAL1 of the genomic regions at 9.2kbs and 3.5kbs upstream miR-223 TSS and at 11.2kbs upstream miR-146b TSS was analyzed by ChIP-qPCR in JURKAT and CCRF-CEM cells (Table 2.5). The promoter region of LCP2 [143] was used as positive control for TAL1 binding and a random intergenic region was used as negative control. TAL1 binding was calculated as the fold enrichment relative to a mock ChIP performed against Fibrillarin.

Table 2.5 **– List of primers used in ChIP-qPCR**

Name	Forward primer	Reverse primer
LCP2 ChIP	AAGGCTGCTTTGGATCTTGAAA	CCTCCAGCCTGGCTGCTA
ChIP223peak-3	CCTGTTGAAGACACCAAGGGC	TTCCCCAGTGCTGAGCCAAC
ChIP223peak-9	GCAGTGGCTATTCACAGGTGACC	CACTCCCACTATTCACATCACACCTG
ChIP146b_peak-11	GTTGATGCTGCCCTCTCTGT	TCAGGCTGAAGGAGGTGAGA
Intergenic region Chip	GGCTAATCCTCTATGGGAGTCTGTC	CCAGGTGCTCAAGGTCAACATC

miRNA target prediction and gene set enrichment analysis. Prediction of microRNA putative targets was performed by MirDIP data integration portal [293], with a minimum threshold of 4 different applications. MicroRNAs experimentally validated human targets were obtained from mirTARbase 3.5, miRrecords and TarBase 6.0. Target genes without matching any Entrez gene identifier in NCBI were discarded. Graphical representation and analysis of miRNA and their cognate targets was done with Navigator software [294]. We compiled a list of high confidence TAL1 positively or negatively regulated genes from publicly available data [143, 295]. For cross-examination of congruent TAL1 regulated protein-coding and miRNA genes, we intersected the predicted targets of TAL1 down-regulated microRNAs with the protein-coding gene targets previously demonstrated to be positively regulated by TAL1, and vice versa, and searched for common hits in both lists.

For biological function and pathway analysis we collected T-lymphocyte and T-ALL related gene sets from Ingenuity Pathway Analysis (IPA), and from the literature [143, 295, 296]. Additional gene sets were downloaded from version 3.1 of the Molecular Signature Database (MSigDB) at the Broad Institute (http://www.broad.mit.edu/gsea/msigdb). We used three categories of gene sets from MSigDB: (C2) all curated gene sets, (C5) GO biological processes and molecular functions, and (C6) all oncogenic signatures gene sets. Gene set enrichment analysis was performed using Genomica software (http://genomica.weizmann.ac.il/). p-values were determined by a hypergeometric test, followed by a false discovery rate correction to account for multiple hypotheses (FDR < 0.05).

Transduction of T-ALL cells for miR-146b over-expression or knockdown. The transduction procedure was the same described above for the pHR-SIN lentiviral transduction. 24h after transduction the reporter expression (RFP) was confirmed by flow cytometry. The resulting cell lines expressing the vectors or the corresponding mock control were sorted for an equivalent RFP expression. The vectors for miR-146b over-expression (pLemiR-146) were kindly provided by the Sai Yendamuri lab. The pLemiR lentiviral vector (Open Biosystems), was modified by inserting the pre-microRNA-146b sequence in the 3' untranslated region of the gene encoding for the TurboRFP red fluorescent protein and driven by the constitutively

active cytomegalovirus (CMV) promoter [297]. The control vector corresponds to the pLemiR vector without an insert. Both mature forms of miR-146b are expressed (miR-146b-5p and miR146b-3p) [297]. The lentiviral vector for miR-146b-5p inhibition (named by us as Sponge-146b), the pEZX-AM03 vector (Tebu-bio), expresses the specific miRNA inhibitor against hsa-miR-146b-5p (the sequence is proprietary) under the control of the H1 promoter. An independent CMV promoter drives the expression of the reporter gene (mCherry). The control vector (named Sponge-SCR) expresses a scramble miRNA inhibitor control sequence.

Assessment of proliferation. T-ALL cell lines were plated (5 x 10^5 cells/mL) in triplicates in flat-bottom 96-well plates at 37°C with 5% CO2 on day zero. Proliferation was measured either by thymidine incorporation or by cell counts. In the case of the first, cells were incubated with ^{3}H-thymidine (1μCi/well) for 8h prior to harvest. Harvesting of the cells was performed at the indicated time points. Proliferation was determined by analysis of DNA synthesis, which was assessed by ^{3}H-thymidine incorporation using a β-scintillation counter. Proliferation was also assessed by counting cells in a hemocytometer using trypan-blue exclusion. Cell counts were performed at the indicated time points. In these experiments cells were plated in RMPI-10 and also RPMI-0 (no serum). Every 48h, cells were counted and seeded at the original concentration, so that medium loss would not compromise their proliferation rate.

Assessment of cell viability. Determination of cell viability was performed by flow cytometry analysis of Forward Scatter versus Side Scatter (FSCxSSC) distribution using LSRFortessa cell analyzer (BD Biosciences). We have previously confirmed that this strategy measures lymphocyte viability as accurately as using Annexin V and propidium iodide staining [55].

Cell migration and invasion assays. Cell migration and invasion studies were done in 24-well transwell cell culture chambers with the upper chamber containing filters of 5μm pore size (Millicell-24 Cell Culture Insert Plate, polycarbonate, Millipore). Cells (10^5) were re-suspended in 100μl of RPMI either in the presence of 10% FBS (RPMI-10) or in the absence of serum

(RPMI-0) and added to the upper wells. In the bottom chamber, 800µl of RPMI-10 or RPMI-0 was added. In each migration or invasion experiment the following conditions were always used:

- RMPI-0 in the upper and lower chambers (R0>R0), used as negative control or indicator of the 'constitutive' migration/invasion of the cells. This condition is not always presented in the results graphs since during the incubation time no cells were found in the lower chamber.
- RMPI-0 in the upper chamber and RMPI-10 in the lower chamber (R0>R10), used as the experimental condition. FBS is used as chemoattractant to evaluate directional migration.
- RMPI-10 in the upper and lower chambers (R10>R10), used as the experimental assay. The chemoattractant is present in both chambers, so as to evaluate random migration.

Each condition was performed in triplicates per assay. Transwells were incubated at 37°C for the appropriated time (DND-41 cells two hours; CCR-CEM and MOLT-4 cells three hours). In the invasion assays, a layer of 200µl of 1mg/ml of Matrigel Growth Factor Reduced Matrix (BD biosciences) was placed in the upper chamber and it was let to solidify overnight at 37°C prior to the assays. The rest of the conditions for the assays are equal to the ones described for the migration assays, except for the incubation time that prolonged to seven hours for both DND-41 and MOLT-4 cells.

The number of migrating/invading cells was determined by counting the cells in the bottom of the lower chamber in a white field microscope (100x magnification). In each chamber, cells were counted in five non-overlapping observation fields. The number of cells per high-power field (HPF) was determined by the average of the five observation areas. In each experiment this number of cells was determined for three chambers per condition and the mean of the technical replicates was calculated. To calculate the migration index, the average number of cells per HPF was compared to those of the mock transduced cells and the fold difference of number of cells that migrated was calculated for each independent migration experiment. The average migration index of several independent experiments was used to plot the results and calculate the significance of the differences between the miR-146b-5p modulation and the mock control cells.

MiR-146b in the development of T-ALL *in vivo*. Ten age-matched (ten weeks old) NOD/SCID mice were used in this experiment: five were injected with CEM cells over-expressing miR-146b (CEM-146b OE) and other five were injected with mock transduced cells (CEM-Empty). The age-matched males and females were equally distributed by the two experimental groups. In the CEM-Empty group three males and two females were injected. In the CEM-146b OE group two males and three females were injected. Leukemia cells (10×10^6) were injected in the tail vein. Mice overall survival was analysed and humane endpoints established (20% weight loss or lethargy) at which the mice were euthanized. These endpoints were used to build the survival curve. The Kaplan-Meier estimator was used to determine the median rate of survival. The p-value was determined using the Log-rank (Mantel-Cox) test. To further assess T-ALL development, peripheral blood samples were collected weekly from the recipient mice (via facial vein puncture) starting two weeks after transplantation and analysed by FACS. The presence of human T-ALL cells was determined weekly by flow cytometry analysis of RFP positive cells. Count beads (5.000 per sample, BD Biosciences) were used to determine the absolute number of RPF+ cell per ml of blood. At the time of the humane endpoints mice were sacrificed with anesthesia overdose (Isoflurane) and selected organs were collected for FACS and histopathology analysis. For FACS analysis, the liver, spleen, lung, kidney, lymph nodes, thymus, long bones and blood (cardiac puncture) were collected. The solid organs were smashed to obtain a single cell suspension. The bone marrow was flushed-off with PBS-2%FBS from the tibia and femurs. When a clear presence of erythrocytes was detectable, samples were incubated ten minutes with RBCL buffer (eBiosciences), according to the manufacturer instructions. The erythrocyte free cell suspension was then acquired on a FACSAria III (BD Biosciences) to detect the presence of RFP+ cells. The cells were also stained for human CD45 (lymphocyte common antigen) to confirm the percentage of T-ALL cells.

Histopathology. Mice were sacrificed with anesthesia overdose and selected organs (liver, spleen, lung, kidney, lymph nodes, thymus, long bones, head and spinal cord) were harvested, fixed in 10% neutral-buffered formalin, embedded in paraffin and 3µm sections

were stained with hematoxylin and eosin (H&E). Tissue sections were examined by a pathologist, blinded to experimental groups, in a Leica DM2500 microscope coupled to a Leica MC170 HD microscope camera. Histological findings were assigned with a semi-quantitative severity score based on the following classification: minimal=1; mild=2; moderate=3; severe=4.

Flow Cytometry. Standard procedures were used to stain the cells with fluorochrome-conjugated antibodies or to verify reporter protein expression (GFP and RFP). The antibodies used in this study were CD1a-APC, CD3-PerCP-Cy5.5, CD4-PE-Cy7, CD8-FITC, CD45-APC (eBioscience). Samples were acquired in LSRFortessa cell analyzer (BD Biosciences), unless stated otherwise. The analyses were performed using the FlowJo software.

Statistical analysis. Statistical analyses were performed using GraphPad Prism version 6.01 for Windows (GraphPad Software). Statistical differences between mean values were evaluated using 2-tailed Student's t-test (paired or unpaired, as appropriate) or One-way ANOVA, when appropriate. Statistical differences in mice survival were evaluated using a Log-rank (Mantel-Cox) test. Differences were considered significant for p-values less than 0.05. When statistical significant differences were determined, the respective p-value or a representative symbol is depicted in the correspondent graphic (*$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$). Further information about specific statistical analysis is detailed in the above sections.

Chapter 3

RESULTS

Chapter 3.1

IDENTIFICATION OF MIRNA FAMILIES THAT REGULATE TAL1 EXPRESSION IN T-ALL

Nádia C. Correia[1], Alice Melão[1], Vanda Póvoa[1], Leonor Sarmento[1], Marta Gómez de Cedrón[2], Marcos Malumbres[2], Francisco J. Enguita[1], João T. Barata[1]

[1] Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisboa
[2] Cell Division and Cancer Group, Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid

Results

To address the possibility of a putative post-transcriptional regulation of *TAL1* by miRNAs happening during normal hematopoiesis, we used T-cell specific conditional Dicer-null mice [218]. The *Dicer* knockout results in substantial deletion of Dicer at the DN3 stage and complete deletion by DN4 stage. This results in a sharp reduction of miRNAs by the DP stage. Thymus of knockout mice have 10 times less TCR$\alpha\beta$ thymocytes than control counterparts, but maintain the normal DN cell numbers, explaining elevated DN percentage in these thymuses [218]. We found that in *Dicer* deficient mice the *Tal1* transcript, normally present until the DN3 stage [30], is increased by more than two fold in the total thymocytes in comparison to the thymocytes of Cre control mice (**Figure 3.1.1**). This suggests that TAL1 may be regulated post-transcriptionally by miRNAs during normal thymic development.

To identify putative miRNAs that target TAL1 we started by performing computational prediction of miRNAs that bind to *TAL1* mRNA. Computational algorithms have been the major driving force in predicting miRNA targets. Several web-based bioinformatics tools (PicTar (4-way), TargetScanS, miRanda (miRBase and microRNA), DIANA-microT algorithm, miRDB and StarBase) were used to perform the preliminary identification of putative regulators of TAL1 (**Supplementary Table 1**).

The criteria used by the different computational methods for miRNA target prediction vary widely, but most frequently they include: 1 – strong base pairing of the 5' seed of the miRNA (nucleotide positions 2–8 of the miRNA) to a complementary site in the 3' untranslated region (UTR) of the mRNA; 2 – conservation of the MRE (miRNA recognizing element); 3 - favorable minimum free energy (MRE) for the local miRNA/mRNA interaction; and sometimes 4 - structural accessibility of the surrounding mRNA sequence [298]. MicroRNAs normally recognize target sites present in the 3'UTR of the transcripts, and generally only this region of the mRNA is considered by the computational algorithms used. The TAL1 3'UTR in humans has around 3.4kb and harbors many possible MRE distributed through the entire sequence (data not shown). To our knowledge, it has never been described an alternative polyadenilation site in the 3'UTR of the *TAL1* gene that could produce a transcript with a shorter 3'UTR.

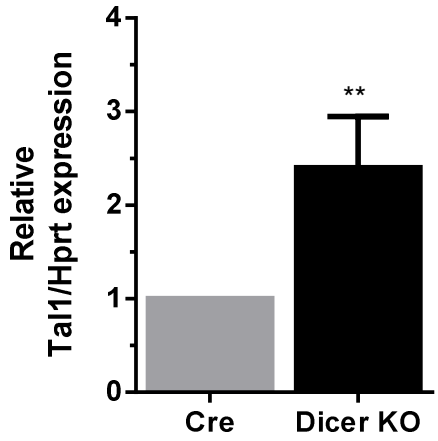


Figure 3.1.1 – ***Tal1* expression is increased in thymocytes from Dicer knockout mice.**

The expression of *Tal1* was determined by qRT-PCR and normalized to *Hprt* expression. The results are presented as fold difference of normalized *Tal1* expression in thymocytes from conditional Dicer ko mice, compared to Cre control mice. Here are represented four independent qRT-PCR experiments, each performed with RNA pooled from at least two aged matched mice per group. Statistical analysis was performed using a Student's t-test (**p<0.01)

From this analysis we compiled an initial list of 90 miRNAs that are candidates to regulate TAL1 mRNA (**Supplementary Table1**). Given the large number of candidate miRNAs we decided to rationally narrow down the list using the following criteria:

a) miRNAs under-expressed in TAL or LMO overexpressing cases.

b) concomitant identification of LMO2 as putative target.

c) more than one predicted target site in the 3'UTR of the TAL1 mRNA, and

d) 8mer (or 9mer) type of seed paring.

e) Identification of the miRNA as regulators of *TAL1* expression by at least two different algorithms.

Any miRNA predicted to fulfill at least one of these criteria was included for further testing. This way, we narrowed down the list of putative miRNAs targeting *TAL1* to 39.

One way to verify that a miRNA/mRNA target can happen intracellularly is through a reporter system with the mRNA under study, transiently transfected to a cell line together

with the miRNA of interest. The binding of a given miRNA to its specific mRNA target site represses the reporter protein production which can be compared to a control [299].

In order to validate the candidate miRNA/TAL1 mRNA interaction, we transiently co-transfected 293T cells with a commercially available reporter plasmid with *TAL1* 3'UTR immediately downstream of the luciferase open reading frame (pLuc-TAL1-3'UTR), together with the candidate miRNAs. These assays were performed with the miRVec library for expression of several microRNA species [286, 287].

We then verified if the reporter expression was decreased when compared to the control plasmid encoding a scramble (SCR) sequence, which is indicative of the miRNA biological activity against TAL1 3'UTR (**Figure 3.1.2**). For further analysis we selected microRNAs that significantly lowered the luciferase expression in 25-50% (**Figure 3.1.3**): miR-101, miR-520d-5p, miR-140-5p, miR-448 and miR-485-5p (check **Figure 3.1.4** and **Supplementary Figure 1** for miRNA binding details). We excluded miR-20a, miR-17 and miR-93 since they belong to the oncogenic miR cluster miR-17-92 (miR-17 and miR-20a) or to the same family (miR-93), and we excluded miR-410 and miR-199* due to the weak effect on luciferase expression.

Next, we mutated the MRE in the 3'UTR of *TAL1* in order to disrupt the miRNA/mRNA binding. We then re-evaluated the capacity of the respective miRNA to silence the reporter, which should be compromised. We co-transfected pLuc-TAL1-3'UTR or the vector carrying mutations in the miR binding sites into 293T cells together with each candidate miRNA.

We have mutated the only binding sequence for miR-101 in *TAL1* 3'UTR and verified that it restores the luciferase expression in the presence of the miRVec-101 (**Figure 3.1.4a** and **5a**). This demonstrates that the mutated sequence corresponds to the recognizing element of miR-101 in TAL1 3'UTR.

The miR-520d-5p has four predicted binding sites in the 3'UTR and we were able to mutate three of them. The triple mutants in TAL1 3'UTR restore partially but not completely the luciferase expression in the presence of miR-520d (**Figure 3.1.4b** and **5b**).

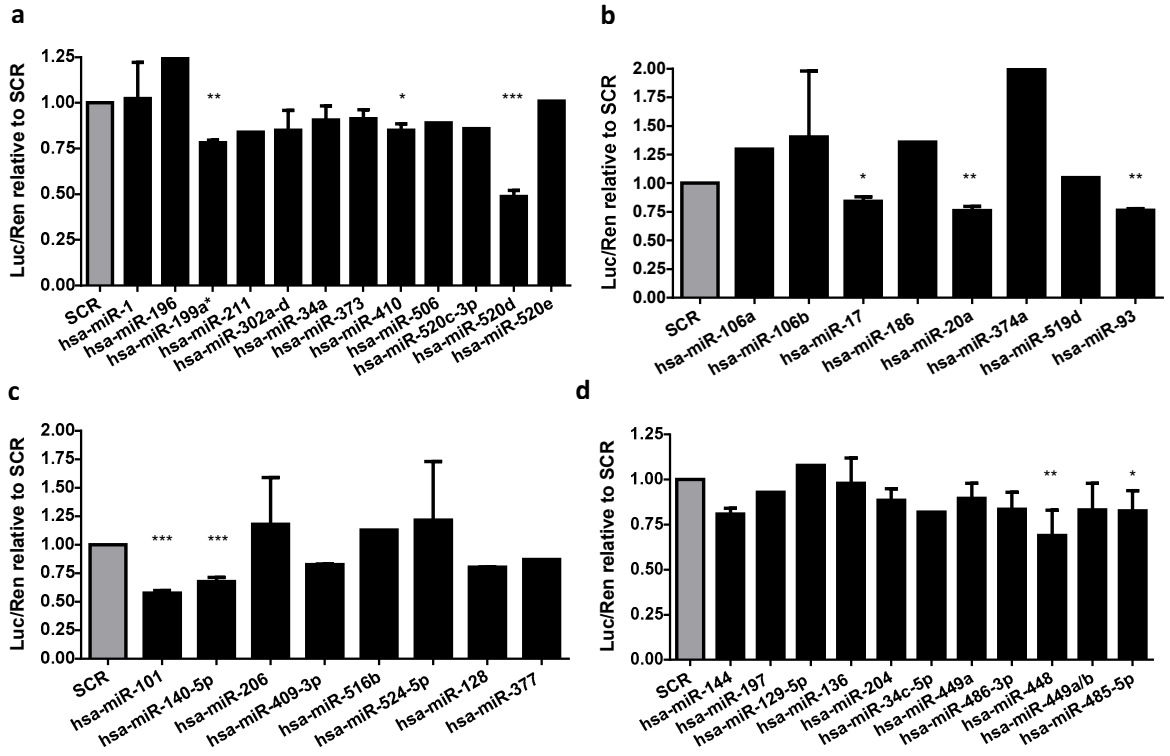


Figure 3.1.2 **– miRNA predicted to target TAL1 and the effect of their expression in TAL1 3'UTR.**

The graph illustrates the relative luciferase activity of a reporter construct carrying the TAL1 3'UTR downstream of the luciferase gene. The pLuc-TAL1-3'UTR was co-transfected into 293T cells with a vector expressing each of the indicated miRNA precursors or a scramble control sequence. Luciferase levels were normalized to Renilla levels, expressed from the same plasmid. All data are normalized to the luciferase levels generated upon co-transfection with the scramble sequence. The graphs represent at least two independent experiments with two replicates; statistical analysis was performed by One-way ANOVA (* $p<0.05$; **$p<0.01$;***$p<0.001$). a) miRNAs under-expressed in TAL1/LMO cytogenetic subgroup; b) miRNAs predicted to also target LMO2; c) miRNAs with more than one predicted target site and/or 8mer (or 9mer) type of seed paring; d) miRNA predicted to target TAL1 by at least two different algorithms

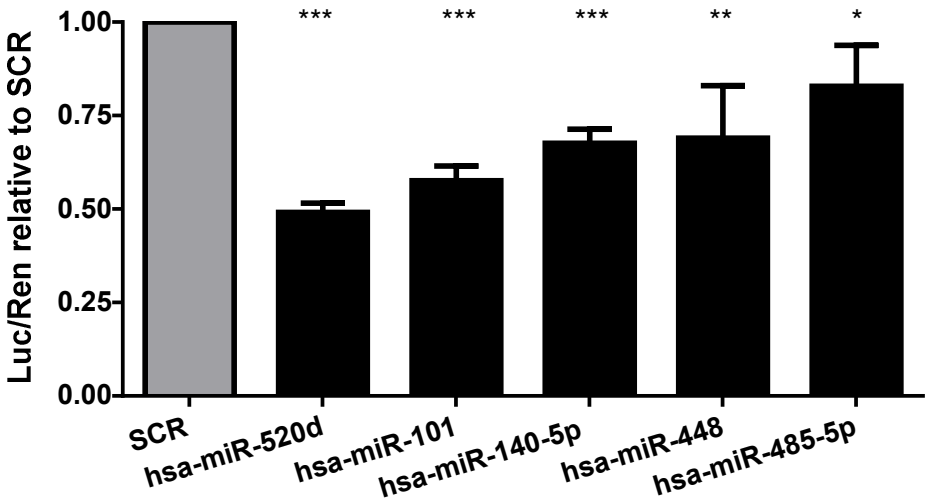


Figure 3.1.3 **– miRNAs that decrease luciferase activity by targeting TAL1 3'UTR.**

The graph illustrates the relative luciferase activity of a reporter construct carrying the *TAL1* 3'UTR downstream of the luciferase gene, as detailed in **Figure 3.1.1**. The graph represents at least 4 independent experiments with two replicates; statistical analysis was performed by One-way ANOVA (* $p<0.05$; **$p<0.01$;***$p<0.001$).

The miR-520d precursor can give rise to two mature forms of the microRNA, the 5p from the 5' arm of the hairpin and 3p from the 3' arm of the hairpin. Both mature forms can be expressed by miRVec vectors. Despite the fact that miR-520d-3p is only predicted by one algorithm (TargetScanS, **Supplementary Table 1**) to target TAL1 3'UTR, we mutated a putative binding site for miR-520d-3p in the 3'UTR (**Supplementary Figure 2a**). We verify that the mutation on the miR-520d-3p putative binding site can increase by 15% the luciferase expression when compared with the non-mutated 3'UTR (**Figure 3.1.5b**). This result suggests that the down-regulation of the reporter expression observed by miR-520d is partly explained by the 3p form. Despite strong efforts, we were not able to mutate the third miR-520d-5p MRE in TAL1 3'UTR (**Figure 3.1.4b**). One possible explanation for this is the fact that this MRE is placed in a region of the 3'UTR with a low (<35%) GC content, hindering the mutagenesis efficacy at this place. We cannot rule out the possibility that this last unmutated

MRE would be responsible for the full recovery of the luciferase expression that we could not achieve. Another possible explanation is that either this miRNA is not targeting directly the 3'UTR of TAL1, or is doing so by means of another MRE that does not have the canonical properties taken into account by the computational predictions.

The miR-140-5p has three predicted binding sites in the 3'UTR and we mutated all of them. By mutating just one of the three MRE we can already restore 90% of the luciferase expression (**Figure 3.1.4b** and **5c**) and the mutation of the three MREs fully restores the reporter expression. This demonstrated that these MRE are true recognition sites for miR-140-5p in the *TAL1* transcript. The miR-140 precursor can also give rise to two mature forms of the microRNA, the 5p and 3p. We also mutated two putative binding site for miR-140-3p in the TAL1 3'UTR (**Supplementary Figure 2b**), but this led to only 10% recovery of the reporter down-regulation, showing that the miR-140-3p is not the specimen responsible for the 50% reduction observed in luciferase expression upon introduction of miRVec-140 (**Figure 3.1.3** and **5c**).

If a given mRNA is a true target of a specific miRNA, then modulation of the miRNA concentration should result in changes in the amount of protein encoded by the target mRNA. Thus, in order to evaluate the physiological importance of the miRNA/TAL1 mRNA pairs we over-expressed the candidate miRNAs in T-ALL cell lines that endogenously over-express TAL1. This was followed by evaluation of the effect of microRNA expression on endogenous TAL1 mRNA and protein levels. To circumvent the difficulty imposed by low efficiency of transfection on the T-ALL cell lines, we co-transfected the miRVec with a GFP expressing vector (pMax) and cells were sorted to enrich for high GFP+ populations.

We observed that by over-expressing miR-520d, miR-101, miR-140, miR-485 and miR-448 in different T-ALL cell lines we down-regulate TAL1 expression at the transcript and protein level in a range of 20-60% (**Figure 3.1.6**). This range is in accordance to the predicted effects of microRNAs in protein expression [192, 193] and it varies depending on the cell line and microRNA specimen. A down-regulation mediated by these microRNAs in the *TAL1* transcript was only observed in PF382 cells (**Figure 3.1.6b**) for miR-520d, miR-101, miR-140 and miR-448 and SUP-T1 cells for miR-520d and miR-140. These results show that ectopic

expression of the selected miRNAs can physiologically target TAL1 by impairing the protein translation in T-ALL cell lines analyzed, without affecting the mRNA stability in the majority of the cases.

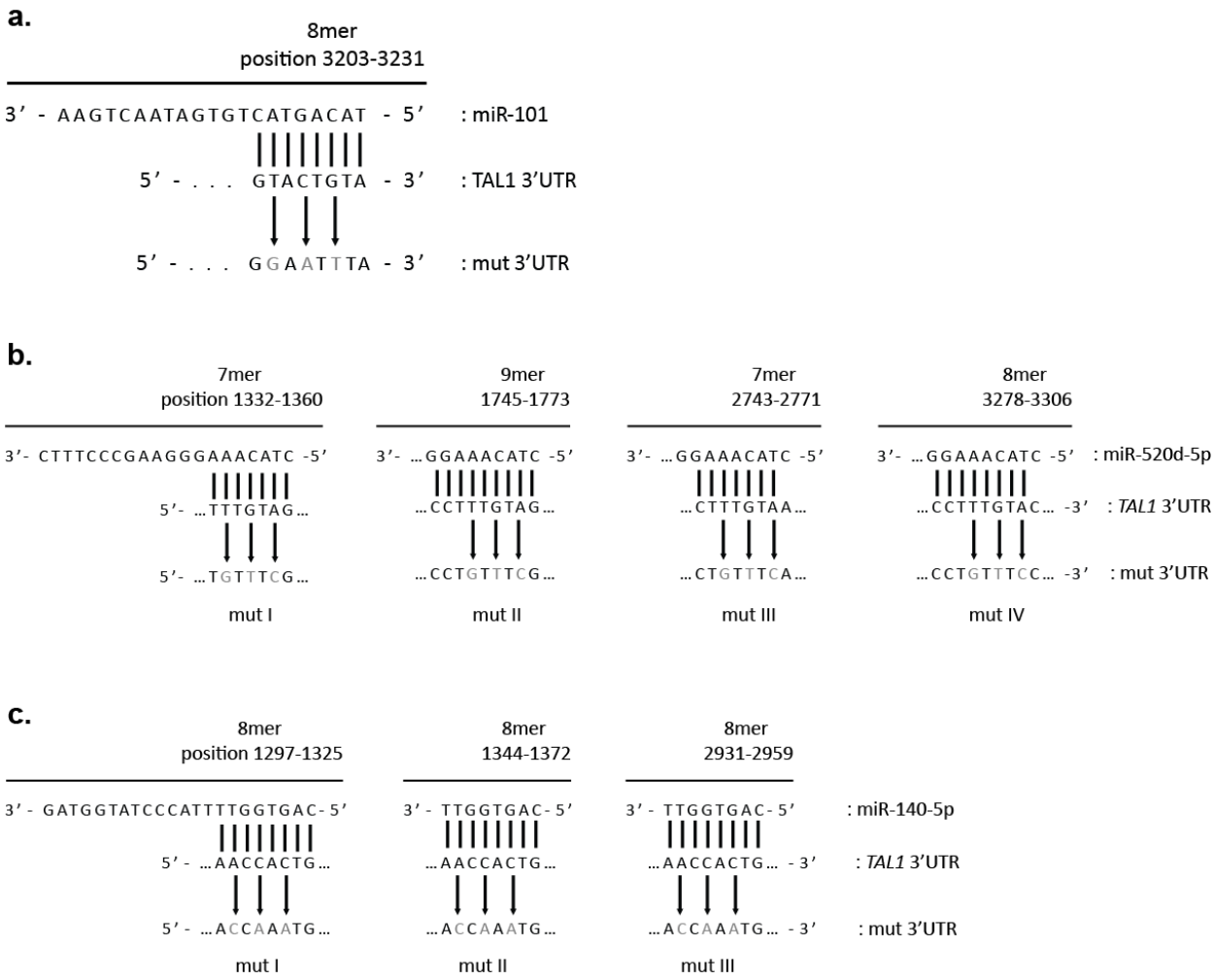


Figure 3.1.4 **– microRNA binding to TAL1 3'UTR and respective MRE mutagenesis.**

Schematic representation of microRNA binding to *TAL1* and the mutagenesis performed to disrupt miRNA seed binding. **a)** miR-101; **b)** miR-520d-5p and **c)** miR-140-5p. miRNA binding to *TAL1* 3'UTR details are depicted according to DianaMicroT [282] target prediction algorithm results.

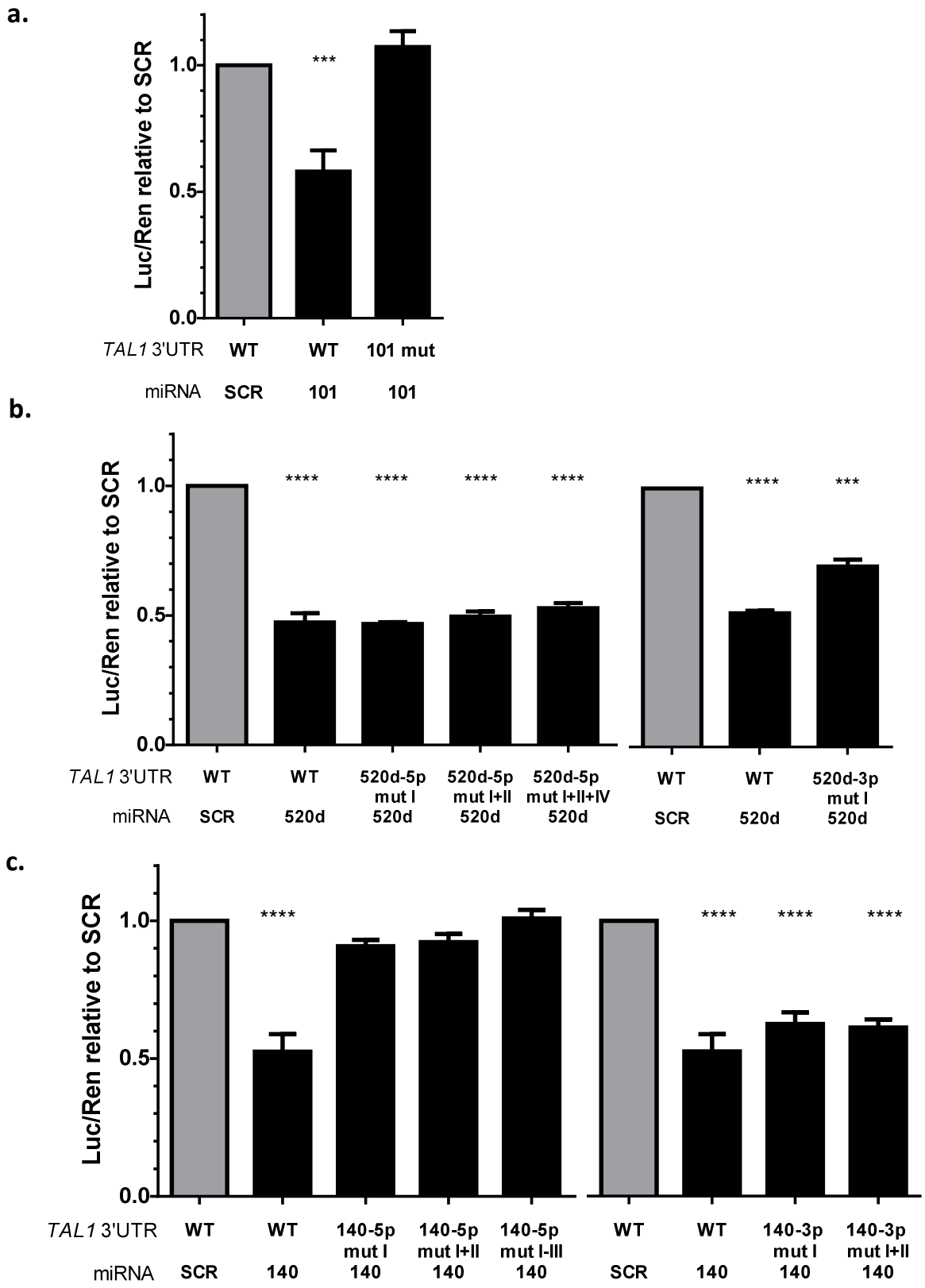


Figure 3.1.5 – **The effect of mutagenesis on microRNA-mediated repression of TAL1 3'UTR.**

Luciferase activity of wild-type (wt) or mutant (mut) TAL1 3'UTRs in the presence of scramble sequence (SCR) or the corresponding miRNAs. **a)** miR-101; **b)** miR-520d and **c)** miR-140. The mutations were made in cumulative manner, meaning mut II was performed on the 3'UTR bearing already mut I and so on. The mutations are named according to **Figure 3.1.4** scheme. The graphs represent 4 independent experiments with two replicates. Statistical analysis was performed by One-way ANOVA (***p<0.001; ****p<0.0001).

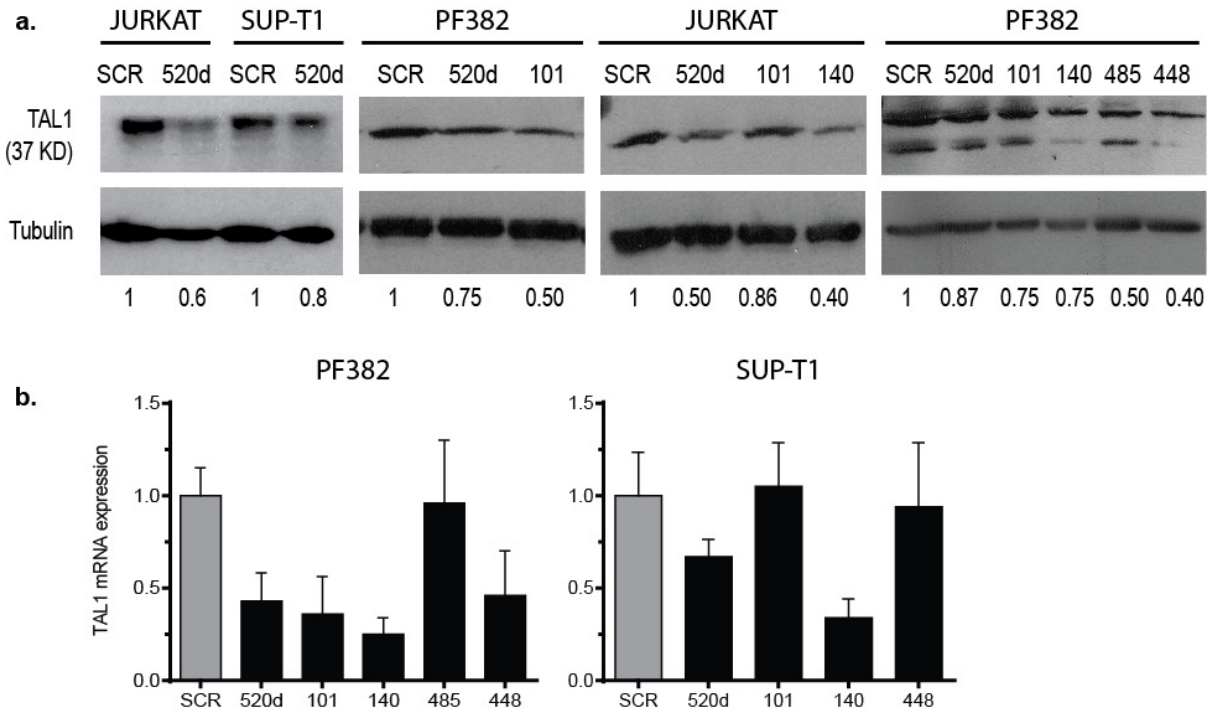


Figure 3.1.6 **– Ectopic expression of miR-520d, 101, 140, 485 and 448 decreases endogenous TAL1 mRNA and protein levels in T-ALL cell lines.**

a) Western blot and densiometric analysis in T-ALL cell lines upon transfection with miRVec vectors. The relative TAL1 protein levels in several T-ALL cell lines were normalized using α-Tubulin as loading control. Values presented result from densiometric quantitation of the bands and are normalized to that measured in the presence of the scramble vector. **b)** TAL1 transcript levels analysis by qPCR upon transfection with miRVec vectors of PF382 and SUP-T1 cells. Values indicate the mean ± lower and upper limit of three technical replicates relatively to the SCR transfection.

The previous experiment has the caveat of using ectopic expression of the miRNAs that goes beyond physiologic levels. To circumvent this, we also performed the reverse approach, where we transfected cells with high endogenous miRNA levels and correspondingly low TAL1 expression, with antisense oligoribonucleotides (ASO) to inhibit the function of the endogenous mature miRNAs and evaluate its impact on TAL1 mRNA and protein expression. We verified that inhibiting miR-520d-5p and miR-101 rescues endogenous TAL1 protein expression by 20 to 40% on average (**Figure 3.1.7a** and **7b**). This increase in protein expression was not always accompanied by a *TAL1* transcript increase (**Figure 3.1.7c**), which is in accordance with the previous experiments. Therefore, miR-520d-5p and miR-101 affect TAL1 mostly at the level of translation in T-ALL cell lines.

Interestingly when we compared by qPCR the expression of the microRNAs between TAL1 positive and negative T-ALL cell lines we verified that they are (miR-101, 520d-5p) or tend to be (miR-140-5p and miR-448) more expressed in the TAL1 negative cell lines, which favors our hypothesis that TAL1 over-expression in some T-ALL cases may result from, or be potentiated by, decreased expression of specific miRNAs (**Figure 3.1.8**). This idea is further strengthened by the fact that these microRNA genes are expressed in normal human thymic populations (**Figure 3.1.9a**) and that their expression is modulated during T-cell differentiation.

We next reasoned that if these microRNAs have a role in TAL1 over-expression in some T-ALL cases their levels should be decreased when compared to normal developing T-cells. We found that that miR-101 and miR-140-5p are less expressed in T-ALL patient samples than in more mature thymocytes. This is in accordance with the fact that TAL1 mRNA is not detected beyond the early thymic progenitor stage. Moreover, comparing the expression of miR-101, miR-140-5p, miR-448 and miR-485-5p with cells from normal bone marrow (**Figure 3.1.9b**), that still express TAL1 (data not shown), we verified that T-ALL patients have lower miRNA expression than the normal counterparts. This favors our hypothesis that down-modulation of miRNAs that target TAL1 could be partially responsible for deregulated TAL1 expression in T-ALL. These data are merely correlative, but provide further support to the notion that the identified miRNAs, which are down-regulated in T-ALL,

are likely upstream regulators of TAL1. Overall, our findings indicate that TAL1 expression is regulated by miRNAs.

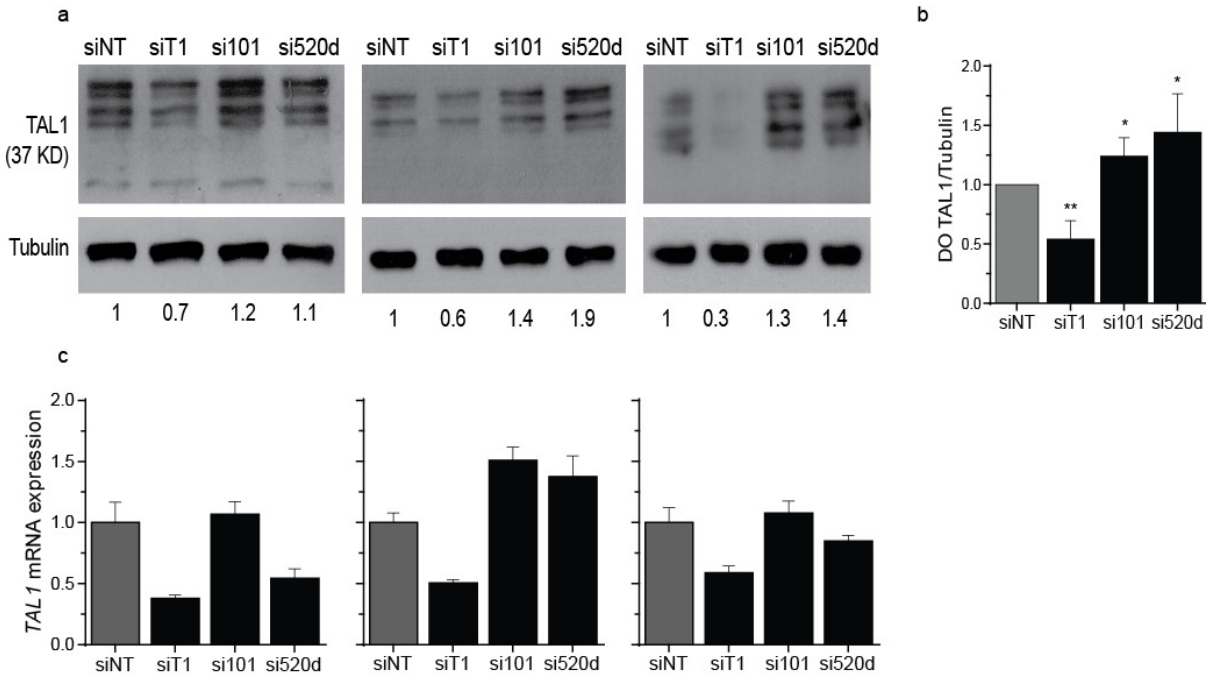


Figure 3.1.7 **–MiR-520d-5p and miR-101 inhibition increases endogenous TAL1 protein levels in T-ALL cells.**

Western blot, densiometric and qPCR analysis of TAL1 expression in CCRF-CEM cells upon nucleofection with microRNA inhibitors (si101 against hsa-miR-101 or si520d against hsa-miR-520d-5p), a siRNA against TAL1 (siT1) or a non-targeting siRNA control (siNT). **a)** 72h after nucleofection, cells were lysed and analyzed by immunoblotting for the expression of TAL1. Tubulin was used as loading control. Here are represented three independent nucleofection experiments. The protein bands intensity was quantified by densiometry and leveled by the tubulin expression. The numeric values depicted bellow represent the densiometric values normalized to the TAL1 expression in the siNT control for each experiment. **b)** Densiometric values (DO) of TAL1 expression were normalized to the Tubulin expression and compared to the control. Values indicate the mean ± standard deviation of four independent experiments and were analyzed using a Student's t-test (* $p<0.05$; **$p<0.01$). **c)** *TAL1* transcript levels analysis by qPCR of the same three independent nucleofection experiments depicted in a). Values indicate the mean ± lower and upper limit of three technical replicates relatively to the scramble nucleofection.

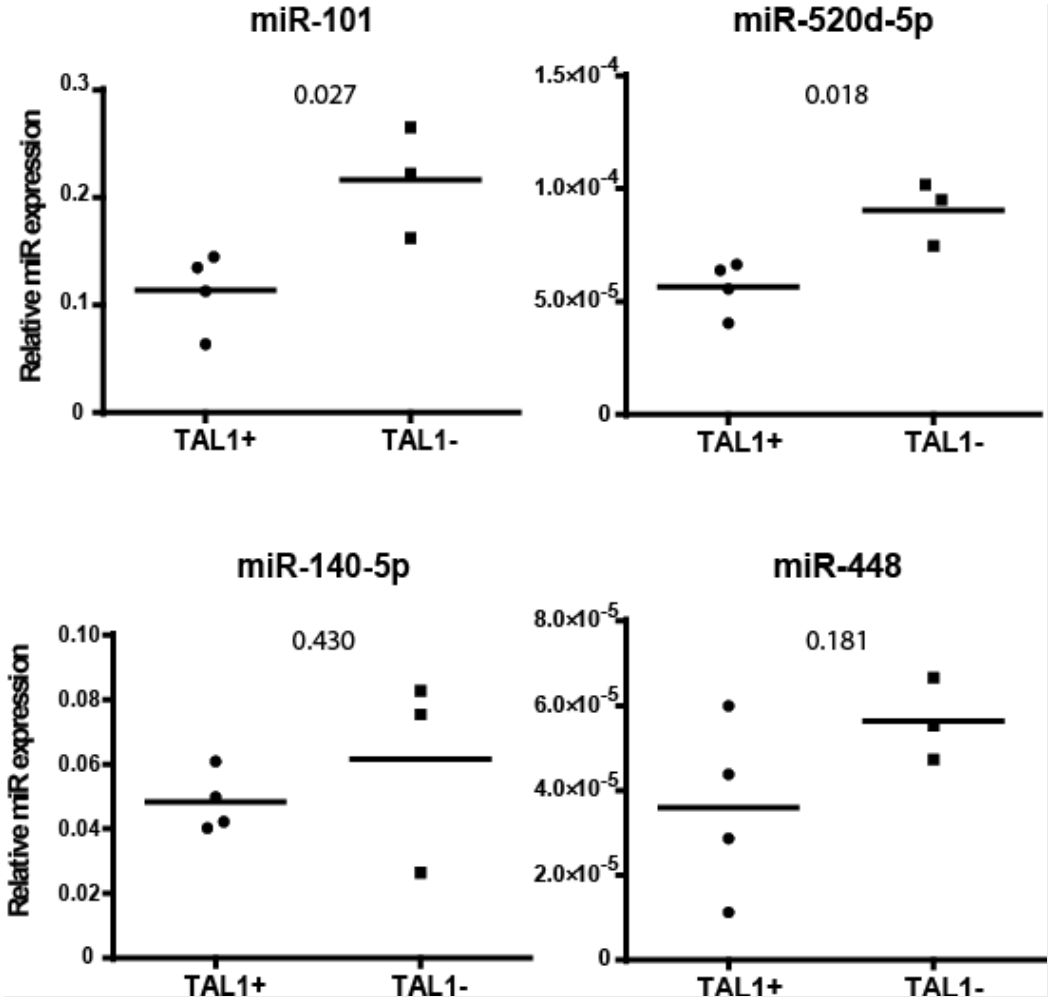


Figure 3.1.8 – **microRNA expression in T-ALL cell lines.**

The miR-101, miR-520d-5p, miR-140-5p and miR-448 expression was determined by qRT-PCR and normalized to *SNORD38B* expression in TAL1 positive (SUP-T1, CEM, TAIL7, PF-382) and negative (HPB-ALL, P12, TALL-1) T-ALL human cell lines. The numeric values depict the p-value of a Student's t-test comparing TAL1+ and TAL- cell lines.

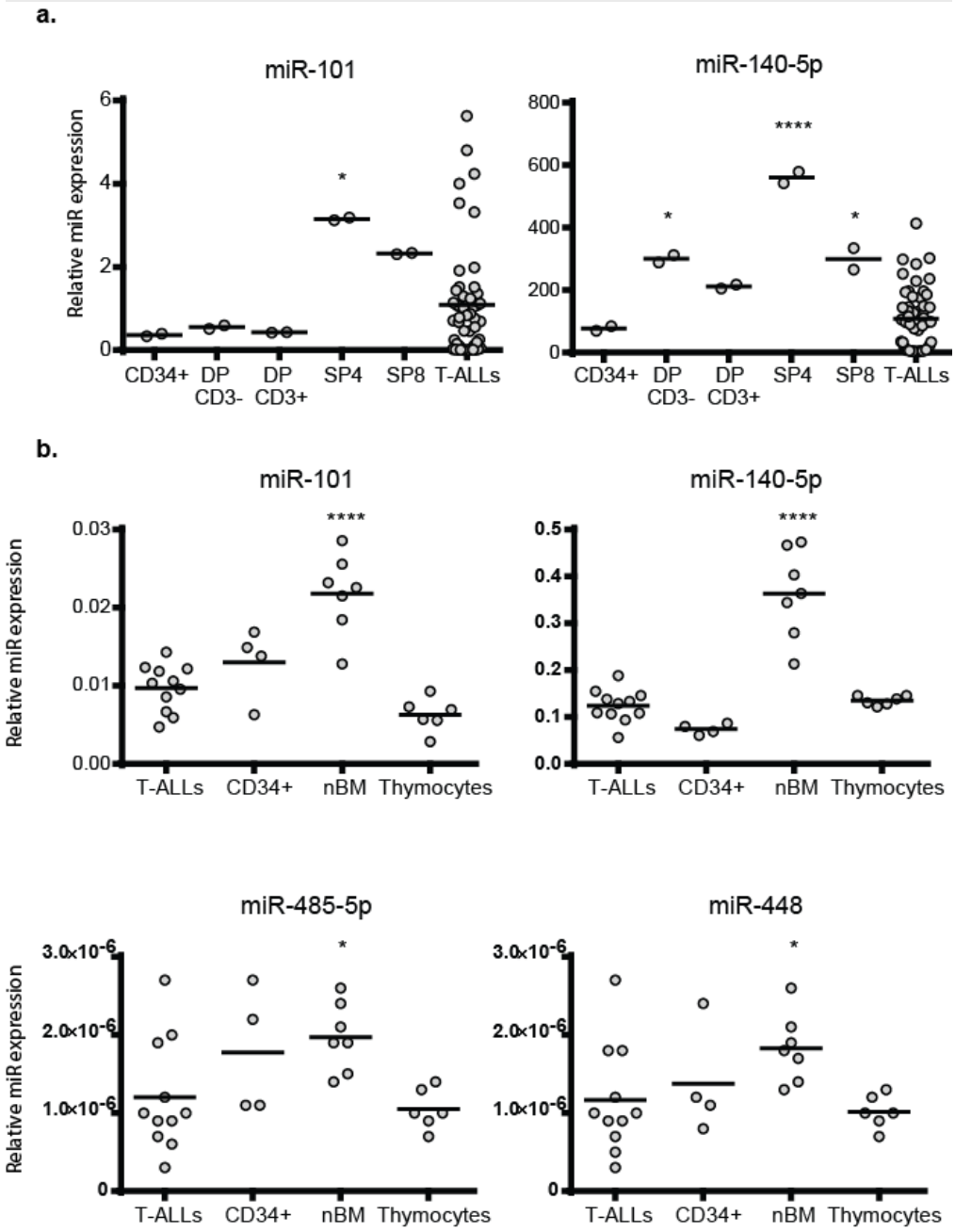


Figure 3.1.9 **– microRNA expression in T-ALL patients and normal counterparts.**

The miR-101, miR-140-5p, miR-485-5p and miR-448 expression was analyzed from publicly available data. **a)** The miRNA expression in T-ALL patients was compared to normal thymic populations (DP – double positive; SP4 – single positive CD4+; SP8 – single positive CD8+ T-cells). Data collected from [261]. **b)** The miRNA expression in T-ALL patients was compared to normal thymocytes, normal bone marrow samples (nBM) and CD34+ peripheral blood pediatric samples. Data collected from [255]. Statistical analysis was performed by One-way ANOVA (*p<0.05; ****p<0.0001).

Chapter 3.2

NOVEL TAL1 TARGETS BEYOND PROTEIN-CODING GENES: IDENTIFICATION OF TAL1-REGULATED MICRORNAS IN T-CELL ACUTE LYMPHOBLASTIC LEUKEMIA

Nádia C. Correia[1], Kaat Durinck[2], Ana Paula Leite[1], Maté Ongenaert[2], Pieter Rondou[2], Frank Speleman[2], Francisco J. Enguita[1], João T. Barata[1]

Adapted from: Correia, N.C. et al, Leukemia, 2013. **27**(7): p. 1603-6.

[1] Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa

[2] Center for Medical Genetics, Department of Pediatrics and Genetics, Ghent University, Ghent

Results

TAL1 is a *bona fide* T-cell oncogene [117]. The relatively small list of known TAL1 target genes included, at the beginning of this thesis, exclusively protein-coding genes. In addition, the aberrant transcriptional circuitry responsible for thymocyte transformation mediated by TAL1 is not yet fully understood. The aberrant expression of microRNAs has been reported in several hematological malignancies and in the case of ALL, microRNA expression signatures are able to delineate leukemia subgroups, [211, 274] and microRNA networks have been implicated in T-ALL [261]. Therefore, it was reasonable to expect that an oncogenic transcription factor such as TAL1 could also drive the expression of microRNA genes.

To identify a TAL1-dependent microRNA gene expression profile, we ectopically expressed TAL1 in the TAL1-negative T-ALL cell line P12 and performed low density array analysis, based on qRT-PCR. For this purpose, we transduced P12 cells with a bicistronic vector driving the concomitant expression of TAL1 and GFP or with the control mock vector. After RNA quality control, gene expression analysis for 372 human miRNA genes and 6 reference genes was performed in three independent samples of P12 mock and P12 expressing TAL1. From 204 microRNAs expressed in the cell line, we identified eight whose expression changed significantly upon TAL1 over-expression (**Figure 3.2.1a, Table 3.2.1, and Supplementary Table 2**).

We then validated these results by quantitative PCR analysis of each microRNA after enforcing or silencing the expression of TAL1 in TAL1-negative and TAL1-positive T-ALL cell lines, respectively. In such way, we confirmed that miR-135a, miR-223 and miR-330-3p were up-regulated by TAL1, whereas miR-146b-5p and miR-545 were down-regulated (**Figure 3.2.1b-f**). TAL1 over-expression in P12 and TALL1 cells resulted in decreased miR-375 expression, and PF382 and RPMI-8402 cells displayed increased miR-491-5p levels after TAL1 knockdown, partially confirming the array data, whereas no significant variation in miR-574 expression was observed (data not shown). These three microRNAs were excluded from subsequent analyses, since we considered as validated targets strictly those genes whose expression was regulated in the predicted manner by both over-expression and silencing of TAL1 in at least one cell line.

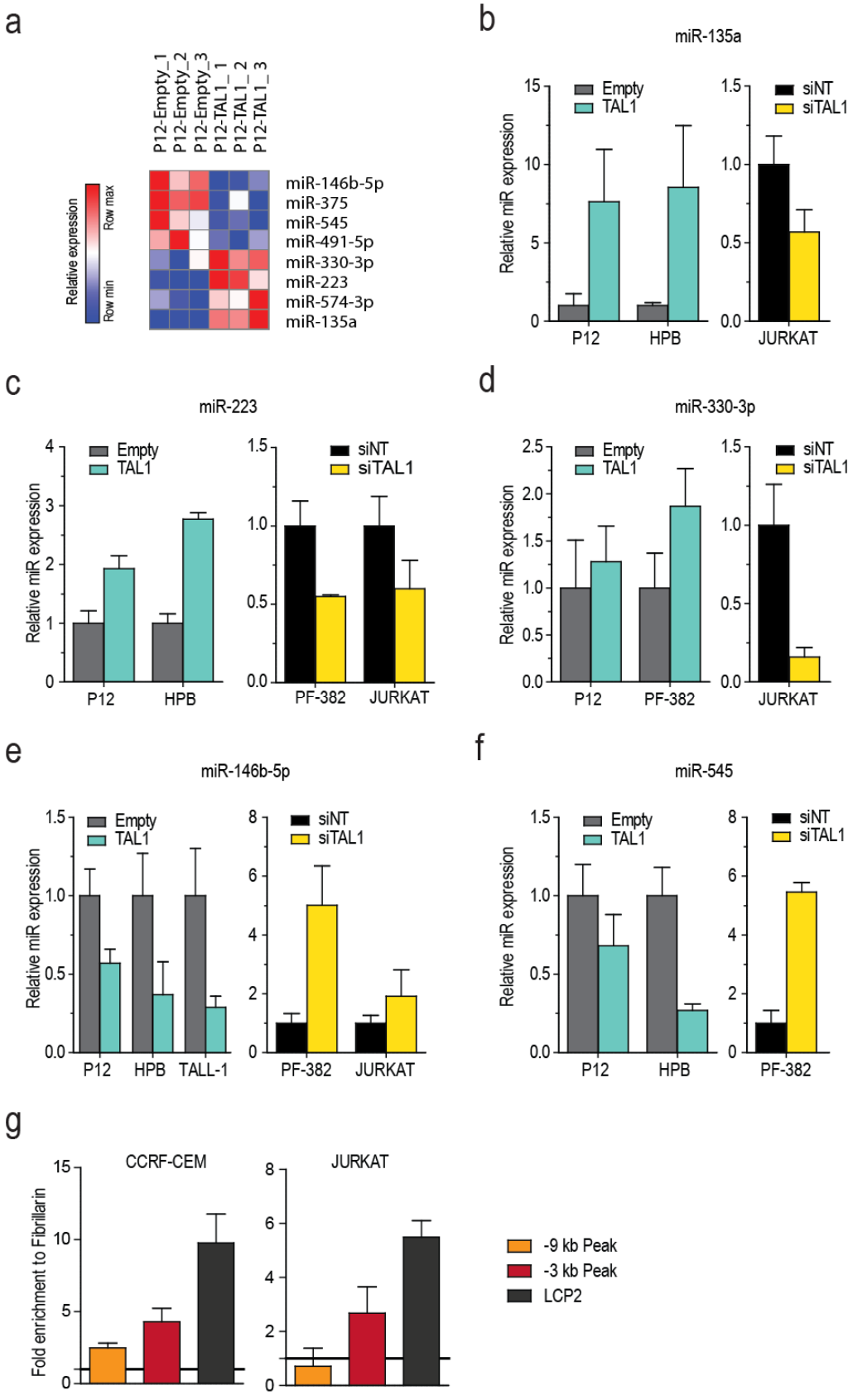

a
P12-Empty_1
P12-Empty_2
P12-Empty_3
P12-TAL1_1
P12-TAL1_2
P12-TAL1_3
Relative expression
Row max
Row min
miR-146b-5p
miR-375
miR-545
miR-491-5p
miR-330-3p
miR-223
miR-574-3p
miR-135a
b
miR-135a
Relative miR expression
Empty
TAL1
siNT
siTAL1
P12
HPB
JURKAT
c
miR-223
PF-382
d
miR-330-3p
e
miR-146b-5p
TALL-1
f
miR-545
g
CCRF-CEM
JURKAT
Fold enrichment to Fibrillarin
-9 kb Peak
-3 kb Peak
LCP2

Figure 3.2.1 **– Identification of TAL1 regulated microRNA genes.**

a) Heat map of differentially expressed microRNAs upon TAL1 over-expression. MicroRNAs were hierarchically clustered (rows, microRNAs; columns, experiments). See Table 3.2.1 for fold-difference values. Levels greater than or less than the mean are shown in shades of red or blue, respectively. **b-f)** qPCR validation of microRNA expression modulation by TAL1. Relative expression of hsa-miR-135a **b)**, hsa-miR-223 **c)**, hsa-miR-330-3p **d)**, hsa-miR-146b-5p **e)** and hsa-miR-545 **f)** normalized to *SNORD38B* in T-ALL cell lines with over-expression (left) or knockdown of TAL1 (right). The bars represent the mean±SD of three independent replicates. siNT - non-targeting siRNA **g)** TAL1 ChIP-qPCR in T-ALL cell lines. The occupancy by TAL1 of the genomic regions 9.2kb and 3.5kb upstream of miR-223 TSS was analyzed by ChIP-qPCR in JURKAT and CCRF-CEM cells. The promoter region of LCP2 was used as positive control for TAL1 binding and a random intergenic region was used as negative control. TAL1 binding is expressed as the fold enrichment relative to a mock ChIP performed against fibrillarin. The error bars represent the CI 95% of the fold enrichment. The horizontal line denotes the fold enrichment detection for the negative control.

Table 3.2.1 **– microRNAs differentially expressed upon TAL1 over-expression**

Rank	Upregulated in	Feature	score	P-value	fold change	Empty Mean	Empty Std	TAL1 Mean	TAL1 Std
1	TAL1	hsa-mir-135a	-5,147	0,009	12,796	0,074	0,016	0,941	0,152
2	TAL1	hsa-mir-223	-2,974	0,022	1,565	94,941	3,038	148,59	14,999
9	TAL1	hsa-mir-330-3p	-1,475	0,042	1,528	0,153	0,036	0,234	0,019
13	TAL1	hsa-mir-574-3p	-1,34	0,041	2,121	0,01	0,003	0,021	0,005
15	Empty	hsa-mir-491-5p	1,282	0,046	1,561	0,167	0,029	0,107	0,018
10	Empty	hsa-mir-545	1,475	0,054	1,533	0,067	0,001	0,044	0,005
4	Empty	hsa-mir-375	1,993	0,032	3,17	0,107	0,009	0,034	0,027
3	Empty	hsa-mir-146b-5p	2,11	0,009	1,906	0,092	0,012	0,048	0,009

Determined by the Comparative Marker Selection suite. Cutoffs for statistical significance were: 1) p-value < 0.05, and 2) fold change > 1.5. Std – standard deviation

Next, we evaluated whether the validated microRNAs were direct targets of TAL1 in T-ALL cells. To this purpose, we scrutinized publicly available TAL1 ChIP-seq data (GEO accession number GSE29181) for two T-ALL cell lines (JURKAT and CCRF-CEM) and two primary T-ALL samples [295] for the presence of TAL1 binding peaks up to 10kb upstream of the transcription start site (TSS) of each microRNA gene. We identified one peak in a putative promoter region for miR-146b, at approximately 11kbs upstream the miRNA TSS (**Figure**

3.2.2a), suggesting that this gene may be a transcriptional target of TAL1. Furthermore, two peaks were observed upstream of miR-223 TSS (**Figure 3.3.2b**). To confirm these findings, we performed TAL1 ChIP-qPCR in JURKAT and CCRF-CEM cells using primers designed for the genomic areas covered by the two peaks in the miR-223 locus. We verified that there is more than 2-fold enrichment, as compared to a mock ChIP performed against fibrillarin, in the amplified area within 3.5kbs upstream of the miR-223 TSS (**Figure 3.2.1g**). These results indicate that miR-223 is a direct target of TAL1 in T-ALL. We did not find evidence from the available TAL1 ChIP-seq data for direct binding of TAL1 to the remaining microRNA genes, suggesting that miR-135a, miR-330-3p and miR-545 might be indirectly regulated by TAL1, at least in the T-ALL cells analyzed.

Interestingly, analysis of microRNA gene expression profiles in different T-ALL subsets [261] revealed that TAL/LMO primary samples (integrating Sil-Tal1+ and LMO+ cases, which frequently express high TAL1 levels) display higher levels of miR-223 ($p=0.035$) and tend to express lower levels of miR-146b-5p ($p=0.092$) than other T-ALL cases (**Figure 3.2.3**), further confirming our screen results. In line with these observations, miR-223 appears to follow the same pattern of expression along normal human thymocyte development as TAL1 [94], with high levels in CD34+ T-cell precursors and sharp down-regulation in more differentiated subsets (CD4+CD8+ thymocytes) (**Figure 3.2.4a**). A similar pattern was observed for miR-135a (**Figure 3.2.4b**), in agreement with the notion that TAL1 positively regulates both genes. Also is worth to notice that the expression of miR-223 in the T-ALL primary samples is elevated when compared to thymocytes (**Figure 3.2.3c**), in agreement with the hypothesis of TAL1 being positively regulating the expression of microRNAs with oncogenic functions.

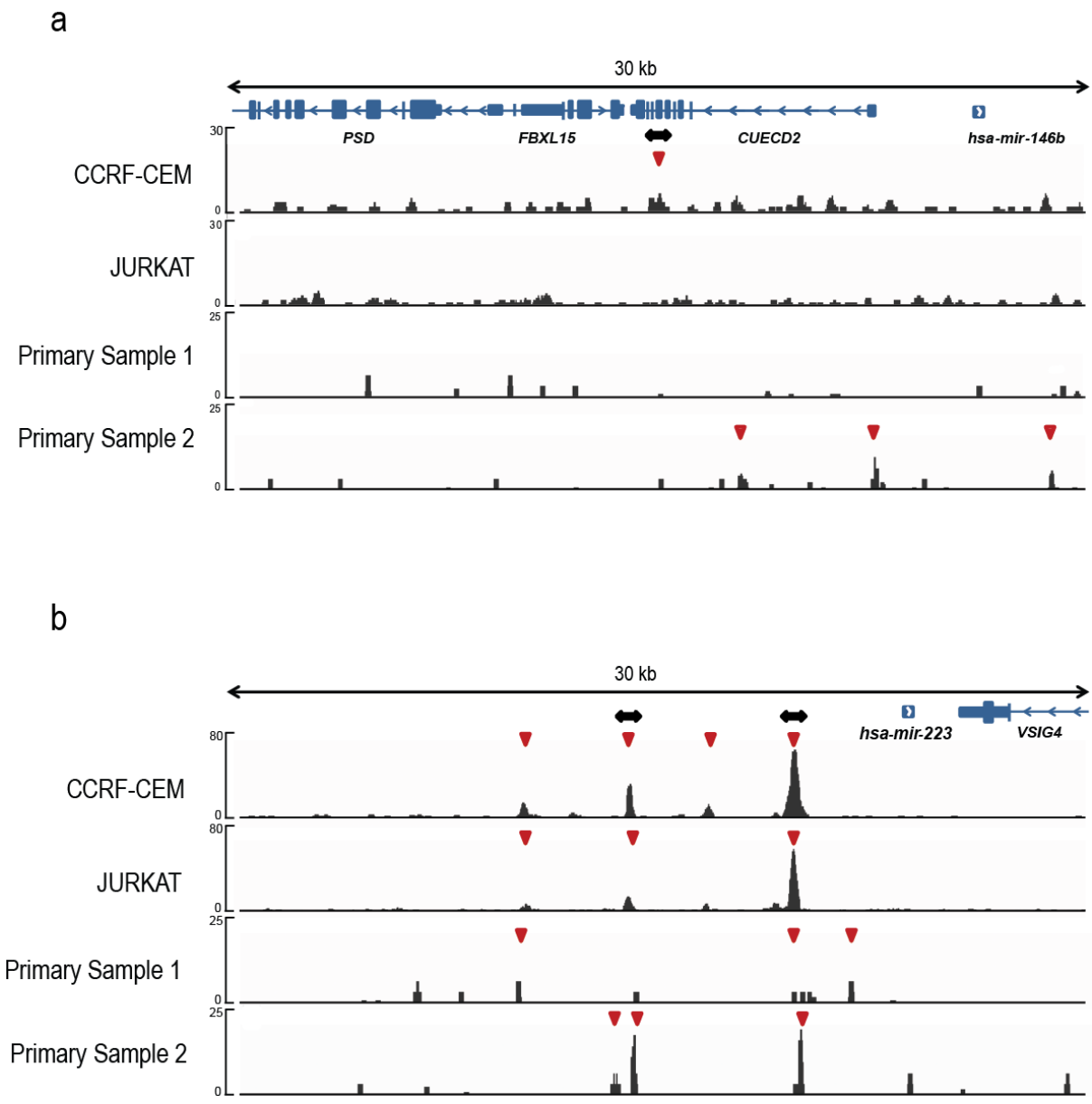


Figure 3.2.2 - **Evidence for direct binding of TAL1 to miR-146b and miR-223 loci.**

Analysis of publicly available ChIP-seq data for JURKAT, CCRF-CEM and two primary T-ALL samples (GEO accession number GSE29181). Representative Integrative Genomics Viewer (IGV) gene tracks show TAL1 binding peaks detected in the genomic area upstream of the miR-146b **(a)** and miR-223 **(b)** TSS, whose direction of transcription is indicated by an arrow. The arrowheads indicate regions bound by TAL1. The top horizontal bars indicate the scale in kilobases (kb). The black double arrows indicate the genomic areas to which primers were designed to validate TAL1 binding by ChIP-qPCR.

In contrast, miR-146b-5p is clearly up-regulated in the double-positive to single-positive transition (**Figure 3.2.4c**). The fact that the levels of miR-146b-5p associate with thymocyte maturation (**Figure 3.2.4c**) and the fact that T-ALL patients display lower levels of this miR compared to normal thymocytes (**Figure 3.2.3d**) is in agreement with a model whereby TAL1 over-expression during leukemogenesis inhibits miR-146b-5p and promotes T-cell developmental arrest.

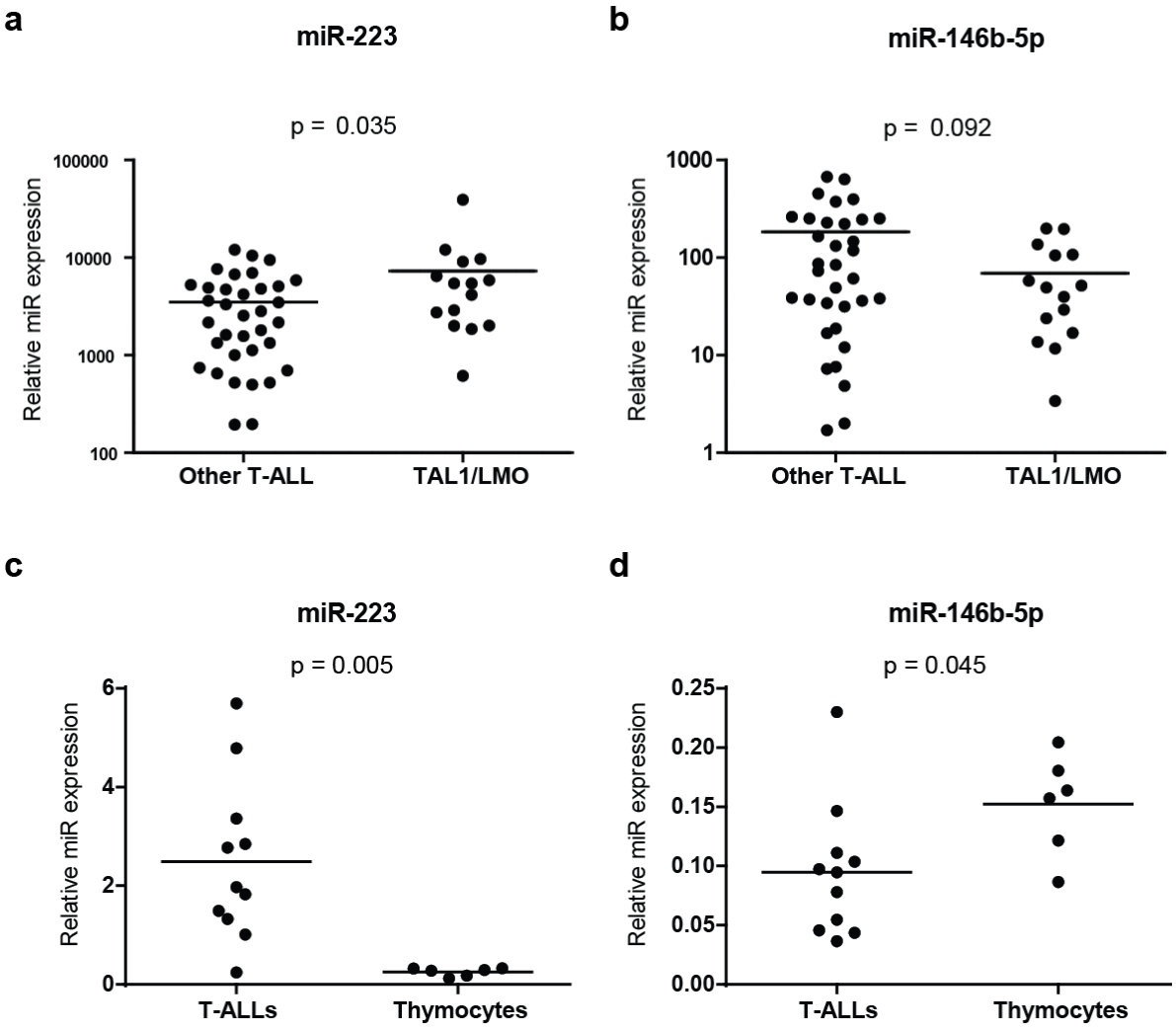


Figure 3.2.3 – **MicroRNA gene expression levels in T-ALL and thymocyte samples.**

TAL/LMO primary samples (integrating Sil-Tal1+ and LMO+) display higher levels of miR-223 **(a)** and lower levels of miR-146b-5p **(b)** than other T-ALL cases. T-ALL primary samples display higher levels of miR-223 **(c)** and lower levels of miR-146b-5p **(d)** than normal thymocytes. Data was collected from [261] (**a** and **b**) and from [255] (**c** and **d**) and further analyzed by us. p-values were calculated using a two-tailed Student's t-test.

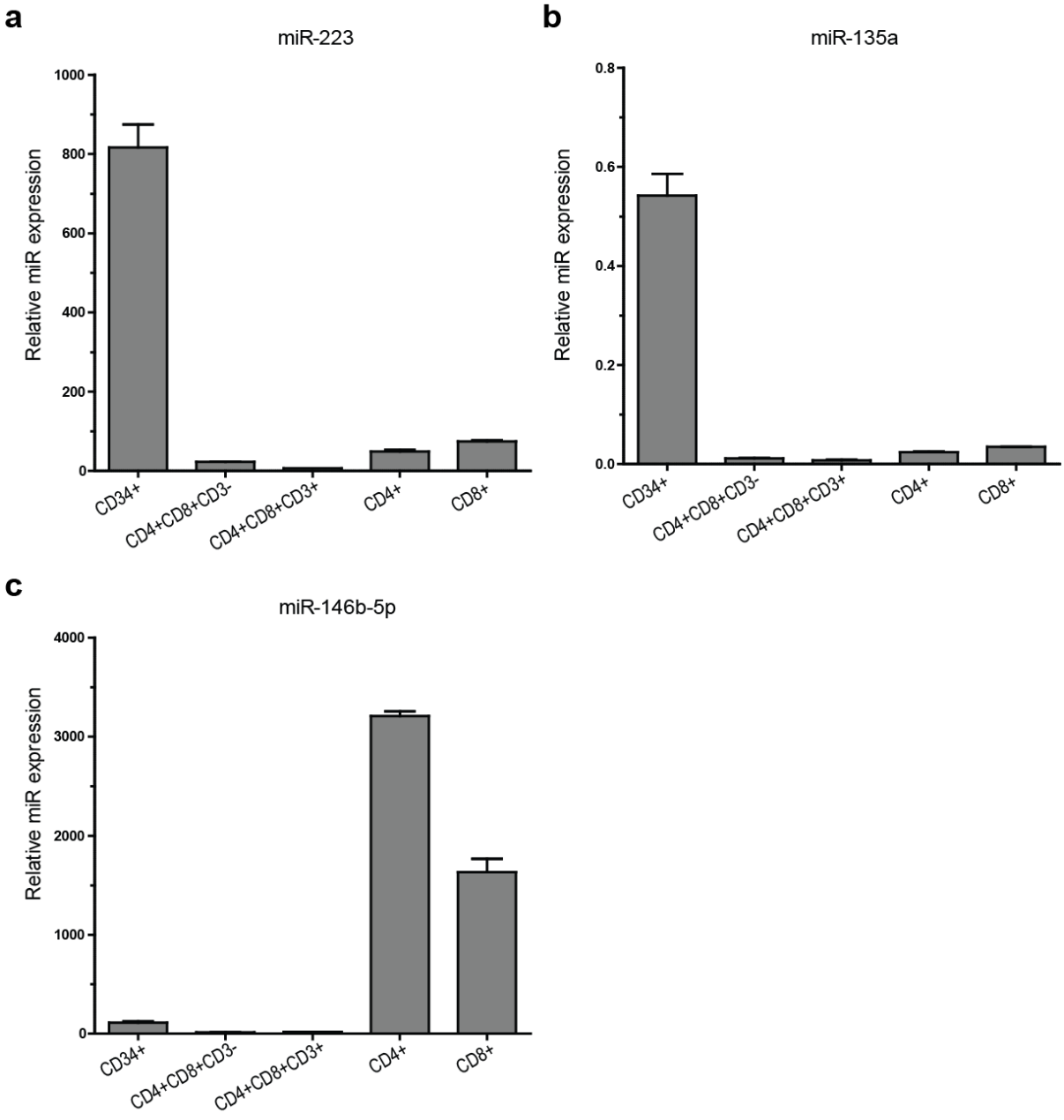


Figure 3.2.4 **– MicroRNA gene expression levels in normal human thymocyte samples.**

miR-223 **(a)** and miR-135a **(b)** are expressed in immature CD34+ cells and their expression is dramatically down-regulated in more mature CD4+CD8+ thymocytes. MiR-146b-5p **(c)** expression is clearly up-regulated in mature SP (CD4+ or CD8+) thymocytes. Data was collected from [261] and further analyzed by us.

Furthermore, we collected information about the validated target genes in human for each microRNA. In order to determine functional relevance of the sets of miRNA targets identified, we tested whether they were enriched for specific biological functions or pathways (**Table 3.2.2**). Gene set enrichment analysis showed an enrichment in biological processes related to inflammation (e.g., NF-kB signaling pathway and IL1/IL1R signaling pathway) and cancer (e.g. pathways in cancer), as detailed in Table 3.2.2 and Supplementary Table 3. Interestingly, the validated targets for the TAL1 down-regulated gene miR-146b-5p include IRAK1, TRAF6 and NFKB1 (all of which are involved in chronic inflammation), as well as the oncogene KIT (**Table 3.2.2**, **Supplementary Table 3** and **Supplementary Figure 3**), thus favoring a possible tumor suppressive role for this miRNA. In contrast, miR-330-3p, up-regulated by TAL1 reportedly targets E2F1 and CDC42, both of which are described to promote apoptosis in different cell types, including JURKAT cells, which is in line with a possible pro-oncogenic role for this miRNA. Moreover, the pro-leukemic role of mir-223 may be also achieved by down-regulating targets such as E2F1, FOXO1, RHOB or EPB41L3, which have been associated with induction of apoptosis and/or have tumor suppressive roles (**Supplementary Table 3**).

Moreover, we compiled a list of high confidence TAL1 positively or negatively regulated genes from publicly available data [143, 295]. We then intersected the predicted targets of TAL1 down-regulated microRNAs with the protein-coding gene targets previously demonstrated to be positively regulated by TAL1, and vice versa, and searched for common hits in both lists.

The data from this cross-examination of congruent TAL1 regulated protein-coding and miRNA target genes, are in line with the notion that the latter could be part of downstream networks collaborating in TAL1-mediated leukemogenesis (**Figure 3.2.5**). Indeed, most TAL1 up-regulated genes that have 3'UTRs predicted as targets for the TAL1 down-regulated miR-146b-5p and miR-545 have a known or putative oncogenic function, namely CD53, PDE3B, ETS-1 and MYB (**Figure 3.2.5a** and **Supplementary Table 4** for details and references). Also of note, three of the four genes (KRT1, Rapgef5, and JAZF1) with predicted targeting sequences for miR-146b-5p and miR-545 are associated with pro-tumoral functions.

In sharp contrast, the TAL1 down-regulated genes that are predicted targets for miR-135a, miR-223 and miR-330-3p display a clear abundance in (putative) tumor suppressors or in genes whose functions are compatible with anti-tumoral effects (SRGAP3, TOX, LRP12) (**Figure 3.2.5b** and **Supplementary Table 5**).

In summary, in this section our studies identify and validate for the first time a small set of non-protein coding TAL1 target genes, implicating microRNA genes as part of the transcriptional network downstream of TAL1 whose role may be important in the context of hematopoiesis and T-cell leukemogenesis. Some like mir-223 have an established oncogenic function in T-ALL, at least in part through FBXW7 down-regulation [261]. For the other microRNAs identified as being regulated by TAL1 still lacks functional information and their role in T-Cell malignancy should be detailed by future investigations, for which this work serves as a valued start point.

Table 3.2.2 **– Validated microRNAs regulated by TAL1 and their experimentally validated human targets.**

microRNA	Modulation by TAL1	Validated Target Genes	Top Enriched Gene Sets (Selection)
hsa-miR-135a	UP	APC, JAK2, NR3C2, FLAP	___
hsa-miR-223	UP	RHOB, NFIX, E2F1, MEF2C, NFIA,LMO2, STMN1, Arid4b, Il6, Lpin2, CHUK, FBXW7, IGF1R, S100B, LIF, SP3, EPB41L3, SLC2A4, IRS1	Genes down-regulated in MEF cells upon TGFB1 stimulation
hsa-mir-330-3p	UP	VEGFA, E2F1, NTRK3, CDC42, CD44	Pathways in cancer
hsa-miR-545	DOWN	LRP1	___
hsa-miR-146b-5p	DOWN	NFKB1, CDKN1A, MMP16, KIT, Card10, Scube2, TRAF6, IRAK1	NF-kB Signaling Pathway IL1/IL1R Signaling Pathway Pathways in cancer

Data obtained from mirTARbase 3.5, miRrecords and TarBase 6.0. Gene set enrichment analysis was performed as described in methods section.

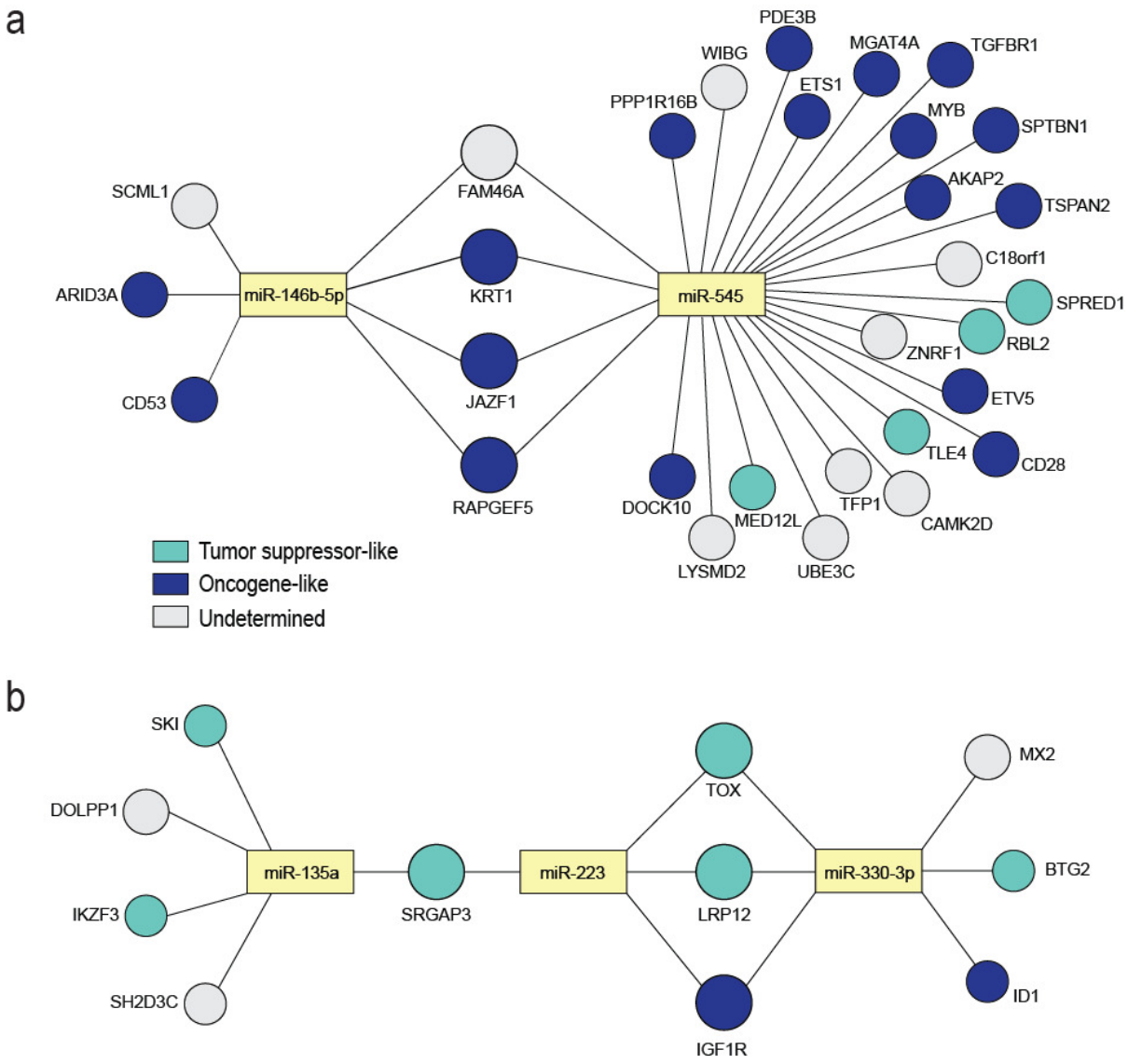


Figure 3.2.5 – **In silico determination of the potential participation of the newly identified TAL1 microRNA target genes in TAL1-mediated leukemogenic pathways.**

Cross-examination of congruent TAL1 regulated protein-coding and miRNA genes was performed as described in Methods section. **(a)** Down-regulated miRNAs and their predicted target genes previously shown to be up-regulated by TAL1. **(b)** Up-regulated miRNAs and their predicted target genes previously shown to be down-regulated by TAL1.

Chapter 3.3

MIR-146B IN T-CELL ACUTE LYMPHOBLASTIC LEUKEMIA: A POSSIBLE TUMOR SUPPRESSOR ROLE

Nádia C. Correia[1], Rita Fragoso[1], Vanda Póvoa[1], Francisco J. Enguita[1], João T. Barata[1]

[1] Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa

Results

Given that microRNA genes are regulated by TAL1 at the level of transcription, we hypothesized that some of these miRNAs can be part of the oncogenic network triggered by TAL1 ectopic expression in leukemia. One of the obvious candidates to explore was miR-223, since it has been recently shown to be one of the highest expressed microRNAs in T-ALL, able to accelerate the onset of the disease in a Notch1-induced mouse model. This effect is partially attributed to its ability to negatively regulate FBXW7 protein, a ubiquitin-ligase involved in Notch degradation [261]. Given that the effects of miR-223 deregulated expression have meanwhile been investigated by others [300], we decided to study other microRNAs regulated by TAL1.

We focused our efforts on miR-146b-5p and evaluated the functional and molecular effects of its deregulation by TAL1 in the context of T-ALL. This microRNA was found to be down-regulated by TAL1 in cell lines and to be less expressed in TAL1 positive T-ALL patients, so we hypothesize that it might have a tumor suppressive role in this disease.

In the previous chapter, we showed that T-ALL patients over-expressing TAL1 have a tendency to express lower miR-146b-5p levels than other T-ALL patients (**Figure 3.2.3b**). Publicly available data from a recent study [301, 302] allowed us to corroborate our results using an independent cohort of T-ALL patients. In fact, in this second cohort we verify that T-ALL patients over-expressing TAL1 (TAL subgroup) significantly express lower miR-146b-5p levels than other T-ALL patients (**Figure 3.3.1a**). Besides, T-ALL patients significantly express lower levels of miR-146b-5p compared to several normal counterpart cells (**Figure 3.3.1b** and **1c**). This suggests that reduced expression of miR-146b-5p might be important in the context of T-ALL, particularly in the TAL1 overexpressing cases.

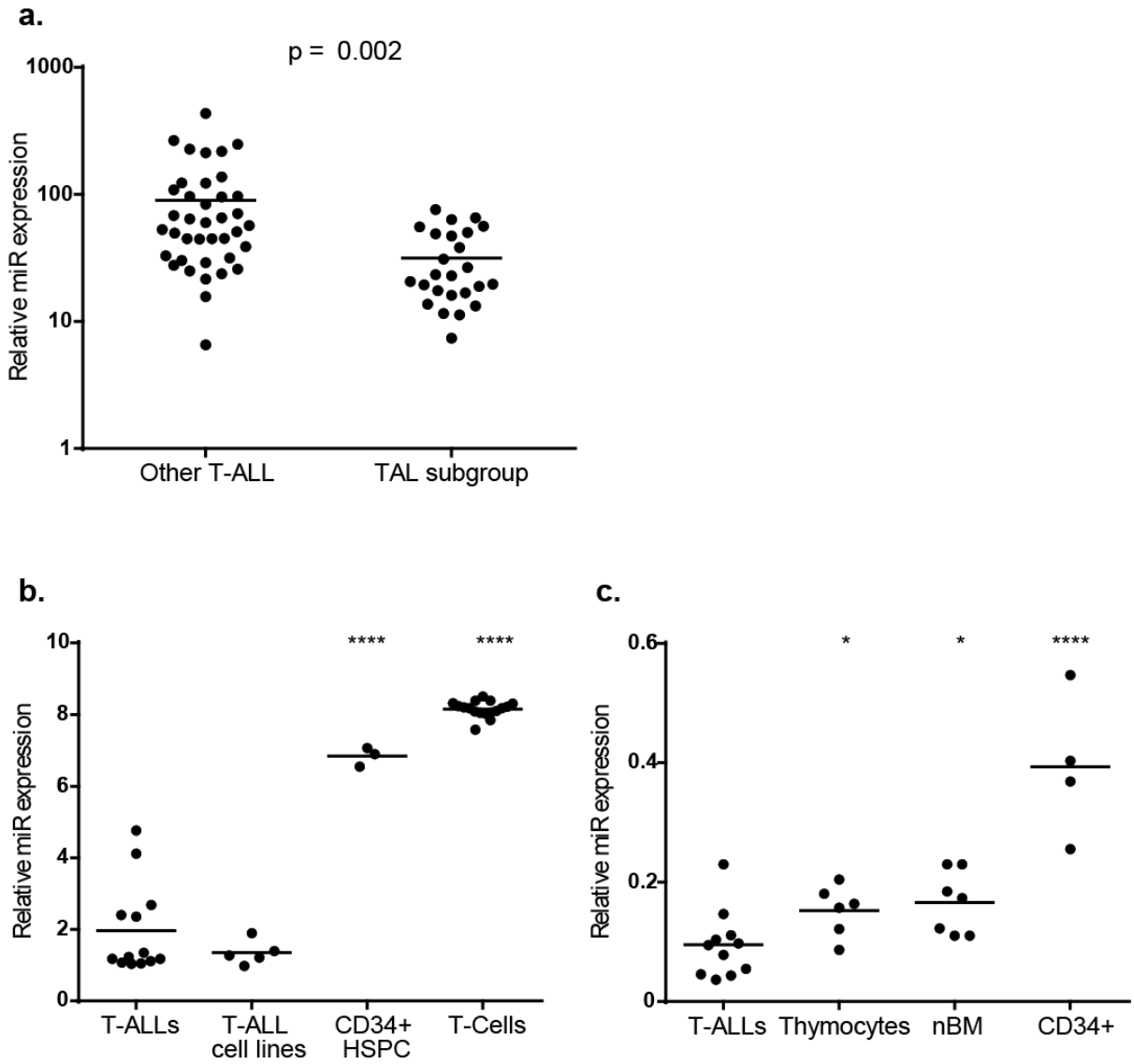


Figure 3.3.1 **– MiR-146b-5p expression is decreased in TAL1+ T-ALL patients and in T-ALL patients in general.**

The miR-146b-5p expression was analyzed from publicly available data. **a)** The miRNA expression was analyzed in a cohort of 64 T-ALL patients [302] comparing TAL1+ T-ALL cases (SIL-TAL, TCR-TAL and other TAL1+ cases) with T-ALL cases carrying other genetic abnormalities (TLX1, TLX3, HOXA and immature subgroups). Statistical analysis was performed using Student's t-test. **b)** The miRNA expression in T-ALL patients and cell lines was compared to normal T-cells or CD34+ HSCP cells from the peripheral blood of healthy donors. Data was collected from GEO database (GSE51908). **c)** The miRNA expression in T-ALL patients was compared to normal thymocytes, normal bone marrow samples (nBM) and CD34+ peripheral blood pediatric samples. Data collected from [255]. Statistical analysis was performed using One-way ANOVA (*$p<0.05$; ****$p<0.0001$).

We assessed the levels of the primary transcript of miR-146b in T-ALL cell lines. We reasoned that if miR-146b is a direct transcriptional target of TAL1, the levels of its primary transcript should correlate with TAL1 status. In fact, TAL1 negative cell lines tend to express more pri-miR-146b than TAL1 positive ones (**Figure 3.3.2a**). In addition, as already shown for the mature miRNA (**Figure 3.2.1e**), the knockdown of *TAL1* in a T-ALL cell line results in a marked up-regulation of the primary transcript (**Figure 3.3.2b**), reinforcing the hypothesis that TAL1 transcriptionally down-regulates miR-146b.

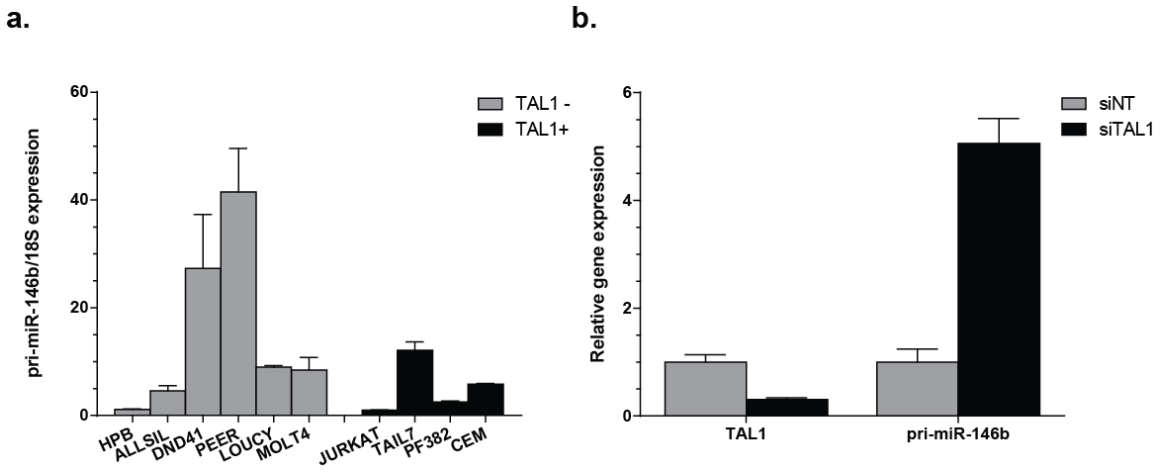


Figure 3.3.2 **– pri-miR-146b is less expressed in TAL1+ human T-ALL cell lines.**

a) The pri-miR-146b expression was assessed by qRT-PCR in several human T-ALL cell lines. The transcript levels were normalized to the ribosomal *18S* expression. Values indicate the mean ± lower and upper limit of three technical replicates relatively to the expression in the JURKAT cell line. The bars representing TAL1- cell lines are colored grey while the bars representing TAL1+ cell lines are colored black. **b)** CCRF-CEM cells were nucleofected with siRNAs against TAL1 (siTAL1) or a non-targeting control (siNT) and the expression of *TAL1* (left) or pri-miR-146b (right) transcript was assessed by qRT-PCR. Values indicate the mean ± lower and upper limit of three technical replicates relatively to the siNT nucleofection.

To verify the functional effects of miR-146b aberrant expression in T-ALL we manipulated its levels in T-ALL cells and assessed the effect on proliferation, cell cycle progression, viability, migration and differentiation *in vitro*. Since we have found that miR-146b is down-regulated by TAL1, T-ALL cell lines that do not express TAL1 (DND-41 and

MOLT-4) were stably transduced with a miRNA inhibitor to down-regulate the human miR-146b-5p expression, thus mimicking the effect of TAL1 (**Supplementary Figure 4a**). On the other hand, TAL1-positive T-ALL cell lines (JURKAT and CCRF-CEM) were stably transduced with a lentiviral vector to ectopically express the human pre-miR-146b (pLemiR-146b [297]), counteracting the effect of TAL1 on miR-146b-5p expression (**Supplementary Figure 4b**). We then analyzed the effects of miR-146b modulation on biological functions relevant for the T-ALL physiology, namely effects on proliferation, viability, migration and differentiation.

Surprisingly, we found that the down-regulation of miR-146b-5p on MOLT-4 or DND-41 cells does not promote any advantage in the proliferation of the leukemic cells. Likewise, the up-regulation of miR-146b on CCRF-CEM or JURKAT cells does not impair proliferation, which we postulated would be a consequence of a putative suppressive role for miR-146b. In fact, no accountable difference in cell proliferation assessed by cell counts (**Figure 3.3.3**) and thymidine incorporation (**Supplementary Figure 5**) was verified in those cell lines either in normal serum conditions (10% FBS) or serum starvation (0% FBS). The same results were obtained when assessing cell viability in the same conditions, either for miR-146b down- or up-regulation (**Figure 3.3.3**).

MiR-146b was previously shown to be associated with thymocyte maturation and highly expressed in mature single-positive thymocytes ([219] and **Figure 3.2.4c**). Thus, we speculated that miR-146b could have an effect on T-ALL cell differentiation. To assess this possibility, we monitored the immunophenotype (assessing the expression of CD3, CD4, CD8 and CD1a) of the cell lines stably over-expressing or stably down-regulating this miRNA during several weeks after their establishment. Once more, we found no obvious alterations (**Supplementary Figure 6**).

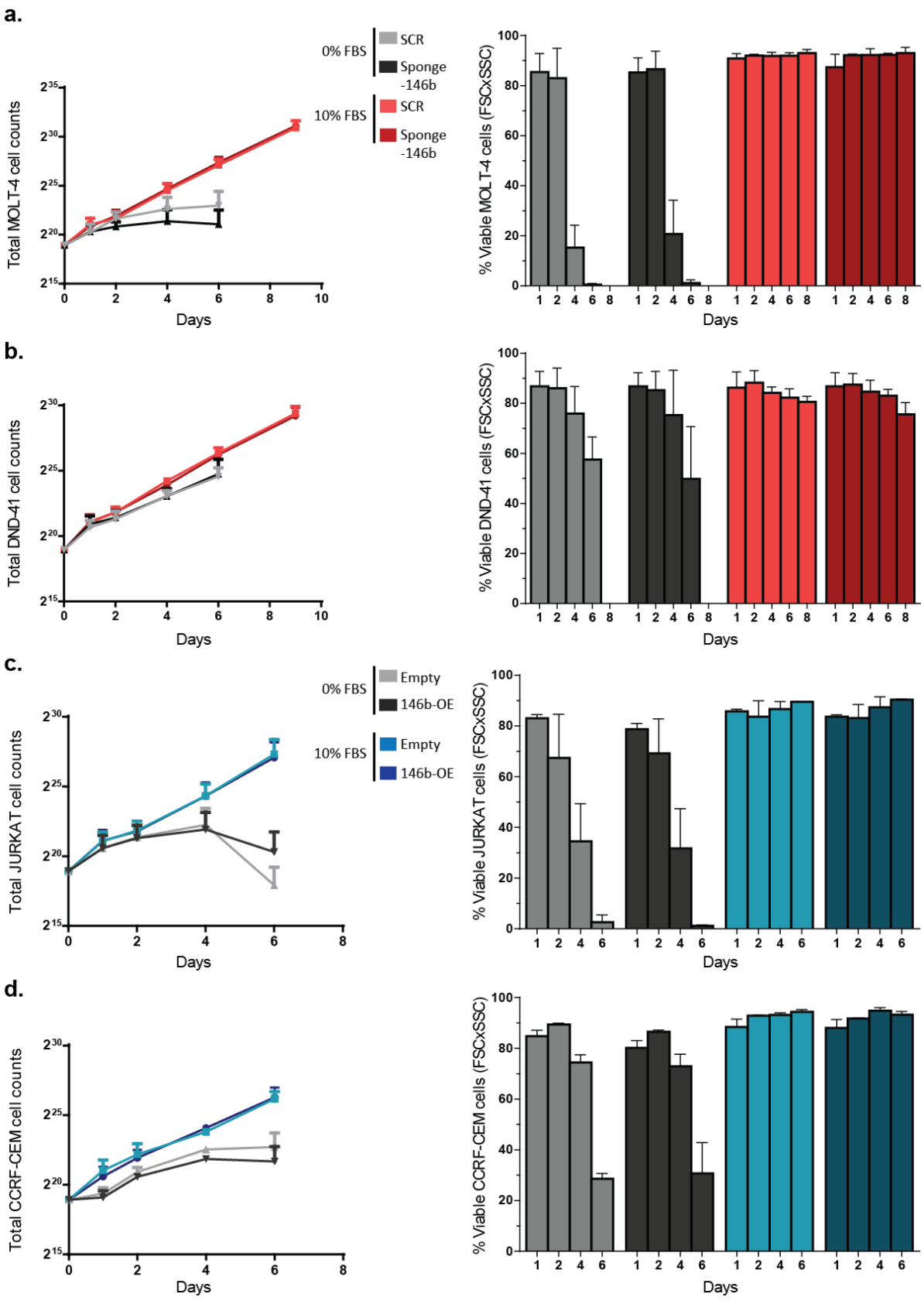

a.
Total MOLT-4 cell counts
Days
0% FBS
SCR
Sponge -146b
10% FBS
SCR
Sponge -146b
% Viable MOLT-4 cells (FSCxSSC)
Days
b.
Total DND-41 cell counts
Days
% Viable DND-41 cells (FSCxSSC)
Days
c.
Total JURKAT cell counts
Days
0% FBS
Empty
146b-OE
10% FBS
Empty
146b-OE
% Viable JURKAT cells (FSCxSSC)
Days
d.
Total CCRF-CEM cell counts
Days
% Viable CCRF-CEM cells (FSCxSSC)
Days

Figure 3.3.3 – **miR-146b modulation has no effect in proliferation and viability of T-ALL cells.**

Proliferation (left), determined by cell counts, and viability (right), assessed by flow cytometry, were analyzed at several time points either in normal medium conditions (10% FBS) or in serum deprivation (0% FBS). MOLT-4 **a)** and DND-41 **b)** cell lines expressing lower levels of miR-146b-5p (Sponge-146b) are compared to the mock transduced cell lines (Sponge-SCR). JURKAT **c)** and CCRF-CEM **d)** cell lines ectopically expressing miR-146b (146b-OE) are compared to the mock transduced cell lines (Empty). The bars graphs represent the mean (± SD) of three independent experiments.

Altered miR-146b expression has been previously linked to malignancy, specifically by altering migration properties of the cancer cells in solid tumors [297, 303-305]. Therefore, we evaluated the effects of miR-146b in the migration capacity of T-ALL cell lines. We verified that down-modulation of miR-146b-5p increases the number of DND-41 and MOLT-4 cells migrating in response to serum present in the medium (**Figure 3.3.4a** and **4c** and **supplementary Figure 7**). On the contrary, CCRF-CEM cells over-expressing miR-146b showed reduced migration capacities in the same conditions, compared to mock transduced cells (**Figure 3.3.4e** and **supplementary Figure 7**). Moreover, we performed invasion assays by introducing a matrigel layer between the T-ALL cells and the chemoattractant medium. In order to bypass this layer, T-ALL cells have to be able to invade the matrigel matrix. We showed that by decreasing miR-146b-5p levels in DND-41 and MOLT-4 cells we can enhance the invasion capacities of the T-ALL cells (**Figure 3.3.4b** and **4d** and **supplementary Figure 7**).

In order to investigate a possible tumor suppressor role for miR-146b in the development of tumors *in vivo* we used a murine xenograft model of human T-ALL that we have previously established [306]. We compared TAL1-positive T-ALL cells over-expressing miR-146b (CCRF-CEM) to mock vector-transduced T-ALL cells in their ability to develop tumors *in vivo*. For this purpose, CCRF-CEM cells ($10x10^6$ cells) were injected intravenously in immunocompromised mice (NOD/SCID). Subsequently, we evaluated leukemia progression and overall survival by analyzing the percentage of leukemic cells in the peripheral blood and time-to-death as indirect marker of leukemogenesis (the mice were euthanized when presenting clear signs of disease). At the time of death, the presence and quantification of

leukemic cells was further evaluated by analysis of histological sections and by flow cytometry analysis of single-cell suspensions of several organs (blood, bone-marrow, thymus, lymph-nodes, spleen, liver, kidney, lungs and CNS). We found that over-expression of miR-146b in human T-ALL cells significantly increased the survival of the mice (**Figure 3.3.5a**). We also observed a lower percentage of leukemic cells in the peripheral blood of mice transplanted with miR-146b OE cells compared to the mice transplanted with mock-transduced cells (**Figure 3.3.5b** and **Supplementary Figure 8a**). These results suggest that the increased overall survival of miR-146b OE mice might result from a delay in the disease progression, and are consistent with the hypothesized tumor suppressor role of miR-146b in T-ALL. At the time of death, no major differences in the infiltration of leukemic cells between the two groups of mice were observed (**Supplementary Figure 8 b)** and **c)**), with the exception of the lymph-nodes of miR-146b OE injected mice that had increased percentage of leukemic cells determined by flow cytometry. Besides that, histopathology analysis (**Supplementary Figure 8b)**) revealed a slight decrease in tumor load in the sections involving the CNS (Bregma, olfactory bulb and cerebellum) and the spinal cord in the mice injected with miR-146b OE cells. Overall, our results indicate that TAL1 down-regulates miR-146b, a microRNA with possible tumor suppressor functions in T-ALL.

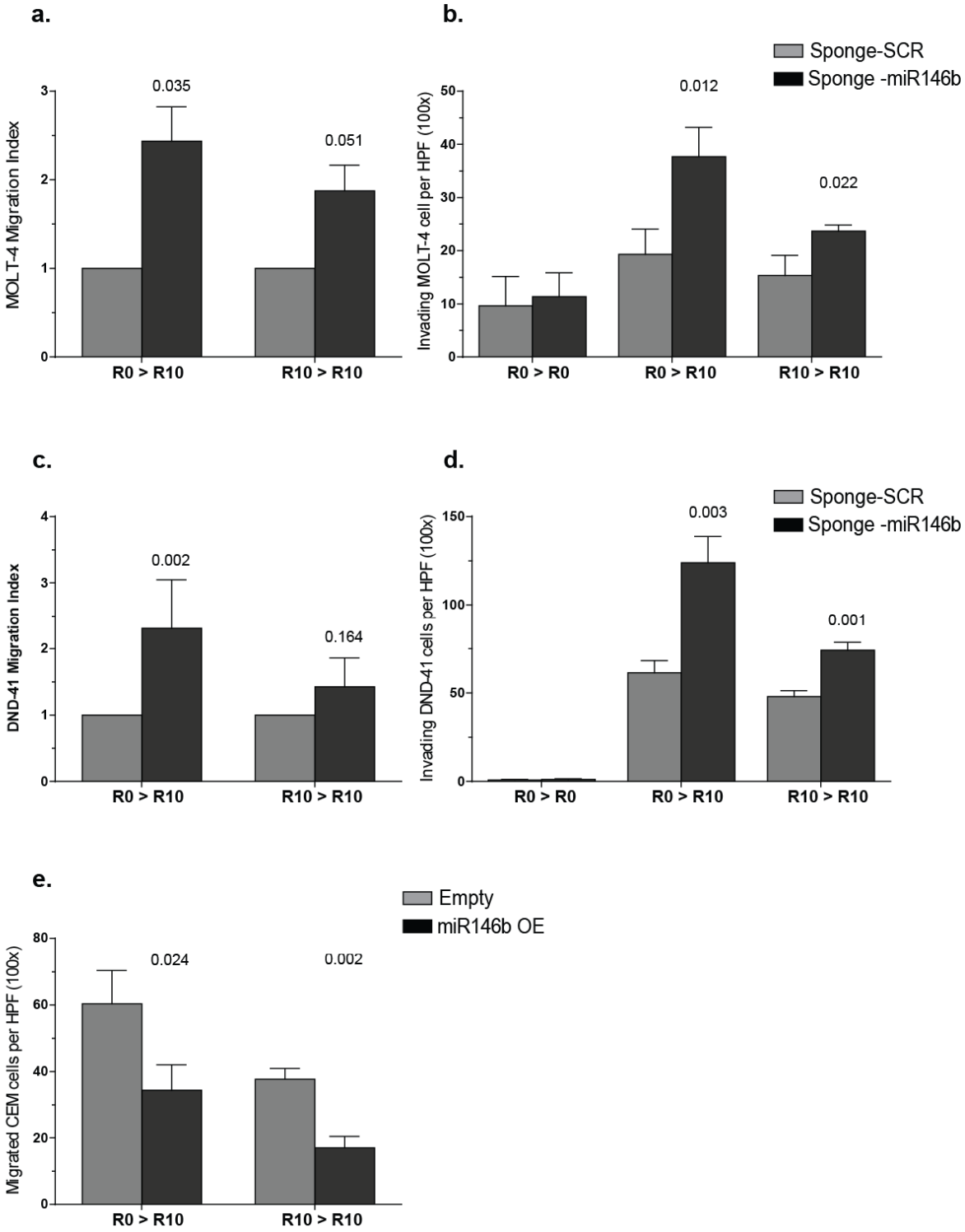


Figure 3.3.4 **– miR-146b-5p down-regulates migration and invasion of T-ALL cells.**

The migration capacities of cells with down-regulation (MOLT-4 **a)** and DND-41 **c)**) or with up-regulation (CCRF-CEM **e)**) of miR-146b expression were assessed through transwell migration assays. The invasion capacities of MOLT-4 (**b)**) and DND-41 (**d)**) were assessed by matrigel coated transwell assays. In both types of experiments serum was used as chemoattractant. A total of 100 x 10^3 cells were plated on the upper chamber of the transwell in culture medium (RPMI) either in the absence of serum (R0) or in the presence of 10% serum (R10). After the appropriate time lapse, cells that migrated/invaded to the lower chamber were counted on a microscope. The migration index of MOLT-4 (**a)**) and DND-41 (**c)**) cells represents the fold change of the number of migrated cells with low miR-146b-5p expression (sponge-146b) compared to mock transduced cells (sponge-SCR) in at least three independent experiments. The graphs **b)** (MOLT-4) and **d)** (DND-41) depict the average number of cells per HPF counted in three transwells of one representative experiment of invasion. The graph **e)** depicts one representative experiment of migration by showing the average number of CCRF-CEM cells over-expressing miR-146b (miR-146b OE) that migrated in three transwells, compared to mock transduced cells (Empty). Error bars represent SD and values depicted represent the p-values calculated using a two-tailed Student's t-test. The symbol > represents the direction of the migration, ie, R0>R10 means cells migrated from RPMI with 0% serum to RPMI with 10% serum.

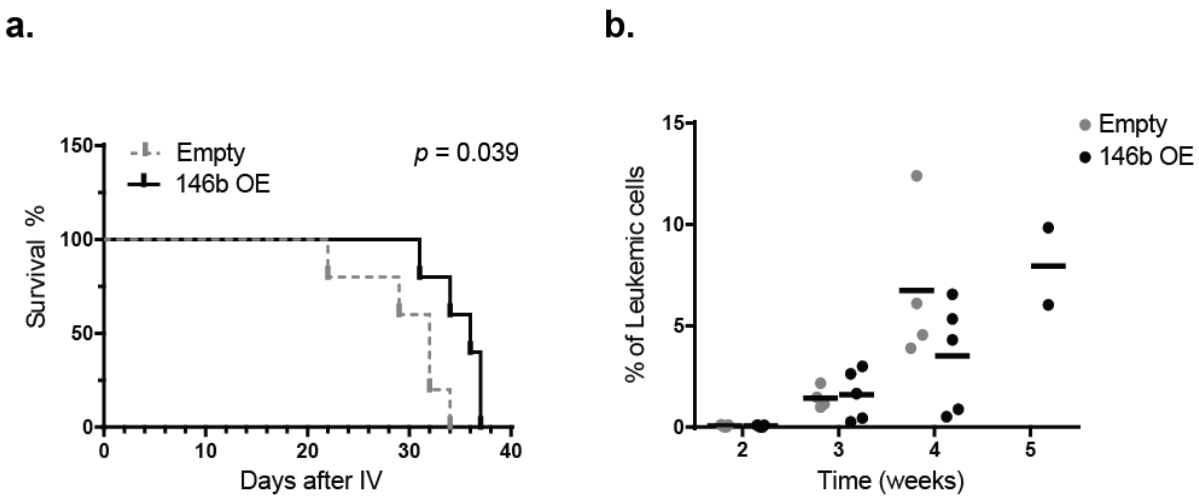


Figure 3.3.5 **- MiR-146b delays the development of T-ALL *in vivo*.**

a) Survival of mice (Kaplan-Meyer survival curve, n=10). NOD/SCID mice were intravenously (IV) injected with CCRF-CEM cells over-expressing miR-146b (miR-146b OE) or mock transduced cells (Empty). Statistical differences were evaluated using a Log-rank (Mantel-Cox) test. Median survival was 32 and 36 days post-injection in the miR-146b OE and control group, respectively. **b)** Percentage of human leukemic cells in the peripheral blood of the mice as determined weekly by flow cytometry evaluation of RFP+ cells. The lines represent the mean for each time point. For RFP+ absolute numbers check supplementary Figure 8a.

Chapter 4

GENERAL DISCUSSION

Discussion

T-cell acute lymphoblastic leukemia (T-ALL) is an aggressive childhood malignancy in which the transformed clone is arrested in T-cell development. Despite the significant improvements in treatment outcome, the challenge is to develop more efficient therapeutic strategies that target the leukemia cells specifically, this way diminishing the toxic effects of the current treatments and their long term effects. To achieve this, it is essential to build on our knowledge regarding the causes, pathophysiology and biology of T-ALL.

More than half of mammalian messenger RNAsare under selective pressure to maintain pairing to miRNAs [191]. For this reason, it may prove difficult to find a biological function or process that is not modulated by miRNA regulation. Leukemogenesis is not an exception [223] and several studies [211, 223, 251] have contributed to the profiling of miRNA genes in ALL patient samples. These led to the notion that specific miRNA gene expression signatures are associated with particular B- and T-ALL oncogenetic subgroups and indicate that the expression of certain miRNAs have prognostic relevance [307]. The participation of miRNA genes in T-ALL has been explored in recent studies and specific miRNAs have been implicated in T-ALL pathogenesis. Importantly, known oncogenes in the pathogenesis of T-ALL have been associated with deregulated microRNA networks in this disease context.

Is TAL1 regulated by microRNAs?

TAL1 is a transcription factor important for the regulation of early hematopoiesis, but rapidly down-regulated upon T-cell lineage commitment [308]. Moreover, aberrant expression of TAL1 in committed T-cell precursors is associated with leukemia development [92]. In fact, TAL1 is a well-established T-ALL oncogene being over-expressed in more than 60% of the patients [30]. For these reasons, we hypothesized that the pathways responsible for TAL1 repression during T-cell differentiation may be inactivated in T-ALL. Given the broad range of biological processes regulated by microRNAs we further hypothesized that the

pathways involved in TAL1 repression may be normally regulated by microRNAs and, de-regulation of these pathways in T-ALL could result in the increased expression of TAL1 observed in this disease.

We believe it is important for a better comprehension of the disease pathophysiology that researchers have a clear understanding of the mechanisms that lead to aberrant expression of TAL1 so frequently in T-ALL, and also if this process involves microRNAs. Targeting microRNAs to inhibit oncogenic signals is an attractive alternative to the targeting of oncogenes themselves, not only as an intrinsic mechanism of tumor suppression but also as a therapeutic strategy for cancer treatment. Oncogenes often have essential cellular functions and therefore are difficult to inhibit without causing substantial harmful effects to normal tissues. Hence, the use of microRNAs that naturally target pro-oncogenic proteins or whose down-regulation is necessary for the transformative progression might reveal an attractive strategy for cancer therapeutics.

Several observations led us to this hypothesis. First, more than 60% of T-ALL patients have ectopic expression of TAL1 mRNA but only 25% of those patients harbor DNA rearrangements that activate TAL1 transcription [104]. Thus, the majority of T-ALL cases do not harbor any detectable cytogenetic or molecular evidence of gene variant affecting TAL1 locus. Still, TAL1 is detected in patient blasts despite the fact that in human normal hematopoietic development TAL1 is not expressed from early thymic progenitors onwards [133]. Second, TAL1 is a putative target of several miRNAs that are up-regulated in hematopoietic stem cells, such as hsa-miR-17-5p, hsa-miR-197, hsa-miR-106 and hsa-miR-20 [309], and of some that are down-regulated in differentiated megakaryocytes, such as hsa-miR-106 and hsa-miR-20 [310], suggesting that miRNAs might regulate TAL1 at different stages of hematopoietic development. Lastly, the LIM only protein LMO2, a TAL1 co-factor, was shown to be regulated by microRNAs, namely during differentiation of erythroid cells by miR-223. The decline of miR-223 and the de-repression of LMO2 protein expression are important events in erythroid differentiation [311]. The majority of T-ALL patients with LMO ectopic expression also overexpress TAL1 [30] and both transcription factors cooperate to

induce leukemia in transgenic mice [92]. These facts contributed to our speculation that, similar to LMO2, TAL1 could be regulated by miRNAs during normal development.

Conditional Dicer deletion in early T-cell development results in sharp reduction of miRNAs by the DP stage, meaning that these cells are not capable of processing pre-miRNA transcripts into mature and functional miRNAs. As a consequence, these mice have ten times less TCR$\alpha\beta$ thymocytes [218]. The Dicer deletion does not affect DN cell numbers, explaining DN elevated percentage in these thymi. We have found that, when comparing *Tal1* expression in whole thymocytes of Dicer KO mice to the WT counterparts, *Tal1* levels are clearly augmented in thymocytes (Figure 3.1.1). This suggests that TAL1 is regulated post-transcriptionally by miRNAs during normal thymic development. Contrary to humans, the TAL1 transcript is found until the DN3 stage in the thymus of adult mice [118] and in the Dicer KO mice there is a substantial deletion of Dicer by the DN3 stage [218]. Given that the absolute numbers of DN cells are maintained in the Dicer transgenic mice [218], the increased *Tal1* levels observed in these mice must come from increased transcript copy number in each DN thymocyte, as a result of impaired microRNA regulation. Importantly, these data are well in frame with our hypothesis that TAL1 ectopic expression in T-ALL from currently unknown reasons could in part be mediated by abnormal down-regulation of particular miRNAs targeting TAL1.

To identify putative miRNAs that target TAL1 we performed computational prediction of miRNAs that bind to TAL1 mRNA. From the 39 tested, only ten microRNAs were able to significantly silence the reporter expression (miR-101, miR-520d-5p, miR-140-5p, miR-448, miR-485-5p, miR-20a, miR-17, miR-93, miR-410 and miR-199*). Therefore, the number of experimentally verified candidates is significantly lower than those generated by target prediction databases (10/39=26%). This confirms the significant level of false positives (74%) generated by computational prediction programs, a percentage that we predict would be higher if all the predicted miRNAs were tested. On the other hand, some studies using a microRNA library based screen with an UTR of interest, have found some miRNA-UTR interactions that are not predicted by the computational algorithms, representing false negatives [266, 302]. With our approach we could not detect these cases. An approach that

allows the testing of more candidate miRNAs to TAL1 targeting would probably uncover new inhibitory miRNA-TAL1-3'UTR interactions than the ones presented in this study. Nevertheless, the miRNA-UTR interactions need to be physiologically validated in the biological context of interest and preferably with the endogenous microRNA and protein of interest. In the case of our study, we hypothesize that TAL1 levels decrease during normal T-cell development at least in part due to miRNA-dependent down-regulation, and therefore TAL1 over-expression in some T-ALL cases could result from decreased miRNA expression. Our hypothesis foresees that these miRNAs would be decreased in the disease context. In other words, these miRNAs would have a tumor-suppressive role by targeting a frequent deregulated oncogene: TAL1. Hence, from the ten microRNAs able to significantly silence TAL1 3'UTR we selected microRNAs that significantly lowered the luciferase expression in 25-50% (Figure 3.1.3) and had putative tumor-suppressive functions: miR-101, miR-520d-5p, miR-140-5p, miR-448 and miR-485-5p. Therefore, we excluded from further analysis miR-20a, miR-17 and miR-93, since they belong to the oncogenic cluster miR-17-92 (miR-17 and miR-20a) or to the same family (miR-93). The cluster miR-17-92 is highly up-regulated in hematopoietic malignancies and has a clearly defined oncogenic function [226, 229, 258] which would be contradictory to a negative regulation of TAL1 in this context. Moreover, we excluded miR-410 and miR-199* due to the weak effect on luciferase expression.

It is well established that the role of a certain miRNAs in tumorigenesis might depend on the cellular context. In other words, in one cell type a miRNA might have a tumor suppressor function, while in a different cell type it might promote tumor development. Nevertheless this appears not to be the case for miR-101 as, so far, several studies demonstrated a tumor suppressive function for this miRNA in the context of several cancer types. In fact, a brief search in an updated database that annotates the experimentally verified oncogenic and tumor-suppressive miRNAs (OncomiRDB) [312] shows that miR-101 has a tumor-suppressor effect in more than seven types of solid tumors, mainly by inhibiting cell proliferation, migration, invasion and tumor growth. The principal validated target for miR-101 is the histone methyltransferase EZH2, a component of the polycomb repressive complex 2, involved in epigenetic silencing. In fact, in prostate cancer miR-101 expression

decreases during cancer progression, paralleling an increase in EZH2 [313]. Decreased miR-101 expression and direct targeting of EZH2 has been revealed in several other tumors, namely in bladder [314], gastric [315, 316], and renal cell [317] carcinomas, pancreatic ductal adenocarcinoma [318] and melanoma [319]. In addition, several studies demonstrate a suppressive role for miR-101 in tumor xenografts [318, 320, 321].

Additionally, in the context of hematopoietic malignancies, miR-101 was found extremely down-regulated in samples of Burkitt lymphoma (BL) patients [322]. In pediatric B-ALL bone-marrow samples, miR-101 was found be down-regulated [255, 323] and to target the anti-apoptotic factor heat shock protein p23 [323]. Deleterious mutations in PRC2 genes (namely in *EZH2*) were found in T-ALL patients, pointing to a tumor suppressor role of PRC2 in T-cell transformation. Nevertheless, these mutations were found mainly associated with the immature type of T-ALL but not with more differentiated cases [32]. Besides, another study showed that adult T-cell leukemia/lymphoma patient cells have increased expression of EZH2 that is inversely correlated with the expression of miR-101 [324]. Hence there is still room for a possible tumor suppressive role of miR-101 in ALL, independently on EZH2 targeting. In our work, we showed that TAL1 is a direct target of miR-101 in T-ALL cell lines. MiR-101 targets TAL1 through a 8mer binding site in its 3'UTR (Figure 3.1.4). The miR-101 ectopic expression led to a 25-50% TAL1 protein decrease and the endogenous miR-101 inhibition led to a 20-40% TAL1 protein rescue in T-ALL cell lines. These results are in accordance with the studies indicating that miR-101 has a tumor suppressor role in hematological malignancies. These are also in agreement with the decreased miR-101 expression that we found in T-ALL patients compared with normal bone marrow samples and mature thymocytes (Figure 3.1.9). Moreover, TAL1 negative cell lines express higher levels of miR-101 than TAL1+ ones (Figure 3.1.8). Given that our results from the chapter 3.2 did not recognize miR-101 as a possible transcriptional TAL1 target, this difference in the cell lines favors our hypothesis that TAL1 ectopic expression could be the consequence of abnormal down-regulation of miRNAs. We can therefore foresee a tumor-suppressive role for miR-101 in T-cell transformation, which might be in part mediated through down-regulation of TAL1.

The miR-140-5p has also been associated with suppression of tumorigenesis. In osteosarcoma and colon cancer cells the miR-140-5p mediates chemo-resistance through the suppression of HDAC4 and consequent reduction of cell proliferation [325]. Moreover, miR-140-5p was found down-regulated in breast tumors [326], hepatocellular carcinoma [327] and lung cancer [328]. Loss of miR-140 is a hallmark of ductal carcinoma in situ (DCIS) lesions and miR-140 is significantly down-regulated in cancer stem cells compared with normal stem cells [329]. The direct targets of this miRNA include stem cell self-renewal regulator SOX2 in breast cancer [326], TGFBR1 and FGF9 in hepatocellular carcinoma [327], SOX9 and ALDH1 in DCIS [329] and IGF1R in lung cancer [328]. Ectopic expression of miR-140 reduced tumor growth *in vivo* [329] and metastasis formation [328]. In our work, we showed that TAL1 is a direct target of miR-140-5p in T-ALL cell lines. The miR-140-5p targets TAL1 through three predicted binding sites in the 3'UTR. The most upstream 7mer binding site is responsible for 90% of the effect on the reporter expression (Figure 3.1.4 and 5) and the mutation of the three MREs fully restores the reporter expression. The other miRNA expressed from the same precursor, miR-140-3p, has a minor effect on TAL1 targeting (Figure 3.1.5), demonstrating that miR-140-5p is the mature form that targets TAL1 post-transcriptionally. In concordance, the miR-140-5p ectopic expression led to a 25-60% TAL1 protein decrease in T-ALL cell lines. Moreover, we observed a decrease in miR-140-5p expression in T-ALL patients compared with normal bone marrow samples and thymocytes (Figure 3.1.9). These results point for a possible tumor suppressive role of miRNA-140-5p also in T-ALL, which might be in part mediated through down-regulation of TAL1.

Recently, miR-520-5p was the first single miRNA shown to be capable of converting malignant or immortalized hepatoma cells to benign or normal cells [330]. Hepatoma cells over-expressing miR-520d-5p expressed Oct4 and Nanog, showed p53 up-regulation and hTERT down-regulation, and lost their migration abilities [330]. No direct target genes were clearly identified, but regulation of malignancy by miR-520d-5p appears to be through the conversion of cancer cells to normal stem cells by maintenance of p53 up-regulation [330]. MiR-520d-5p has four putative binding sites in the 3'UTR of TAL1 and it is the miRNA that most strongly inhibits luciferase expression (Figure 3.1.3), which we anticipated to be a

consequence of the several predicted binging sites. Nevertheless, mutation of three of the four MREs did not fully restore the inhibition mediated by this miRNA. Neither the mutation of a miR-520d-3p MRE in TAL 3'UTR can fully explain the inhibitory effect of miR-520d (Figure 3.1.5). We were not able to mutate a fourth miR-520d-5p MRE in TAL1 3'UTR (Mut III in Figure 3.1.4b). We cannot rule out the possibility that this last unmutated MRE would be responsible for the full recovery of the luciferase expression. The more updated versions of the prediction databases used in this study actually predict a fifth binding site (data not shown) for miR-520d-5p, closely placed to the most upstream. This fifth MRE could be the one responsible for the majority of the inhibition conducted by this miRNA, or other until now non-predicted binding sites can exist in the 3'UTR of TAL1 that could justify the effect. MiR-520d belongs to a family of poorly conserved miRNAs in vertebrates which could explain the inaccuracy of the actual prediction algorithms. The hypothesis of the luciferase inhibition effect being caused by the binding of the miR-520 to sequences in the reporter vector other than TAL1 3'UTR can be ruled out by the effect of miR-520d ectopic expression in T-ALL cell lines (Figure 3.1.6). In fact, not only ectopic expression of miR-520d can decrease endogenous TAL1 protein in 25-50% but also inhibition of endogenous miR-520d-5p expression can rescue TAL1 protein by 40-90%. This leads us to believe that the effect of this miRNA on TAL1 post-transcriptional regulation is specific. We therefore speculate that either we could not find the MRE in the TAL1 3'UTR or that the effect of this miRNA is not direct. Nevertheless, TAL1 negative cell lines express higher levels of miR-520d-5p, than TAL1+ ones (Figure 3.1.8) favoring our hypothesis that TAL1 ectopic expression is, in some cases, consequence of abnormal down-regulation of miR-520d-5p. This miRNA is expressed at very low levels in thymocytes and leukemic cells, making its detection difficult (data not shown). Despite the evidences that this miRNA can biologically inhibit TAL1, we could not find clear indications for a possible tumor suppressive role in T-ALL.

In the case of miR-485-5p, no validated target genes have been so far associated to cancer development. Nevertheless, down-regulation of miR-485-5p was observed in ovarian epithelial tumors, which significantly correlated with clinical variables and histological subtypes, hinting for a potential importance of this microRNA as diagnostic biomarker [331].

Moreover, the over-expression of miR-485 in breast carcinoma cells resulted in a significant decrease in cell growth, cell colony formation, and cell migration [332]. Furthermore, it was found that the presence of a SNP in the miR-485-5p binding site in the 3'UTR of the HPGD gene is associated with breast cancer risk [333]. In respect to miR-448, it is the most down-regulated microRNA in breast cancer following epithelial-mesenchymal transition induced by chemotherapy. In this context, release of the repression of SATB1 by miR-448 leads to NF-κB activation. On the other hand, NF-κB binds directly to the promoter of miR-448 suppressing its transcription, suggesting a positive feedback loop between NF-κB and miR-448 [334]. Despite the lack of knowledge on miR-485-5p and miR-448 functions, the existing studies suggest it may have tumor suppressive functions.

Each of these microRNAs has a putative binding site in TAL1 3'UTR, not yet validated. Due to lack of time, we were unfortunately not able to test their validity by mutagenesis. Nevertheless, we showed that ectopic expression of miR-485 and miR-448 in PF-382 cell line can decrease TAL1 protein in 50 and 40%, respectively. Moreover, the expression of these miRNAs is decreased in T-ALL patients compared to normal bone-marrow cells (Figure 3.1.9). Regarding TAL1 targeting, the results are promising but need further validation. Therefore, a putative tumor suppressive role for miR-485 and miR-448 in T-ALL warrants further confirmation.

The mild effects observed in protein expression mediated by the miRNAs studied are in accordance to what is described for the microRNA post-transcriptional regulation [192, 193]. Individual miRNA targeting sites usually reduce protein output by less than a half and often by less than a third. Moreover, the post-transcriptional mechanism for TAL1 regulation (either mRNA stability regulation or translational impairment) might be not only microRNA dependent, but especially cell line dependent. In fact, we only observed a marked down-regulation of the TAL1 transcript in PF-382 cells, with the exception of miR-485-5p (Figure 3.1.6), despite the effect on the protein levels in other cell lines. Moreover, upon miR-101 and miR-520d-5p inhibition no clear effects in the TAL1 mRNA levels are observed in CCRF-CEM cells. So, the mechanism of microRNA-mediated TAL1 targeting remains to be determined, but it appears to be mainly through translational impairment. The regulation

mediated by translational repression is a very attractive biological strategy since it allows potential translational reactivation of the repressed mRNAs, instead of irreversibly remove them through mRNA turnover.

Given the mild effects of the miRNAs on TAL1 protein expression, we do not foresee that their deregulated expression can fully justify the high levels of *TAL1* ectopic expression observed in T-ALL patients. Nevertheless, given the strong evidences presented for the TAL1 post-transcriptional regulation by microRNAs we cannot help speculating that TAL1 ectopic expression in T-ALL could, in some cases, be amplified by abnormal down-regulation of miRNAs targeting. Thus our results led us to speculate a model whereby microRNA could be part of TAL1 inhibition during normal development. In early stem cell progenitor cells TAL1 is normally expressed. Upon commitment to specific hematopoietic lineages the regulation of *TAL1* expression involves enhancer/promoter interactions, epigenetic alterations and trans-acting mechanisms. This complex regulation should result in *TAL1* silencing in the lymphoid lineage. A possibility is that a modest mRNA destabilization perpetuated by the miRNAs quickly yields substantial repression of protein output after transcription of the mRNA ceases. A similar mode of action has been described for other miRNAs in other physiological conditions [194]. In this model, depending on the threshold level for protein function, the mRNA decay rate, and the protein decay rate, modest miRNA-mediated repression can lead to substantially reduced protein and a much more rapid transition to the off state. If the miRNA also mediates translational repression, the transition to the off state is further accelerated [194]. If we then consider that in a pre-leukemogenic stage the normal mechanisms that lead to *TAL1* transcriptional silencing are decreased/disrupted, a coincident miRNA-TAL1 interaction decrease would augment/facilitate the up-regulation of TAL1 protein. This is our proposed mechanism for miRNA-TAL1 contribution for the disease development, which needs further confirmation. We propose as next steps to show that TAL1 regulation by miRNA indirectly affects downstream targets of TAL1 and suppresses TAL1 phenotypic effects in T-ALL. We also would like to investigate the upstream regulators of these microRNAs not only in T-ALL but in the normal context and clarify the mechanisms responsible for the abnormal microRNA down-regulation in T-ALL.

MicroRNAs transcriptionally regulated by TAL1

TAL1 appears to be on the top of a transcriptional network that, in transformed thymocytes, drives the expression of genes involved in abnormal proliferation, differentiation and survival. Yet, the pathways downstream of TAL1 that contribute to leukemia development are still poorly characterized. Having that in mind, miRNA genes are attractive candidates to fulfill the role of TAL1 targets with important consequences for leukemogenesis.

LMO2, a common factor in the transcriptional complex involving TAL1, can negatively regulate the *in vivo* expression of miR-142, which is known to promote T-cell development [335]. In the erythroid cell context, LMO2 was shown to be negatively regulated by miR-223, and this consequently releases miR-142 from LMO2 inhibition, demonstrating a regulatory pathway with pivotal functions in differentiation [336]. These discoveries support the unexplored possibility that TAL1-mediated transcriptional regulation of miRNAs may occur in T-ALL and have functional impact on the pathophysiology of T-ALL. Moreover, an explanation is still lacking for the majority of aberrantly expressed miRNAs in T-ALL [274], and miRNA regulatory networks regulated by T-ALL oncogenes have not been so far extensively studied. Therefore we speculated that TAL1-mediated transcriptional regulation of miRNAs may occur in T-ALL and have functional impact on the pathophysiology of this malignancy. Through characterization of a TAL1-dependent microRNA gene expression profile, we identified eight miRNAs whose expression changed significantly upon TAL1 over-expression in a T-ALL cell line (Figure 3.2.1).

Most studies performed so far to explore the transcriptional network of coding genes downstream of TAL1 usually characterize the expression profile upon knockdown of TAL1 [124, 143, 144]. We know from the literature [124, 143] and our own experience (data not shown) that the knockdown of TAL1 has a profound effect in the viability of the cells, with more than 50% decrease on the viability of the T-ALL cells in the first 48h. We reasoned that a cell line stably over-expressing TAL1 could give some false positive results as consequence

of the elevated TAL1 expression, but it would prevent us from finding altered miRNAs related to effects in apoptosis rather than to TAL1 transcription.

Importantly, it is known that TAL1 can act both as a transcriptional activator and inhibitor, depending on the cellular context and target gene. Therefore, we expected that miRNA genes would be both up- and down-regulated by TAL1 and that is exactly what we have found. Following validation of the microRNA screen profile results upon enforced or silenced expression of TAL1 in T-ALL cell lines, we confirmed that miR-135a, miR-223 and miR-330-3p were up-regulated by TAL1, whereas miR-146b-5p and miR-545 were down-regulated by TAL1 (Figure 3.2.1).

We considered stringent criteria to validate the miRNA screen profile results by enforcing or silencing expression of TAL1 in several T-ALL cell lines. In fact, we accepted as validated only miRNAs that were regulated in the predicted manner by both over-expression and silencing of TAL1 in at least one cell line. Actually most of them were validated in two different T-ALL cell lines (Figure 3.2.1). Given this stringent criteria, we confidently consider that miR-135a, miR-223, miR-330-3p, miR-146b-5p and miR-545 expression is modulated by TAL1. The next obvious step was to determine if this control is mediated by a direct regulation of these miRNA transcriptions by TAL1 in T-ALL cells. Given that the promoter region for the majority of miRNA genes is not yet defined, we examined ChiP-seq data for the presence of TAL1 binding peaks up to 10kb upstream of the transcription start site (TSS) of each microRNA gene. We were able to verify by TAL1 ChIP-qPCR in T-ALL cells that TAL1 binds to a genomic area 3.5kbs upstream of the miR-223 TSS, which indicates that miR-223 is a direct target of TAL1 in T-ALL (Figure 3.2.1g and 3.2.2). We did not find evidence, from the available TAL1 ChIP-seq data, for direct binding of TAL1 to the remaining microRNA genes, suggesting that miR-135a, miR-330-3p and miR-545 might be indirectly regulated by TAL1, at least in the T-ALL cells analyzed. It should be noted, however, that some TAL1 binding sites differ between cell lines and between cell lines and primary samples [144] and thus we cannot exclude the possibility that TAL1 may regulate directly these miRs in other T-ALL cells. Interestingly, TAL1 appears to bind upstream of miR-223 to a previously described region containing a conserved proximal genomic element with binding sites for the transcription

factor C/EBPα and PU.1 [337]. Furthermore, miR-223 appears to follow the same pattern of expression along normal human thymocyte development as TAL1 [94], with high levels in CD34+ T-cell precursors and sharp down-regulation in more differentiated subsets of CD4+CD8+ thymocytes (Figure 3.2.4a). This is also in agreement with the hypothesis that TAL1 positively regulates the expression of this microRNA with oncogenic functions.

Can miRNAs regulated by TAL1 be part of an oncogenic transcriptional network downstream of TAL1?

The miRNA-223 is a highly conserved miRNA preferentially expressed in the hematopoietic system [213]. Specifically, miR-223 is a myeloid-specific [213, 337] microRNA essential for normal neutrophil maturation and for granulocyte differentiation. This miRNA negatively regulates progenitor proliferation via MEF2C down-regulation [338]. In agreement, miR-223 functions as a tumor suppressor in acute myeloid leukemia [339], and appears to be repressed in chronic myeloid leukemia, allowing for the expression of MEF2C [340]. Also in agreement, miR-223 expression in HSC progenitors is elevated, but it sharply decreases during lymphocyte lineage development (Figure 3.2.4). In addition to proper granulocyte differentiation and function, miR-223 is also essential for proper erythroid development, likely in part by directly targeting LMO2 [336]. Importantly, TAL1 is essential for maturation of erythroid and megakaryocytic precursors [127]. Loss of *Tal1* in the embryo results in reduced capacity towards myeloid differentiation [126] and decreased expression of *TAL1* in human CD34+ cells compromised the erythroid and myeloid lineage [134]. Given the importance of TAL1 and miR-223 in the myeloid compartment, a possible regulation of miR-223 by TAL1 in the normal myeloid lineage commitment, that was never addressed, would certainly be an interesting subject for future studies.

Much of the research on miR-223 has focused on its role in myeloid cell differentiation but interestingly, over-expression of miR-223 in mouse hematopoietic progenitor cells leads to a specific increase in the T-lymphoid lineage cells [213]. Moreover,

miR-223 is one of the most expressed microRNAs in T-ALL [341] and we verified that it is highly over-expressed compared to normal precursors (Figure 3.2.3c). Also, a subset of adult T-ALL has been identified that is characterized by myeloid-like genetic features and miR-223 over-expression [342]. In childhood T-ALL, miR-223 was shown to cooperate with Notch1 to accelerate the onset of disease in a Notch-induced leukemia mouse model. This effect was proposed to be due, at least in part, to inhibition of FBXW7, a negative regulator of NOTCH signaling [341]. The ubiquitin ligase FBXW7 has already been identified previously as target of miR-223 in MEFs, where it regulates the cell cycle by inducing cyclin E ubiquitination and subsequent degradation [343].

Interestingly, analysis of microRNA gene expression profiles in different T-ALL subsets [261] revealed that primary samples which frequently express high TAL1 levels display higher levels of miR-223 (Figure 3.3.3). This observation not only strengthened our results showing that TAL1 transcriptionally activates miR-223, but also suggests that the oncogenic function of this microRNA may extend beyond mere collaboration in Notch-induced leukemia. In fact, FBXW7 targets to degradation other oncogenic proteins, such as c-MYC, MYB, cyclin E, mTOR, HIF-1α and MCL-1 [343-348]. Importantly, FBXW7 is recognized as an important tumor-suppressor in T-ALL, with 20% of patients carrying loss-of-function mutations [72, 73] Furthermore, *Fbxw7* knockout mice developed thymic lymphoma, with c-Myc accumulation and increased proliferation of immature T-cells, but also with p53-dependent cell-cycle arrest and apoptosis in mature T-cells [80]. Furthermore, the oncogenic transcription factor MYB, which is known to be involved in malignant hematopoiesis [42], was recently identified as a direct target of TAL1 [144]. Many of the downstream targets of MYB are also controlled by TAL1. Therefore, these two transcription factors activate an overlapping oncogenic network to work in concert to maintain a deregulated gene expression program in T-ALL [144]. Given that FBXW7 targets MYB to degradation and that TAL1 might down-regulate FBXW7 through miR-223, it would be interesting to explore the possibility of TAL1 regulating MYB expression levels not only transcriptionally [144] but also at the level of protein stability through a miR-223/FBXW7 axis. Of note, while our manuscript was under review another study was published that corroborated our results regarding miR-223 regulation by TAL1 [300]. In this

work, the link between TAL1 and FBXW7 down-regulation mediated by miR-223 was firmly established as a mechanism leading to down-regulation of FBXW7 in the majority of T-ALL cases that lack gene-specific FBXW7 inactivating mutations or deletions [300].

The oncogenic role of miR-223 in T-ALL might not be restricted to FBXW7. The abnormal expression of miR-223 in thymocytes can aberrantly decrease the expression of other key proteins that are not normally regulated by this miRNA in this hematopoietic lineage. Moreover, the pro-leukemic role of miR-223 may be also achieved by down-regulating targets such as E2F1, FOXO1, RHOB or EPB41L3, which have been associated with induction of apoptosis and/or have tumor suppressive roles (Table 3.2.2). Interestingly, the possibility that miR-223 may act downstream of TAL1 to negatively regulate MEF2C, recently identified as an oncogene in T-ALL [349], is in line with the fact that TAL1 and MEF2C tend to segregate, defining two discrete T-ALL subsets [349].

The fact that TAL1 positively regulates miR-223, which in turn down- regulates an important tumor suppressor such as FBXW7 [261] and possibly other anti-tumoral genes, indicates that TAL1 is at the leading edge of a transcriptional network involving stabilization of important proteins with oncogenic functions in T-ALL. These findings enabled answering our initial question regarding the knowledge of miRNAs regulated by TAL1 and their part in an oncogenic transcriptional network downstream of TAL1 in T-ALL. TAL1 effects appear to be mediated in part by miR-223, but our studies suggest that other microRNAs may have a similar role as effectors of TAL1-mediated T-cell oncogenesis (Figure 3.2.5).

Interestingly, miR-223 may be transcriptionally regulated by other oncogenes than TAL1 in T-ALL. Recently, miR-223 was shown to be a direct target of NOTCH in T-ALL cells [270]. The binding of NOTCH to miR-223 promoter region requires NF-kB activation. NOTCH1, NOTCH3 and p65 are directly recruited to the promoter region of miR-223 and, as TAL1, also bind to the previously described C/EBPα and PU.1 binding region. Contrarily to our results (Figure 3.3.3), in T-ALL patient samples the authors were unable to show a direct correlation between miR-223 expression levels and the up-regulation of NOTCH or NF-kB [270]. This result suggests that miR-223 expression may be maintained by other pathways aberrantly activated in T-ALL, as we have shown by the activation of miR-223 by TAL1. On the other

hand, the fact that T-ALL blasts consistently express high levels of miR-223 irrespectively of TAL1 expression [261] tells us that yet other mechanism of miR-223 regulation could be altered in T-ALL.

One miRNA may target several components of a single signaling pathway but also several miRNAs can regulate in additive or redundant ways several genes in the same functional pathway. To explore the possibility of networking coordination by the microRNAs altered by TAL1 over-expression we collected information about the validated human target genes for each microRNA. In order to determine the functional relevance of the sets of miRNA targets identified, we tested whether they were enriched for specific biological functions or pathways. Gene set enrichment analysis showed an enrichment in biological processes related to inflammation (e.g., NF-kB signaling pathway and IL1/IL1R signaling pathway) and cancer (e.g. pathways in cancer) (Table 3.2.2 and Supplementary Table 3).

There is only one report associating miR-330-3p expression with cancer development, where miR-330-3p expression was associated with development of brain metastases in patients with NSCLC [350]. Interestingly, miR-330-3p reportedly targets E2F1 and CDC42, both of which are described to promote apoptosis [351, 352]. This is in line with a possible pro-oncogenic role for this miRNA up-regulated by TAL1. Moreover, the pro-leukemic role of miR-223 may also be achieved by down-regulating E2F1. Therefore, the cell cycle regulator E2F1 is a putative target of two miRNAs up-regulated by TAL1 (miR-223 and miR-330-3p). Reduced E2F1 has been associated to enhanced survival of leukemic blasts [265] and development of hematopoietic malignancy [229]. E2F1 regulation can be mediated by miRNAs in T-ALL, namely by miR-17-5p and miR-20a that are components of the miR-17-92 oncogenic polycistron [229]. Here, we uncover the possibility that networking coordination by the microRNAs up-regulated by TAL1 might mediate E2F1 down-regulation and enhance survival of leukemic cells. Since T-ALL cells tend to display several complementary cell cycle promoting lesions (e.g. p16 inactivation and PI3K signaling pathway activation) down regulation of E2F1 may essentially prevent cell death without significantly compromising cell division.

We hypothesized that if miRNAs regulated by TAL1 have an impact in TAL1 promotion of leukemogenesis, not only their targets genes should reinforce TAL1 pro-leukemogenic role, as some of the miRNAs targets could be also regulated by TAL1. In this manner, TAL1 could reinforce its oncogenic role through the indirect modulation of certain genes via regulation of specific miRNA-target axes. We have already discussed above this possibility for the case of MYB and the miR-223/FBXW7 axis, and we wondered if this could be extended to other miRNAs regulated by TAL1 and other miRNA-downstream targets. Therefore, we compiled a list of high confidence TAL1 positively or negatively regulated genes from publicly available data [143, 295]. We then intersected the predicted targets of TAL1 down-regulated microRNAs with the protein-coding gene targets previously demonstrated to be positively regulated by TAL1, and vice versa, and searched for common hits in both lists.

The data from this cross-examination of congruent TAL1 regulated protein-coding and miRNA target genes, are in line with the notion that the latter could be part of downstream networks collaborating in TAL1-mediated leukemogenesis (Figure 3.2.5). Indeed, most TAL1 up-regulated genes that have 3'UTRs predicted as targets for the TAL1 down-regulated miR-146b-5p and miR-545 have a known or putative oncogenic function (Figure 5a). For example, CD53 was shown to protect JURKAT cells from apoptosis [353], PDE3B appears to be involved in glucocorticoid resistance in CEM cells [354], ETS-1 participates in the T-cell maturation arrest mediated by TLX genes in T-ALL [355], and MYB is a bona-fide oncogene in T-ALL [295] (Supplementary Table 4).

There is only one report associating miR-545 to cancer where this miRNA was found to be less abundant in cancerous lung tissues and associated with tumor suppressive functions [356]. In fact, ectopic miR-545 expression suppressed cell proliferation by cell cycle arrest at the G0/G1 phase and induced cell apoptosis in lung cancer cells by targeting cyclin D1 and CDK4 genes [356]. We predict a tumor-suppressive function for miR-545 in T-ALL, given that this gene is down-regulated by TAL1 (Figure 3.2.1). One predicted target of miR-545 is the oncogene MYB. Once more, it would be interesting to explore the possibility of TAL1 regulating MYB expression levels not only transcriptionally [144] but also post-transcriptionally through the down-regulation of miR-545. Moreover, miR-545 might also

exert a tumor suppressive role targeting cyclin D1 and CDK4 genes. Therefore, another hypothesis is that by down-regulating this miRNA in T-ALL, TAL1 might be promoting cell cycle progression through up-regulation of cyclin D:CDK4 complexes. Also of note, three of the four genes (*KRT1*, *Rapgef5*, *JAZF1*) with predicted 3'UTR seed sequences for both miR-146b-5p and miR-545 are associated with pro-tumoral functions (Supplementary Table 4).

In sharp contrast, the TAL1 down-regulated genes that are predicted targets for miR-135a, miR-223 and miR-330-3p display a clear abundance in (putative) tumor suppressors or in genes whose functions are compatible with anti-tumoral effects (Figure 5b and Supplementary Table 5). This is evident, for instance, in the case of the four genes potentially regulated by two of these microRNAs, in which only one has an oncogenic role (*IGF1R*) and three likely have tumor suppressive functions (*SRGAP3*, *TOX*, *LRP12*).

There is a current challenge to reveal the mechanisms of regulation of miRNA gene expression, in order to understand why and how the miRNAs are deregulated during the development of miRNA-associated cancer. In this thesis we identified and validated for the first time a small set of non-protein coding TAL1 target genes, implicating microRNA genes as part of the transcriptional network downstream of TAL1 whose role may be important in the context of hematopoiesis and T-cell leukemogenesis. With the exception of miR-223, which has been intensively studied in the context of T-ALL, we still lack functional information for the other microRNAs identified as being regulated by TAL1. Their role in T-cell malignancy should be detailed by future investigations, for which our work serves as a starting point.

A possible tumor suppressor role for miR-146b in T-ALL

Having shown that microRNA genes are regulated by TAL1 at the level of transcription, we hypothesized that some of these miRNAs can be part of the oncogenic network triggered by TAL1 ectopic expression in leukemia, as discussed. We focused our efforts on miR-146b-5p and evaluated the functional and molecular effects of its deregulation by TAL1 in the context of T-ALL.

When analyzing our own data or data from an independent set of human T-ALL cell lines (data not shown) we did not find significant difference in miR-146b-5p levels between TAL1+ and TAL1- cell lines. Nevertheless this microRNA is one of the most down-regulated miRNAs upon TAL1 over-expression (Figure 3.2.1). Moreover, similar results were obtained when assessing miR-223 expression in T-ALL cell lines (data not shown) and, nonetheless, we and others [300] have convincingly shown that miR-223 is directly regulated by TAL1. Maybe a larger set of T-ALL cell lines would provide data more in accordance to what is obtained when analyzing primary T-ALL samples. In fact, analysis of the different T-ALL subsets of two independent cohorts revealed that primary samples of the TAL1 subgroup (integrating Sil-Tal1+ and LMO+ cases [261] or SIL-TAL, TCR-TAL and other TAL1+ cases [305]) express lower levels of miR-146b-5p than other T-ALL cases (Figure 3.2.3 and Figure 3.3.1a), further confirming our screen results. Therefore, we hypothesize that TAL1 negatively regulates miR-146b-5p, which may have a tumor suppressive role in the T-ALL context.

T-ALL patients express significantly lower levels of miR-146b-5p than normal controls (Figure 3.3.1), and these results were obtained for two independent cohorts of patients, irrespectively of TAL1 expression. T-ALL cell lines also have decreased miR-146b-5p expression (Figure 3.3.1a). This suggested that reduced expression of miR-146b-5p might be an important factor in T-ALL that might extend beyond TAL1 regulation.

Through the ChIP-PCR experiments we did not find a putative TAL1 binding site upstream of the miR-146b locus, although we evaluated the putative binding region depicted in Figure 3.2.2a. A negative result in ChIP-PCR experiments does not necessarily mean this microRNA is not a direct target of TAL1. It might instead just indicate that we are not scrutinizing the correct genomic area for TAL1 binding. Therefore, we assessed the levels of the primary transcript of miR-146b in T-ALL cell lines. We reasoned that if miR-146b is a direct transcriptional target of TAL1, the levels of its primary transcript should correlate with TAL1 status. We verify that TAL1 negative cell lines tend to express more pri-miR-146b than TAL1 positive ones (Figure 3.3.2a). Moreover, the knockdown of *TAL1* in a T-ALL cell line results in marked up-regulation of the primary transcript (Figure 3.3.2b), anticipating the up-regulation of the mature miRNA expression upon TAL1 knockdown (Figure 3.2.1). Although

merely correlative, these results argument in favor of a direct regulation of miR-146b transcription by TAL1.

Furthermore, miR-146b-5p is clearly up-regulated in the double-positive to single-positive transition (Figure 3.2.4c). This miRNA was shown to be amongst the five most up-regulated microRNAs in mature single-positive thymocytes and amongst the five with the highest change in the transition from the DP to SP stage of T-cell development [219]. The fact that the levels of miR-146b-5p associate with thymocyte maturation (Figure 3.2.4c) is in agreement with a model whereby TAL1 over-expression during malignant transformation inhibits miR-146b-5p and promotes T-cell developmental arrest.

Unexpectedly, modulation of miR-146b-5p levels did not affect T-ALL cell viability (Figure 3.3.3). These results were unforeseen since TAL1 expression in mouse or human leukemic cells is associated with increased cell proliferation and resistance to apoptosis [114, 143]. Therefore, if as we postulated, mir-146b-5p acts downstream of TAL1 as a tumor suppressor, its overexpression should impact on critical biological processes such as cell proliferation or viability. Given that miRNA-146b-5p was shown to be highly up-regulated during the later stages of thymocyte maturation [219], we reasoned that modulation of its expression could have an effect on T-ALL cell differentiation. Therefore, we monitored the immunophenotype of the cells stably over-expressing or down-regulating this miRNA during several weeks after their establishment. Once more, no obvious change in the phenotype of any cell line was verified upon miR-146b modulation, irrespectively of the developmental stage of the cells (Supplementary Figure 6).

In contrast, we found that miR-146b-5p modulation alters the migration and invasion capacities of the T-ALL cells *in vitro* (Figure 3.3.4 and supplementary Figure 7). In fact, we have shown that down-modulation of miR-146b-5p increases the number of migrating T-ALL cells in response to serum presence in the medium (Figure 3.3.4a and 4c and supplementary Figure 8). Moreover, we showed that by decreasing miR-146b-5p levels T-ALL cells have enhanced capacity to invade a matrigel matrix (Figure 3.3.4b and 4d and supplementary Figure 8). These results are in agreement with reports linking miR-146b expression to

malignancy, specifically linking its down-modulation to enhanced migration of cancer cells in solid tumors [297, 303-305].

We then analyzed the role of miR-146b in the development of tumors *in vivo*. To that end, we transplanted T-ALL cells over-expressing miR-146b into immunodeficient mice. Since this miRNA modulation counteracts the effects of TAL1, this model allows us to address the role of this microRNA *in vivo* but also its therapeutic potential. In fact, we were able to show that over-expression of miR-146b in human T-ALL cells significantly increased the survival of leukemic mice compared to mock transduced T-ALL cells transplanted mice (Figure 3.3.5a). Given the role of this microRNA in T-ALL cells migration *in vitro*, these results suggest a delayed development of the disease. In agreement with this hypothesis, we observed a lower percentage of leukemic cells in the peripheral blood of the miR-146b mice at the fourth week after transplantation, when practically all mice were still alive (9/10) (Figure 3.3.5b and Supplementary Figure 8a). Therefore, these results are consistent with the hypothesized tumor suppressor role of miR-146b in T-ALL.

Given our *in vitro* results in migration and invasion and the increased survival observed in our *in vivo* experiments, we were expecting to observe differences in leukemia cells infiltration of hematopoietic or non-hematopoietic organs. Nevertheless, we did not find those differences. However, since this analysis was done at the time of death, and the mice were sacrificed only when presenting clear signs of disease, we speculate that those differences were already lost at that point. Therefore, given the aggressiveness of the human T-ALL cells, a strategy to evaluate differential organ invasion at early time points after injection would probably be more informative. In those conditions we might speculate that the mock transduced cells would invade earlier the organs affected. Furthermore, histopathology analysis (Supplementary Figure 8b) revealed a slight decrease in tumor load in the sections involving the CNS (Bregma, olfactory bulb and cerebellum) and the spinal cord in the mice injected with cells over-expressing miR-146b. The significance of these results requires obviously further confirmation. In this regard, it would be very relevant to explore the possibility of using increment of miR-146b levels in T-ALL cells as adjuvant therapy to prevent CNS involvement, which is one factor of poor prognosis in this malignancy

Interestingly, a potential therapeutic use of miR-146b was investigated by intra-tumor injection of exosomes derived from miR-146b-expressing mesenchimal marrow stromal cell. This strategy was able to significantly reduce glioma xenograft growth in a rat model of primary brain tumor [357] paving the way for the therapeutic use of this microRNA for cancer treatment.

In the future we would like to reveal the miR-146b-5p downstream targets that contribute to the enhanced migration capacities phenotype in T-ALL cells upon this miRNA repression. One obvious strategy would be to perform transcriptome analysis upon modulation of the miRNA in the T-ALL cells, followed by identification of the genes that are down-modulated by miR-146b-5p. The intersection of these with a list of miR-146b-5p predicted targets might help to narrow down the candidates. A gene ontology analysis of miR-146b-5p predicted targets genes (by TargetScan, data not shown) indicate that the most enriched biological process is transcription, followed by signaling transduction. Nevertheless, a manual search through the predicted targets foresee several genes with putative functions associated with migration and motility, namely *RAB10, RAB8B, IQGAP3, IQGAP1, ANKRD28, MARK1, EFNB2, ROBO1, TLN2, NRP2, FLNA, MYO6, RND2* and *CLASP2*. Interestingly, the validated targets for the TAL1 down-regulated gene miR-146b-5p include IRAK1, TRAF6 and NFKB1 (all of which are involved in chronic inflammation), as well as the oncogene KIT (Table 3.2.2, Supplementary Table 3 and Supplementary Figure 3), thus favoring a possible tumor suppressive role for this miRNA.

Interestingly miR-146b-5p has been implicated in a mechanism linking chronic inflammation and cancer promotion also with tumor-suppressor function. In cancer cells, the STAT3-induction of miR-146b is disrupted, often through methylation of the miR-146b promoter, leading to the release of miR-146b target, NF-κB. NF-κB activation leads to production of IL-6, which activates STAT3, leading to enhanced migration and invasion in breast cancer cells [358]. Several other studies have implicated miR-146b-5p as a tumor suppressor whose decreased expression promotes directly or indirectly enhanced migration properties of cancer cells. For instance, in osteosarcoma cells some tumor suppressor functions of miR-146b-5p are mediated through the direct targeting of the mRNA-binding

protein AUF1, leading to suppressed invasion/migration and proliferation abilities [359]. Over-expression of miR-146b into breast cancer cells also suppressed experimental lung metastasis [305]. Moreover, miR-146b-5p reduces glioma cell migration and invasion by direct targeting MMP16 [360] and EGFR [303]. MiR-146b-5p mediated targeting of MMP16 and inhibitory effects on cell migration and invasion was also observed in pancreatic cancer [304].

More importantly, a tumor suppressor function for miR-146b was recognized in the context of PTEN-deficient T-lymphomagenesis in mice [361]. This miRNA is up-regulated in the thymus of PTEN-deficient mice as part of the cellular response against transformation and when over-expressed can effectively inhibit progression to malignancy. Tumor suppression was mediated by miR-146b attenuation of TCR signaling through repression of its direct target TRAF6, an important activator of NF-κB, inhibiting downstream NF-κB-dependent induction of c-Myc [361]. Also regarding hematological malignancies, over-expression of miR-146b-5p inhibited diffuse large B-cell lymphoma (DLBCL) cell proliferation and miR-146b-5p lower expression in DLBCL patients was associated with poor prognosis [362].

Overall, our results point to a direct down-regulation of miR-146b-5p by TAL1, microRNA with a very probable tumor suppressive role in T-cell acute lymphoblastic leukemia.

Concluding Remarks

The improvement of our knowledge regarding the causes, pathophysiology and regulation of T-ALL is necessary to develop more efficient and less toxic therapeutic strategies that target the leukemia cells. A growing number of evidences place microRNAs as critical players in the complex biology of cancer cells and show that they actively contribute to T-cell acute lymphoblastic leukemia progression.

Nevertheless, despite the growing knowledge on miRNA biology, little is known about the transcriptional regulation of miRNAs and an explanation is still lacking for the majority of aberrantly expressed miRNAs in T-ALL. Moreover, miRNA regulatory networks regulated by T-ALL oncogenes have not been so far extensively studied.

In this thesis we have shown that TAL1 sits at the top of a transcriptional network that in transformed thymocytes drives the expression of microRNA genes that are involved in key biological processes for leukemia development. In particular, our results point to a direct down-regulation of miR-146b-5p by TAL1, a microRNA that demonstrated to have a tumor suppressive function in our *in vitro* and *in vivo* experiments.

Importantly, we have also shown that TAL1 is post-transcriptionally regulated by microRNAs. We therefore speculated a model whereby in a pre-leukemogenic stage the normal mechanisms that lead to *TAL1* transcriptional silencing are decreased/disrupted and a coincident decrease in miRNA-TAL1 interaction would augment the up-regulation of TAL1 protein that we find in some T-ALL cases.

In all, our findings support the existence of a cross-talk between TAL1 and microRNA genes, involving upstream epigenetic regulation of TAL1 by specific miRNAs and downstream transcriptional regulation of miRNA genes by TAL1. Our studies contribute to the understanding of the mechanisms involved in TAL1 over-expression in T-ALL and to the clarification of the tumorigenic network downstream of TAL1, namely concerning the identification of miRNA genes transcriptionally regulated by TAL1. Whether the interactions identified here may be explored therapeutically for the benefit of T-ALL patients remains a challenge for the future.

Based on the results described in this PhD thesis we propose the model depicted in Figure 4.1.

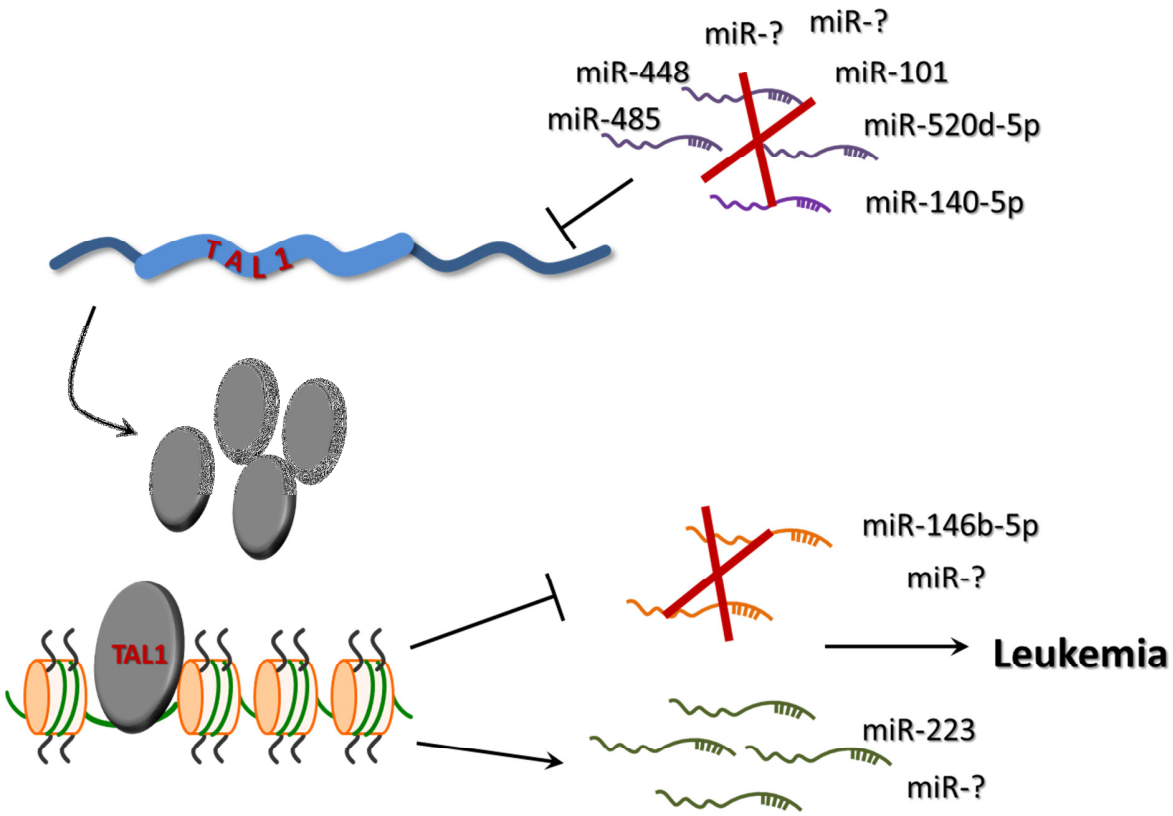


Figure 4.1 **– Dual-way talk between TAL1 and microRNA genes.**

TAL1 is post-transcriptionally regulated by the microRNAs described (miR-101, miR-140-5p, miR-520d-5p, miR-448 and miR-485) and may be by other miRNAs not yet identified (miR-?). We propose that in a pre-leukemogenic stage the normal mechanisms that lead to TAL1 transcriptional silencing are disrupted and a coincident impairment of the miRNA-TAL1 interaction could augment the up-regulation of TAL1 protein. On the other hand, ectopic expression of TAL1 in thymocytes leads to down-regulation of tumor-suppressive miRNAs (such as miR-146b-5p and other yet unidentified miRNAs – miR-?) and over-expression of oncogenic miRNAs (such as miR-223 and other yet unidentified miRNAs – miR-?). We propose that this transcriptional disruption of the normal microRNome predispose the lymphocytes to the development of leukemia through the altered expression of oncogenes and tumor-suppressor protein coding genes that are targets of the TAL1-regulated microRNAs.

SUPPLEMENTARY DATA

Supplementary Table 1 – Computational prediction of TAL1 3'UTR targeting by microRNAs.

miRNA	Program	Binding type	3'UTR position	Conservation	LMO2 targeting	DIANA-miRPath
hsa-miR-1	miRDB(MirTarget2); TargetScanS; miRanda (microrna)	8mer	2815	*		Adherent junction; glioma; thyroid cancer
hsa-miR-101	miRDB(MirTarget2); TargetScanS; miRanda (microrna); PicTar (4-way); DianaMicroT	8mer	3174	9 *		MAPK signaling pathway; Colorectal cancer; Renal cell carcinoma
hsa-miR-103	miRanda (microrna)		507	*		
hsa-miR-106a	TargetScanS; PicTar (4-way); DianaMicroT	7mer-m8	3156	8	✔	MAPK signaling pathway; Chronic myeloid leukemia; TGF-beta signaling pathway
hsa-miR-106b	TargetScanS; PicTar (4-way); DianaMicroT	7mer-m8	3156	8	✔	TGF-beta signaling pathway; Bladder cancer; MAPK signaling pathway
hsa-miR-107	miRanda (microrna)		507	*		
hsa-miR-1184	miRDB(MirTarget2)		523			
hsa-miR-124	TargetScanS	7mer-m8	545			Gap junction; Melanogenesis; Axon guidance
hsa-miR-128	TargetScanS	7mer-m8	3271, 1595			
hsa-miR-1283	StarBase					
hsa-miR-1285	miRDB(MirTarget2)		2225, 2526			

hsa-miR-1291	miRDB(MirTarget2)		556			
hsa-miR-129-5p	miRDB(MirTarget2); miRanda (microrna)		3332	*		
hsa-miR-1308	miRDB(MirTarget2)		2847, 3353			
hsa-miR-136	miRDB(MirTarget2); miRanda (microrna)		187	*		Nicotinate and nicotinamide metabolism; Butanoate metabolism; Starch and sucrose metabolism
hsa-miR-140-5p	miRDB(MirTarget2); miRanda (microrna); StarBase; DianaMicroT	8mer	1318, 1365, 2952	*		adherent junction; Pancreatic cancer; Regulation of actin cytoskeleton
hsa-miR-144	TargetScanS; PicTar (4-way); DianaMicroT	7mer-1A	3175	6		Wnt signaling pathway; Maturity onset diabetes of the young; Reductive carboxylate cycle (CO2 fixation)
hsa-miR-148a	miRanda (microrna)		279	*		
hsa-miR-17	DianaMicroT	7mer-m8	3156	8	✔	Bladder cancer; Axon guidance; Pancreatic cancer
hsa-miR-17-5p	TargetScanS; PicTar (4-way); DianaMicroT	7mer-m8			✔	
hsa-miR-186	miRDB(MirTarget2); miRanda (microrna; miRBase)		2979, 3022	*	✔	TGF-beta signaling pathway; Long-term potentiation; Wnt signaling pathway
hsa-miR-1915	miRDB(MirTarget2)		154, 647			
hsa-miR-196	miRanda (miRBase)					Cholera - Infection; ABC transporters - General; Novobiocin biosynthesis
hsa-miR-197	TargetScanS; miRDB(MirTarget2);	7mer-m8	3309	*		

	miRanda (microrna)					
hsa-miR-1972	miRDB(MirTarget2)		305, 416, 2262			
hsa-miR-199a*	PicTar (4-way)		3226			
hsa-miR-199a-5p	miRanda (microrna)		2452, 2678	*		
hsa-miR-199b-5p	miRanda (microrna)		2678	*		
hsa-miR-20	PicTar (4-way)		3207			
hsa-miR-204	PicTar (4-way); miRanda (microrna)		3409	*		Long-term potentiation; Chronic myeloid leukemia; Neurodegenerative Diseases
hsa-miR-206	TargetScanS; miRDB(MirTarget2); miRanda (microrna)	8mer	2815-2821	*		adherent junction; Focal Adhesion; Glioma
hsa-miR-20a	TargetScanS; PicTar (4-way); DianaMicroT	7mer-m8	3156	**8**	✔	TGF-beta signaling pathway; Bladder cancer; pancreatic cancer
hsa-miR-20b	TargetScanS; PicTar (4-way); DianaMicroT	7mer-m8	3156	8	✔	Bladder cancer; Axon guidance; pancreatic cancer
hsa-miR-211	PicTar (4-way); miRanda (microrna; miRBase)		3409	*		Long-term potentiation; Chronic myeloid leukemia; Neurodegenerative Diseases
hsa-miR-216a	miRanda (microrna)		2263	*		
hsa-miR-302a	TargetScanS; PicTar (4-way)	7mer-m8	3155			Chronic myeloid leukemia; TGF-beta signaling pathway; prostate cancer
hsa-miR-302b	TargetScanS; PicTar (4-way)	7mer-m8	3155			Prostate cancer; TGF-beta signaling pathway; Chronic

						myeloid leukemia
hsa-miR-302c	TargetScanS; PicTar (4-way)	7mer-m8	3155			Chronic myeloid leukemia; TGF-beta signaling pathway; Focal adhesion
hsa-miR-302d	TargetScanS; PicTar (4-way)	7mer-m8	3155			Prostate cancer; TGF-beta signaling pathway; Chronic myeloid leukemia
hsa-miR-302e	TargetScanS					Ubiquitin mediated proteolysis; Huntington's disease; Amyotrophic lateral sclerosis (ALS)
hsa-miR-31	TargetScanS		2312	*		
hsa-miR-342-3p	miRanda (microrna)					
hsa-miR-346	miRanda (microrna)		2682	*		
hsa-miR-34a	miRanda (microrna); StarBase		1355	*		
hsa-miR-34c-5p	miRanda (microrna); StarBase		1356	*		
hsa-miR-372	TargetScanS; PicTar (4-way)	7mer-m8	3155			Bladder cancer; melanoma; glioma
hsa-miR-373	TargetScanS; PicTar (4-way)	7mer-m8	3155		✔	TGF-beta signaling pathway; Axon guidance; mTOR signaling pathway
hsa-miR-374a	PicTar (4-way); miRanda (microrna)		3361	*	✔	
hsa-miR-377	DianaMicroT	various (3)	728, 2305, 3309	0		Wnt signaling pathway; Small cell lung cancer; TGF-beta signaling pathway;

hsa-miR-409-3p	miRDB(MirTarget2); DianaMicroT	7mer(2);8mer	1753, 2863, 2925	0-2-0		Colorectal cancer; TGF-beta signaling pathway; adherent junction
hsa-miR-410	miRanda (microrna)		3359	*	✔	
hsa-miR-429	miRanda (microrna)		1030	*		
hsa-miR-433	miRanda (microrna)		2094	*	✔	
hsa-miR-448	miRanda (microrna); StarBase		1138	*		
hsa-miR-449a	miRanda (microrna); StarBase		1355	*		
hsa-miR-449b	miRanda (microrna); StarBase		1355	*		
hsa-miR-485-5p	miRanda (microrna); StarBase		1074	*		
hsa-miR-486-3p	StarBase		3316	*		
hsa-miR-505	miRanda (microrna)					
hsa-miR-506	TargetScanS; DianaMicroT	7mer-m8	545	3		Axon guidance; Melanogenesis; Acute myeloid leukemia
hsa-miR-516b	DianaMicroT	7mer;8mer(2)	1376, 2271, 2556	2-1-0		Circadian rhythm; Keratin sulfate biosynthesis; Notch signaling pathway
hsa-miR-519d	TargetScanS; DianaMicroT	7mer-m8	3156	8	✔	Bladder cancer; pancreatic cancer; Axon guidance
hsa-miR-520a	TargetScanS					Pancreatic cancer; Bladder cancer; Chronic myeloid leukemia

hsa-miR-520a-3p	TargetScanS	7mer-m8	3155			
hsa-miR-520b	TargetScanS	7mer-m8	3155			Pancreatic cancer; Bladder cancer; glioma
hsa-miR-520c-3p	TargetScanS	7mer-m8	3155			Pancreatic cancer; Bladder cancer; glioma
hsa-miR-520d-5p	miRDB(MirTarget2); DianaMicroT	7mer(2);8mer; 9mer	1332, 1745, 2743 , 3278	7-4-0-1		Pancreatic cancer; Bladder cancer; TGF-beta signaling pathway
hsa-miR-520d-3p	TargetScanS	7mer-m8	3155			adherent junction; GnRH signaling pathway; MAPK signaling pathway
hsa-miR-520e	TargetScanS	7mer-m8	3155			Pancreatic cancer; Bladder cancer; glioma
hsa-miR-520f	DianaMicroT	7mer(1);8mer(2)	2562, 3183	8-0		TGF-beta signaling pathway; Focal adhesion; Wnt signaling pathway
hsa-miR-524-5p	miRDB(MirTarget2); StarBase; DianaMicroT	7mer(2);8mer; 9mer	1766, 2764, 3299	7-4-0-1		adherent junction; MAPK signaling pathway; GnRH signaling pathway
hsa-miR-539	TargetScanS					Acute myeloid leukemia; TGF-beta signaling pathway; Chronic myeloid leukemia
hsa-miR-544	TargetScanS; miRanda (microrna)	7mer-m8	3214	*		
hsa-miR-545	miRanda (microrna)		949			Regulation of actin cytoskeleton; Melanogenesis; Renal cell carcinoma
hsa-miR-548a-5p	DianaMicroT	7mer;8mer pos1	197, 2463	5-11		Ubiquitin mediated proteolysis; Colorectal cancer; Pancreatic cancer
hsa-miR-548b-5p	DianaMicroT	7mer;8mer pos1	197, 2463	5-11		TGF-beta signaling pathway; colorectal cancer; pancreatic

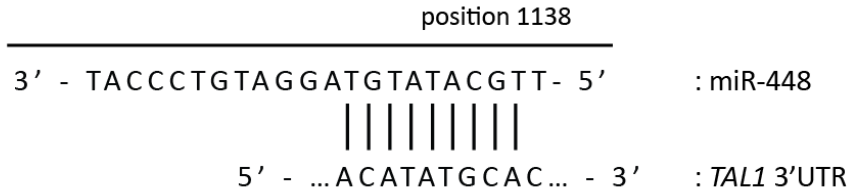
						cancer
hsa-miR-548c-5p	DianaMicroT	7mer;8mer pos1	197, 2463	5-11		TGF-beta signaling pathway; Ubiquitin mediated proteolysis; Colorectal cancer
hsa-miR-548d-5p	DianaMicroT	7mer;8mer pos1	197, 2463	5-11		Colorectal cancer; pancreatic cancer; TGF-beta signaling pathway
hsa-miR-548n	miRDB(MirTarget2)		220, 2485			
hsa-miR-576-3p	miRDB(MirTarget2)		2968			Olfactory transduction; Adipocytokine signaling pathway; Long-term depression
hsa-miR-599	miRanda (microrna)		982	*		
hsa-miR-603	DianaMicroT	7mer pos1	2287	0	✔	Axon guidance; adherent junction; Renal cell carcinoma
hsa-miR-604	miRanda (microrna)		887			ErbB signaling pathway; Type II diabetes mellitus; Thiamine metabolism
hsa-miR-612	miRDB(MirTarget2); DianaMicroT	7mer (2); 9mer	2204 , 2505, 3179	0-0-7		Basal cell carcinoma; Wnt signaling pathway; Endometrial cancer
hsa-miR-613	TargetScanS; miRanda (microrna)	8mer	2815	*		mTOR signaling pathway; Dorso-ventral axis formation; adherent junction
hsa-miR-640	miRDB(MirTarget2)		1994			Basal cell carcinoma; Melanogenesis; Renal cell carcinoma
hsa-miR-7	miRanda (miRBase)					Axon guidance; ErbB signaling pathway; Insulin signaling pathway

hsa-miR-766	DianaMicroT	7mer; 9mer(2)	263, 286, 2339	0		Cell adhesion molecules; Type I diabetes mellitus; Glycerophospholipid metabolism
hsa-miR-874	miRanda (microrna)		2643	*		
hsa-miR-93	TargetScanS; PicTar (4-way); DianaMicroT	7mer-m8	3156	8*	✔	

Several web-based bioinformatics tools (PicTar (4-way), TargetScanS release 4.2, miRBase, microRNA.org, DIANA-microT algorithm, miRDB and StarBase) were used to perform the identification of putative regulators of TAL1. The putative microRNAs binding type, when defined, were listed according to the TargetScanS or DianaMicroT (for 9mer) results for the prediction: 7mer-m8 site comprises the seed plus a match to miRNA nucleotide 8 (miRNA nucleotides 2–8); 7mer-A1 site comprises the seed supplemented by an Adenine at target position 1; 8mer site that comprises a match to miRNA nucleotide 8, the miRNA seed and the A at position 1; and 9mer defines an exact match on the positions 1-10. The 3'UTR position of the putative binding is also depicted. They may vary slightly according to the program used, here we listed the ones predicted by TargetScanS, when available, or the by the program listed in the second column. The conservation score was listed according to DianaMicroT program prediction, when available: it depicts the number of species in which the binding nucleotides of this target site are conserved. We also listed the target sites of conserved microRNA with good mirSVR score according to microrna program, that are signed by the * symbol. The mirSVR scores are based on a regression method for predicting likelihood of target mRNA down-regulation from sequence and structure features in microRNA/mRNA predicted target sites [279]. Additionally we also listed the microRNAs that are also predicted to target LMO2. Finally, we used the DIANA-miRPath tool to predict the main biological pathways were the putative targets of the microRNA are involved.

Supplementary Figure 1

a.



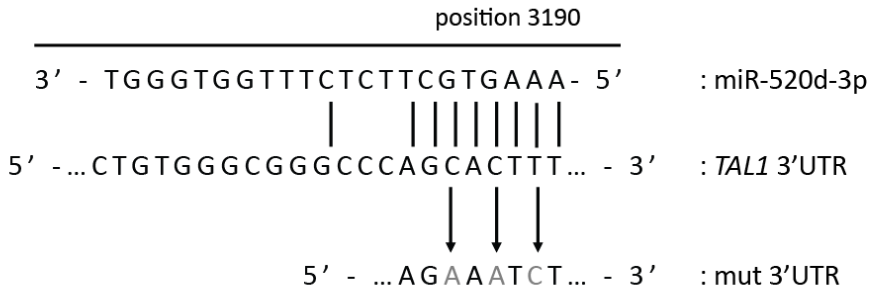
b.

Supplementary Figure 1 – Schematic representation of microRNA binding to TAL1 3'UTR.

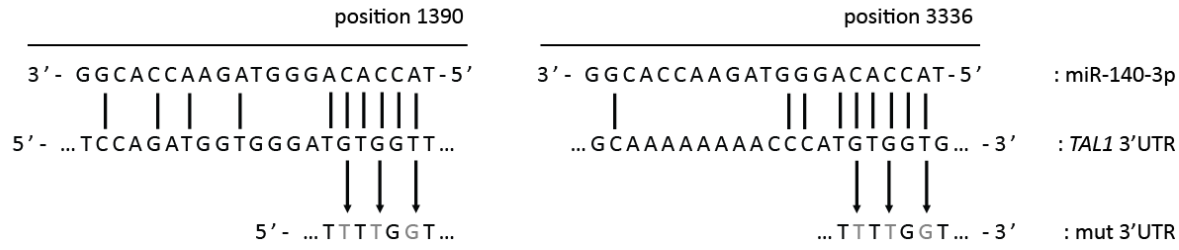
a) miR-448 and **b)** miR-485-5p. miRNA binding to *TAL1* 3'UTR details are depicted according to microRNA.org target prediction algorithm results.

Supplementary Figure 2

a.



b.



Supplementary Figure 2 – Schematic representation of microRNA binding to TAL1 3'UTR and respective mutagenesis.

a) miR-520-3p and **b)** miR-140-3p. These miRNAs are not predicted to bind to *TAL1* 3'UTR, nevertheless the putative MRE were mutated as depicted.

Supplementary Table 2. microRNA gene expression analysis.

The second to seventh columns present the normalized CT values for each detected microRNA in the samples screened. Normalization was performed by the mean expression value normalization method. p-values were calculated using two-tailed Student's t-test.

Detector Name	Ct Empty 1	Ct Empty 2	Ct Empty 3	Ct Tal1 1	Ct Tal1 2	Ct Tal1 3	ddCT	fold change	p-value
hsa-miR-135a	3.82	3.45	4.09	0.20	0.26	-0.16	-3.69	12.90	0.000082
hsa-miR-223	-6.62	-6.54	-6.54	-7.33	-7.26	-7.04	-0.64	1.56	0.002056
hsa-miR-146b-5p	3.26	3.64	3.45	4.62	4.43	4.11	0.94	0.52	0.007100
hsa-miR-503	-0.51	-0.66	-0.67	-1.18	-0.99	-0.93	-0.42	1.34	0.010567
hsa-miR-652	-2.30	-2.51	-2.45	-2.78	-2.74	-3.01	-0.42	1.34	0.016594
hsa-miR-545	3.65	3.97	4.11	4.52	4.37	4.67	0.61	0.66	0.019659
hsa-miR-148a	-3.36	-3.42	-3.37	-3.21	-3.03	-2.89	0.34	0.79	0.021373
hsa-miR-20b	-2.59	-2.67	-2.62	-2.78	-2.94	-2.78	-0.21	1.15	0.021580
hsa-miR-574-3p	6.25	6.75	7.19	5.75	5.87	5.24	-1.11	2.16	0.028828
hsa-miR-20a	-8.16	-8.27	-8.09	-8.02	-7.82	-7.90	0.26	0.84	0.029180
hsa-miR-491-5p	2.63	2.33	2.82	3.16	3.53	3.03	0.64	0.64	0.034590
hsa-miR-375	3.09	3.34	3.25	5.56	3.94	6.06	1.96	0.26	0.038284
hsa-miR-150	-2.63	-2.48	-2.64	-3.14	-2.94	-2.76	-0.37	1.29	0.038396
hsa-miR-330-3p	2.69	3.11	2.41	1.97	2.20	2.12	-0.64	1.55	0.040909
hsa-miR-181a	-7.31	-7.02	-7.38	-7.71	-7.58	-7.52	-0.37	1.29	0.042010
hsa-miR-146a	-1.19	-0.90	-1.12	-1.62	-1.30	-1.44	-0.38	1.30	0.042392
hsa-miR-501-5p	1.52	1.40	1.49	1.61	1.91	1.69	0.27	0.83	0.049712
hsa-miR-29b-2*	0.13	0.51	0.18	-0.28	0.04	-0.18	-0.41	1.33	0.054029
hsa-miR-153	-4.17	-4.06	-4.08	-4.25	-4.20	-4.18	-0.11	1.08	0.055977
hsa-miR-219-5p	0.16	0.02	-0.29	0.20	0.54	0.43	0.43	0.74	0.061502
hsa-miR-30d	-1.17	-0.85	-1.09	-0.33	-0.79	0.09	0.70	0.62	0.062895
hsa-miR-193b	-0.59	0.48	0.62	1.14	1.03	1.78	1.14	0.45	0.063713
hsa-miR-107	-5.49	-5.41	-5.38	-5.31	-5.06	-5.27	0.21	0.86	0.063745
hsa-miR-425	-2.99	-3.28	-2.81	-3.51	-3.64	-3.27	-0.45	1.36	0.064679
hsa-miR-130b	-1.28	-0.99	-1.34	-0.98	-0.67	-0.89	0.35	0.78	0.065254
hsa-miR-27b	-1.65	-1.46	-1.73	-1.37	-1.47	-1.30	0.23	0.85	0.066819
hsa-miR-484	-3.67	-3.51	-3.66	-3.53	-3.28	-3.39	0.22	0.86	0.067794
hsa-miR-514	6.25	6.21	6.11	7.27	6.42	6.75	0.62	0.65	0.069056
hsa-miR-590-5p	-2.13	-2.26	-1.92	-1.92	-1.81	-1.81	0.26	0.84	0.069945

hsa-miR-500	1.50	2.02	1.30	0.77	1.19	1.04	-0.61	1.52	0.070630
hsa-miR-425*	-0.82	-0.53	-0.69	-0.43	-0.35	-0.54	0.24	0.85	0.073797
hsa-miR-766	-2.47	-2.41	-2.43	-2.25	-2.16	-2.40	0.17	0.89	0.074513
hsa-miR-151-5p	0.19	-0.07	0.33	-0.70	-0.18	-0.16	-0.50	1.41	0.078148
hsa-miR-509-3p	5.82	7.44	7.01	4.58	5.39	5.94	-1.45	2.73	0.081093
hsa-miR-362-5p	4.26	3.81	4.09	3.74	3.67	3.80	-0.32	1.25	0.082577
hsa-miR-142-3p	-9.38	-9.53	-9.58	-9.64	-10.09	-9.75	-0.33	1.26	0.087948
hsa-miR-92b	-1.62	-1.93	-1.75	-0.91	-1.50	-1.50	0.46	0.73	0.103038
hsa-miR-328	0.66	1.11	0.96	0.51	0.58	0.73	-0.30	1.23	0.111700
hsa-miR-450a	-0.33	0.06	0.01	-0.24	-0.39	-0.50	-0.29	1.22	0.113072
hsa-miR-589	2.75	3.00	3.17	2.83	2.69	2.46	-0.32	1.24	0.125328
hsa-miR-101	-2.58	-2.42	-2.54	-2.31	-2.49	-2.34	0.14	0.91	0.129458
hsa-miR-423-5p	-2.29	-2.33	-2.52	-2.25	-1.88	-2.22	0.26	0.83	0.129647
hsa-miR-199b-5p	3.94	3.17	3.69	3.28	2.82	3.21	-0.49	1.41	0.139039
hsa-miR-760	4.42	4.29	3.95	5.25	6.46	4.38	1.14	0.45	0.139471
hsa-miR-22*	3.54	3.25	3.39	3.58	3.68	3.48	0.19	0.88	0.139646
hsa-miR-26a-2*	5.99	5.13	5.39	6.26	6.54	5.69	0.66	0.63	0.139965
hsa-miR-361-5p	5.62	5.89	6.11	6.04	6.99	6.31	0.57	0.67	0.142388
hsa-miR-424	-4.27	-4.34	-4.08	-4.51	-4.32	-4.38	-0.17	1.13	0.149796
hsa-miR-194	2.48	2.27	2.54	1.27	1.71	2.44	-0.63	1.54	0.150582
hsa-miR-98	6.45	7.21	6.60	5.71	6.56	6.24	-0.59	1.50	0.158025
hsa-miR-18b	-4.71	-5.05	-4.97	-4.56	-4.73	-4.81	0.21	0.86	0.170112
hsa-miR-188-5p	3.81	3.39	3.52	3.83	3.87	4.98	0.66	0.63	0.171733
hsa-let-7g	-2.45	-2.09	-2.04	-2.01	-1.93	-1.99	0.21	0.86	0.176532
hsa-miR-151-3p	1.10	1.11	2.07	1.67	2.79	2.07	0.75	0.59	0.177609
hsa-miR-196b	0.22	0.40	0.24	0.30	0.48	0.68	0.20	0.87	0.179888
hsa-miR-140-3p	-1.74	-1.58	-1.96	-2.27	-2.04	-1.81	-0.28	1.21	0.180941
hsa-miR-15b	-5.79	-6.03	-6.01	-6.05	-6.58	-6.09	-0.30	1.23	0.184828
hsa-miR-128	-4.52	-4.51	-4.24	-4.28	-4.14	-4.34	0.17	0.89	0.185386
hsa-miR-374b*	5.79	5.30	4.54	5.65	7.51	5.79	1.11	0.46	0.187952
hsa-miR-642	4.77	4.58	5.33	5.16	5.17	5.58	0.41	0.75	0.193765
hsa-miR-365	-0.16	0.07	0.16	0.20	0.10	0.80	0.34	0.79	0.221041
hsa-miR-502-5p	-1.78	-2.12	-2.26	-2.40	-2.06	-3.03	-0.45	1.36	0.232702
hsa-miR-181d	1.65	1.26	1.23	1.56	1.50	1.75	0.22	0.86	0.233556
hsa-miR-570	3.36	2.13	2.38	2.27	2.00	1.99	-0.54	1.45	0.235832
hsa-miR-196a	0.31	0.43	0.42	0.20	0.42	0.20	-0.11	1.08	0.256236
hsa-miR-20b*	0.34	0.63	0.47	0.47	0.82	0.66	0.17	0.89	0.267089
hsa-miR-103	-6.97	-6.79	-6.91	-7.03	-6.90	-7.02	-0.09	1.06	0.268748

hsa-miR-30c	-5.00	-4.78	-5.32	-5.23	-5.31	-5.18	-0.21	1.15	0.268926
hsa-miR-200c	3.54	3.08	3.57	3.34	3.96	3.86	0.32	0.80	0.270059
hsa-miR-23a	-0.85	-0.72	-0.43	-1.16	-1.04	-0.60	-0.27	1.20	0.271755
hsa-miR-199a-3p	4.39	3.74	3.47	4.54	4.02	4.22	0.39	0.76	0.277362
hsa-miR-126	1.40	1.56	1.37	1.05	1.43	1.36	-0.16	1.12	0.277853
hsa-miR-9	0.27	0.89	0.94	-0.10	0.80	0.10	-0.44	1.35	0.278480
hsa-miR-141	5.00	5.06	4.50	4.82	4.47	4.47	-0.27	1.20	0.279235
hsa-miR-30c-2*	4.03	2.62	2.59	4.83	3.53	3.35	0.82	0.57	0.283335
hsa-miR-222	-0.48	-1.39	-1.35	-1.58	-1.35	-1.42	-0.38	1.30	0.284723
hsa-miR-181b	-4.43	-4.91	-5.07	-4.99	-4.92	-5.31	-0.27	1.21	0.296664
hsa-miR-193a-5p	6.31	5.78	6.32	6.31	6.62	6.22	0.25	0.84	0.308875
hsa-miR-340	1.61	1.80	1.26	1.03	1.14	1.66	-0.28	1.21	0.326468
hsa-miR-138	3.09	2.33	2.81	3.01	3.66	2.71	0.39	0.76	0.338066
hsa-let-7f	-2.25	-2.25	-1.96	-2.16	-1.83	-2.05	0.14	0.91	0.352603
hsa-miR-193a-3p	5.85	7.30	7.10	6.92	5.71	5.71	-0.63	1.55	0.354666
hsa-miR-18a	-3.85	-4.01	-4.15	-3.81	-3.76	-4.05	0.13	0.91	0.355906
hsa-miR-629	1.05	1.06	0.62	0.39	0.58	1.02	-0.25	1.19	0.359113
hsa-miR-296-5p	6.17	4.98	7.32	6.35	4.62	4.81	-0.90	1.86	0.359666
hsa-miR-132	0.13	0.19	0.32	0.18	0.28	0.76	0.19	0.88	0.360467
hsa-miR-182	-2.53	-2.32	-2.34	-2.63	-2.48	-2.37	-0.10	1.07	0.370423
hsa-miR-602	4.03	5.35	4.03	3.70	3.51	4.57	-0.54	1.46	0.377500
hsa-miR-96	-2.17	-2.02	-2.19	-2.16	-1.92	-2.04	0.09	0.94	0.387327
hsa-miR-215	1.65	1.48	1.88	1.65	1.39	1.56	-0.14	1.10	0.389938
hsa-miR-23b	-0.83	-0.58	-0.58	-0.73	-0.33	-0.52	0.13	0.91	0.391212
hsa-miR-625*	0.18	-0.33	-0.17	-0.42	0.69	0.47	0.35	0.78	0.398829
hsa-miR-374a	0.29	0.26	-0.30	0.32	0.39	0.11	0.19	0.88	0.416088
hsa-miR-26a	-4.99	-4.63	-4.66	-4.67	-4.83	-4.15	0.21	0.86	0.417115
hsa-miR-19a	-3.73	-3.91	-3.89	-3.94	-4.43	-3.73	-0.19	1.14	0.418386
hsa-miR-361-3p	1.07	0.97	0.92	1.01	0.83	0.94	-0.06	1.04	0.418523
hsa-miR-622	6.03	4.99	5.29	6.78	6.14	5.00	0.53	0.69	0.429399
hsa-miR-550	3.09	3.40	3.58	3.09	2.66	3.56	-0.26	1.19	0.434541
hsa-miR-31	2.54	1.31	2.01	2.66	2.17	2.06	0.35	0.79	0.438209
hsa-miR-572	-0.17	0.01	-0.23	0.05	-0.39	0.99	0.35	0.78	0.445683
hsa-miR-532-5p	1.28	0.50	0.02	0.25	0.54	-0.01	-0.34	1.26	0.448130
hsa-miR-324-3p	-2.27	-2.22	-2.16	-2.33	-2.30	-2.17	-0.05	1.04	0.450359
hsa-miR-30e	-1.97	-1.35	-1.32	-1.74	-1.16	-0.95	0.26	0.83	0.452450
hsa-miR-106a	-7.88	-7.87	-7.88	-7.93	-7.84	-7.93	-0.03	1.02	0.453527
hsa-let-7e	5.03	5.35	5.14	5.14	5.28	5.40	0.10	0.93	0.464845

hsa-miR-663	2.74	1.94	1.73	2.67	2.10	2.47	0.28	0.83	0.469921
hsa-miR-671-5p	1.33	1.28	0.88	1.27	1.55	1.11	0.15	0.90	0.470775
hsa-miR-221	-2.52	-2.48	-1.95	-2.71	-2.37	-2.38	-0.17	1.13	0.475720
hsa-miR-139-5p	4.08	4.42	4.49	5.13	4.18	4.40	0.24	0.84	0.481149
hsa-miR-93	-7.09	-6.78	-6.95	-7.14	-6.93	-7.01	-0.08	1.06	0.481216
hsa-miR-665	2.37	1.71	1.02	2.25	0.26	0.96	-0.54	1.46	0.483505
hsa-let-7d	3.47	5.99	4.19	3.85	4.01	4.07	-0.57	1.49	0.487519
hsa-miR-200a	5.96	7.12	5.21	6.12	6.66	6.84	0.45	0.73	0.497809
hsa-let-7a	-0.89	-2.12	-2.08	-2.04	-1.84	-2.14	-0.31	1.24	0.500801
hsa-miR-30e*	-0.62	-1.31	-0.78	-0.78	-0.56	-0.86	0.17	0.89	0.502214
hsa-miR-301b	-1.14	-0.56	-0.68	-1.31	-1.07	-0.60	-0.20	1.15	0.505207
hsa-let-7d*	1.33	1.28	1.09	1.30	1.15	1.03	-0.08	1.05	0.513348
hsa-miR-345	1.17	0.89	1.86	0.96	1.17	1.16	-0.21	1.16	0.520043
hsa-miR-197	-2.46	-2.37	-2.53	-2.67	-2.45	-2.43	-0.06	1.05	0.525333
hsa-miR-195	2.83	2.97	5.29	3.29	3.17	3.00	-0.54	1.46	0.535421
hsa-miR-301a	-3.82	-3.86	-3.88	-4.03	-3.65	-3.62	0.09	0.94	0.542762
hsa-miR-28-5p	-0.65	-0.02	-0.06	-0.44	0.15	0.12	0.19	0.88	0.544032
hsa-miR-660	1.28	1.21	0.89	2.56	1.14	0.78	0.37	0.78	0.546601
hsa-miR-145	2.45	2.64	2.35	2.55	2.41	2.74	0.08	0.94	0.550921
hsa-miR-16	-6.98	-6.93	-6.87	-6.18	-6.94	-7.11	0.18	0.88	0.556475
hsa-miR-363*	3.61	4.63	4.68	4.46	3.92	3.76	-0.26	1.20	0.557152
hsa-miR-30b*	1.88	1.31	1.83	2.07	1.18	1.09	-0.23	1.17	0.564083
hsa-miR-548b-3p	6.17	5.93	5.12	6.23	5.30	6.66	0.32	0.80	0.566582
hsa-miR-615-3p	3.87	4.56	3.85	4.25	3.89	4.79	0.21	0.86	0.573312
hsa-miR-21*	2.80	2.45	2.86	2.81	2.50	2.52	-0.09	1.07	0.590173
hsa-miR-106b	-3.21	-3.50	-3.44	-2.82	-3.49	-3.46	0.13	0.91	0.609844
hsa-miR-34c-3p	2.27	2.89	2.61	2.12	3.17	1.69	-0.26	1.20	0.612025
hsa-miR-29a	-0.79	-1.21	-1.40	-1.36	-1.08	-1.29	-0.11	1.08	0.612360
hsa-miR-190	3.18	3.05	3.27	4.06	2.92	3.11	0.19	0.88	0.620201
hsa-miR-505	0.46	0.32	-0.03	0.44	0.44	0.15	0.09	0.94	0.621448
hsa-miR-181c	0.68	0.61	1.18	0.79	1.06	0.92	0.10	0.93	0.622658
hsa-miR-155	-2.32	-1.99	-2.10	-2.24	-2.00	-1.96	0.07	0.95	0.623293
hsa-miR-216b	5.39	5.15	5.09	5.38	5.46	4.08	-0.24	1.18	0.629242
hsa-miR-346	3.84	3.32	4.04	3.56	3.72	4.47	0.18	0.88	0.637894
hsa-miR-7	0.51	0.05	-0.41	-0.01	0.59	0.07	0.17	0.89	0.639984
hsa-miR-27a	-1.85	-1.81	-1.67	-1.77	-1.86	-1.78	-0.03	1.02	0.644751
hsa-miR-148b	-3.14	-2.85	-2.83	-2.93	-3.12	-2.95	-0.06	1.04	0.649274
hsa-miR-22	1.51	1.72	1.38	1.72	1.95	1.26	0.11	0.93	0.657996

hsa-miR-95	1.43	1.92	1.19	1.44	2.73	1.13	0.25	0.84	0.662929
hsa-miR-191	-3.90	-4.07	-4.12	-4.14	-3.91	-4.20	-0.05	1.04	0.672211
hsa-miR-140-5p	-0.77	-1.01	-0.92	-1.16	-0.84	-0.87	-0.06	1.04	0.675227
hsa-miR-181a*	-0.41	-0.45	-0.50	-0.03	-0.64	-1.08	-0.13	1.09	0.692030
hsa-miR-185	-2.27	-2.44	-2.56	-2.48	-2.35	-2.31	0.04	0.97	0.693710
hsa-miR-720	-5.57	-5.08	-5.08	-5.76	-5.27	-5.03	-0.11	1.08	0.700137
hsa-miR-33a	-3.96	-3.47	-3.62	-4.05	-3.94	-3.37	-0.10	1.07	0.705818
hsa-miR-210	-5.12	-4.52	-4.56	-4.87	-4.99	-3.87	0.15	0.90	0.723296
hsa-miR-454	-0.47	-0.44	-0.74	-0.65	-0.63	-0.50	-0.04	1.03	0.725827
hsa-miR-26b	-2.50	-2.58	-2.55	-2.47	-2.64	-2.45	0.02	0.98	0.726570
hsa-miR-576-3p	4.66	4.10	4.34	5.64	3.39	4.84	0.25	0.84	0.727678
hsa-let-7i	2.96	2.92	2.80	2.54	3.28	2.60	-0.09	1.06	0.728789
hsa-miR-192	0.38	0.78	0.32	0.25	0.51	0.54	-0.06	1.04	0.742232
hsa-miR-32	-4.25	-4.22	-4.08	-4.26	-4.15	-4.07	0.03	0.98	0.742343
hsa-miR-25	-4.37	-4.59	-4.63	-4.69	-4.39	-4.37	0.05	0.97	0.743510
hsa-let-7c	2.30	3.07	3.35	2.69	4.12	2.53	0.21	0.87	0.747065
hsa-miR-18a*	-1.20	-1.54	-1.64	-1.46	-1.46	-1.60	-0.05	1.03	0.753884
hsa-miR-338-3p	4.82	3.71	5.09	4.45	4.45	5.18	0.15	0.90	0.766918
hsa-miR-627	3.40	3.98	4.07	4.11	3.43	4.19	0.10	0.93	0.775678
hsa-miR-17	-3.63	-4.22	-4.11	-3.49	-4.33	-3.86	0.09	0.94	0.778927
hsa-miR-199a-5p	6.86	6.48	6.27	6.85	5.21	7.02	-0.18	1.13	0.782708
hsa-miR-744	1.37	1.01	0.67	1.36	0.91	0.97	0.07	0.95	0.800982
hsa-miR-542-5p	1.42	1.43	2.02	1.58	1.68	1.45	-0.05	1.04	0.819452
hsa-miR-421	0.45	-0.26	0.11	-0.03	0.17	0.00	-0.05	1.04	0.820257
hsa-miR-92a	-7.53	-7.22	-7.70	-7.45	-7.60	-7.50	-0.03	1.02	0.827535
hsa-miR-934	5.67	5.48	5.09	5.34	5.28	5.77	0.05	0.96	0.833448
hsa-miR-149	3.81	4.58	3.85	4.08	4.44	3.92	0.06	0.96	0.836850
hsa-miR-33b	0.79	0.65	2.13	1.97	1.11	0.87	0.13	0.92	0.836993
hsa-miR-9*	4.15	3.26	4.43	3.64	4.24	4.21	0.09	0.94	0.840513
hsa-miR-940	-0.47	-0.53	-0.64	0.10	-0.94	-0.60	0.07	0.95	0.841596
hsa-miR-373*	5.31	4.42	3.84	5.13	4.13	3.96	-0.11	1.08	0.848090
hsa-miR-877	1.52	0.67	0.70	0.96	0.86	0.91	-0.06	1.04	0.852306
hsa-miR-142-5p	-5.19	-5.34	-4.52	-5.05	-4.96	-5.19	-0.05	1.04	0.858633
hsa-miR-1979	-7.68	-8.70	-8.31	-7.81	-8.13	-9.02	-0.09	1.06	0.859521
hsa-miR-21	-5.53	-4.41	-4.82	-5.43	-4.11	-4.96	0.09	0.94	0.870639
hsa-miR-885-5p	4.40	4.01	3.89	3.57	4.62	3.93	-0.06	1.04	0.873436
hsa-miR-431	4.51	3.70	3.97	4.50	4.35	3.52	0.06	0.96	0.880718
hsa-let-7b	0.30	0.52	0.06	0.15	0.67	-0.04	-0.04	1.03	0.886549

hsa-miR-339-5p	0.70	0.55	0.68	0.61	0.58	0.71	-0.01	1.01	0.890510
hsa-miR-326	2.80	2.21	3.45	3.32	2.50	2.46	-0.06	1.04	0.903217
hsa-miR-374b	-1.66	-1.59	-2.34	-1.81	-1.91	-1.96	-0.03	1.02	0.909933
hsa-miR-331-3p	-2.90	-2.89	-3.02	-3.01	-2.98	-2.84	-0.01	1.01	0.910861
hsa-miR-320a	-5.47	-5.41	-5.42	-5.52	-5.59	-5.14	0.02	0.99	0.911347
hsa-miR-324-5p	-2.77	-2.58	-2.56	-2.82	-2.55	-2.50	0.01	0.99	0.919008
hsa-miR-92a-1*	-0.63	-0.48	-0.48	-0.77	-0.56	-0.30	-0.01	1.01	0.929499
hsa-miR-143	1.57	2.51	2.31	1.87	1.96	2.66	0.04	0.98	0.930560
hsa-miR-183	-0.41	0.50	0.40	0.79	-0.41	0.00	-0.04	1.03	0.938704
hsa-miR-363	-2.80	-2.69	-3.64	-2.51	-3.44	-3.29	-0.03	1.02	0.940303
hsa-miR-31*	3.76	4.15	5.13	4.48	4.65	3.80	-0.04	1.02	0.945267
hsa-miR-29c	-4.66	-4.69	-3.97	-4.34	-4.50	-4.53	-0.02	1.01	0.952637
hsa-miR-126*	3.84	5.48	4.67	4.53	4.18	5.19	-0.03	1.02	0.955280
hsa-miR-342-3p	-3.58	-3.66	-3.58	-3.93	-3.33	-3.59	-0.01	1.01	0.956477
hsa-miR-378	-1.75	-2.44	-2.57	-2.31	-2.06	-2.43	-0.01	1.01	0.963598
hsa-miR-30b	-5.05	-4.74	-4.84	-5.05	-4.80	-4.80	-0.01	1.00	0.963711
hsa-miR-15a	-7.02	-7.21	-7.13	-7.16	-7.07	-7.14	0.00	1.00	0.971982
hsa-miR-497	3.47	3.54	3.21	3.31	3.15	3.77	0.00	1.00	0.984350
hsa-miR-29b	-3.24	-2.97	-2.94	-2.96	-3.01	-3.18	0.00	1.00	0.984619
hsa-miR-19b	-8.39	-8.29	-8.41	-8.49	-8.39	-8.21	0.00	1.00	0.991189
hsa-miR-24	-1.72	-1.59	-1.34	-1.87	-1.37	-1.42	0.00	1.00	0.992020
hsa-miR-423-3p	-4.54	-4.40	-4.31	-4.51	-4.48	-4.26	0.00	1.00	0.992399
hsa-miR-130a	-0.31	-0.39	-0.19	-0.22	-0.37	-0.31	0.00	1.00	0.994561

Supplementary Table 3 – Gene set enrichment analysis for the experimentally validated targets of TAL1 regulated microRNAs.

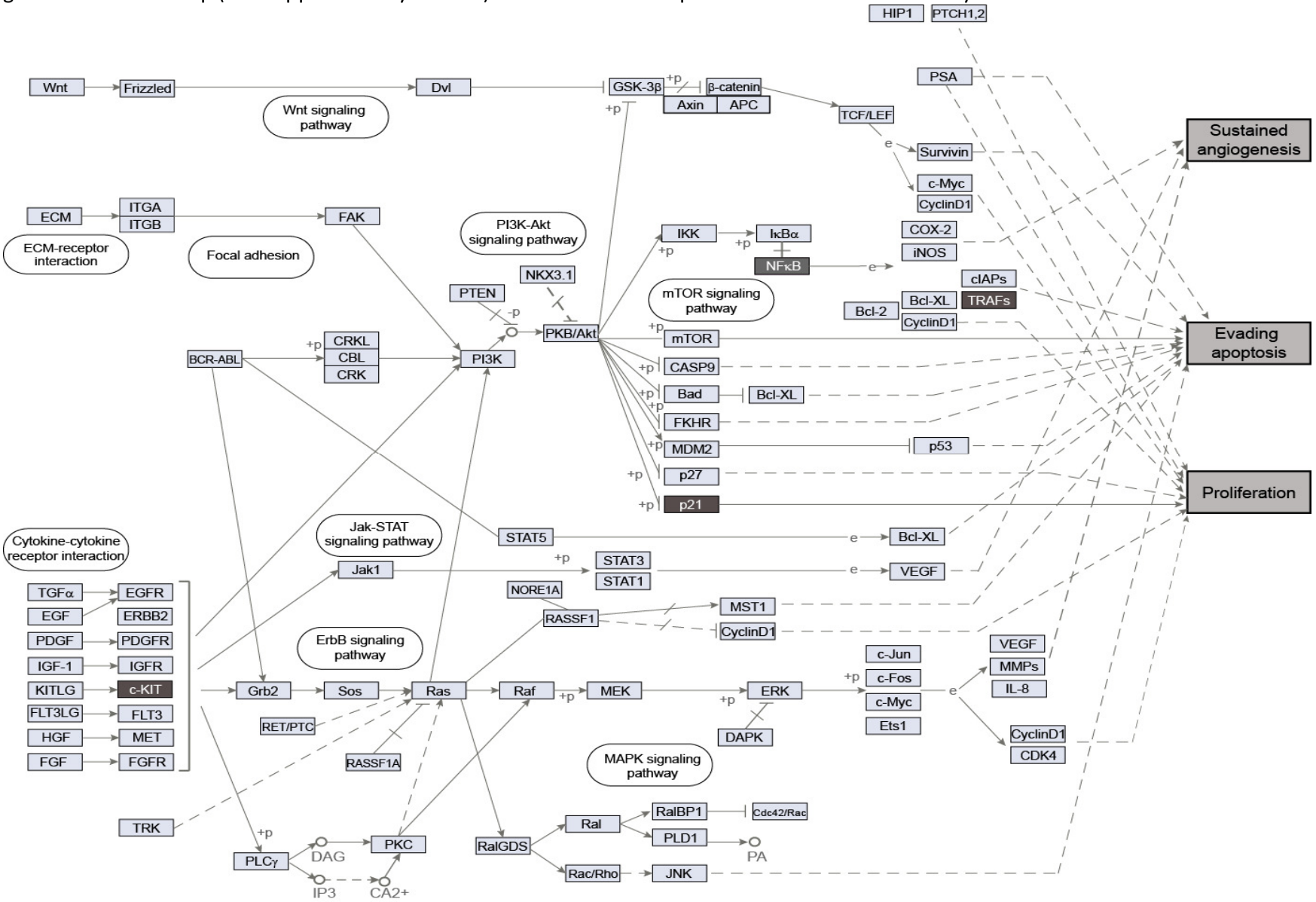
Analysis was performed using Genomica software (http://genomica.weizmann.ac.il/). p-values were determined by a hypergeometric test, followed by a false discovery rate correction to account for multiple hypotheses (FDR <0.05).

microRNA	Modulation by TAL1	Enriched Set	Description	Pvalue	Set Hits	Targets of the individual miR	Set Size	Set Hits (%)	Total Hits	Total Size	Total Hits (%)
hsa-miR-146b-5p	Down	BIOCARTA_NFKB_PATHWAY	NF-kB Signaling Pathway	5,22E-08	3	IRAK1, NFKB1, TRAF6	6	50	23	15958	0,15
		BIOCARTA_IL1R_PATHWAY	Signal transduction through IL1R	1,60E-07	3	IRAK1, NFKB1, TRAF6	6	50	33	15958	0,21
		PID_IL1PATHWAY	IL1-mediated signaling events	1,76E-07	3	IRAK1, NFKB1, TRAF6	6	50	34	15958	0,22
		BIOCARTA_TOLL_PATHWAY	Toll-Like Receptor Pathway	2,28E-07	3	IRAK1, NFKB1, TRAF6	6	50	37	15958	0,24
		KEGG_LEISHMANIA_INFECTION	Leishmania infection	1,74E-06	3	IRAK1, NFKB1, TRAF6	6	50	72	15958	0,46
		KEGG_PATHWAYS_IN_CANCER	Pathways in cancer	2,54E-06	4	CDKN1A, KIT, NFKB1, TRAF6	6	66,67	328	15958	2,06
		KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY	Pathways in cancer	5,00E-06	3	IRAK1, NFKB1, TRAF6	6	50	102	15958	0,64
		KEGG_NEUROTROPHIN_SIGNALING_PATHWAY	Neurotrophin signaling pathway	9,45E-06	3	IRAK1, NFKB1, TRAF6	6	50	126	15958	0,79
		REACTOME_SIGNALLING_BY_NGF	Genes involved in Signalling by NGF	4,81E-05	3	IRAK1, NFKB1, TRAF6	6	50	217	15958	1,36
		LI_INDUCED_T_TO_NATURAL_KILLER_UP	Genes up-regulated in ITNK cells (T-lymphocyte progenitors (DN3 cells) reprogrammed to natural killer (NK) cells by ablation of BCL11B [GeneID=64919] gene), compared to the parental DN3 cells.	9,43E-05	3	CDKN1A, NFKB1, KIT	6	50	272	15958	1,71

microRNA	Modulation by TAL1	Enriched Set	Description	Pvalue	Set Hits	Targets of the individual miR considered	Set Size	Set Hits (%)	Total Hits	Total Size	Total Hits (%)
hsa-mir-330-3p	UP	KEGG_PANCREATIC_CANCER	Pancreatic cancer	8,03E-07	3	CDC42, E2F1, VEGFA	5	60	70	15958	0,44
		KEGG_PATHWAYS_IN_CANCER	Pathways in cancer	8,34E-05	3	CDC42, E2F1, VEGFA	5	60	328	15958	2,06

microRNA	Modulation by TAL1	Enriched Set	Description	Pvalue	Set Hits	Targets of the individual miR considered	Set Size	Set Hits (%)	Total Hits	Total Size	Total Hits (%)
hsa-miR-223	UP	PLASARI_TGFB1_TARGETS_10HR_DN	Genes down-regulated in MEF cells (embryonic fibroblast) upon stimulation with TGFB1 [GeneID=7040] for 10 h.	2,29E-06	5	IRS1, STMN1, MEF2C, NFIA, NFIX	16	31,25	230	15958	1,45
		RIZ_ERYTHROID_DIFFERENTIATION _CCNE1	Selected gradually up-regulated genes whose expression profile follows that of CCNE1 [GeneID=898] in the TLX1 [GeneID=3195] Tet On iEBHX15-4 cells (pro-erythroblasts).	7,99E-06	3	E2F1, NFIX, SP3	16	18,75	40	15958	0,26
		STARK_PREFRONTAL_CORTEX_22Q11 _DELETION_UP	Genes up-regulated in prefrontal cortex (PFC) of mice carrying a hemizygotic microdeletion in the 22q11.2 region.	2,40E-05	4	MEF2C, NFIX, SP3, IGF1R, Lpin2	16	25	177	15958	1,11
		KEGG_ADIPOCYTOKINE_SIGNALING _PATHWAY	Adipocytokine signaling pathway	3,81E-05	3	CHUK, IRS1, SLC2A4	16	18,75	67	15958	0,42
		PID_P75NTRPATHWAY	p75(NTR)-mediated signaling	4,16E-05	3	RHOB, CHUK, E2F1	16	18,75	69	15958	0,44
		WESTON_VEGFA_TARGETS_3HR	Genes up-regulated in MMEC cells (myometrial endothelium) at 3 h after VEGFA	4,73E-05	3	RHOB, LMO2, SP3	16	18,75	72	15958	0,46
		REACTOME_NFKB_AND_MAP_KINASES_ ACTIVATION_MEDIATED_BY_TLR4 _SIGNALING_REPERTOIRE	Genes involved in NFkB and MAP kinases activation mediated by TLR4 signaling repertoire	4,73E-05	3	MEF2C, CHUK, S100B	16	18,75	72	15958	0,46
		REACTOME_TRIF_MEDIATED_TLR3 SIGNALING	Genes involved in TRIF mediated TLR3 signaling	5,13E-05	3	MEF2C, CHUK, S100B	16	18,75	74	15958	0,47
		RIZ_ERYTHROID_DIFFERENTIATION	Selected gradually up-regulated genes in the TLX1 [GeneID=3195] Tet On iEBHX15-4 cells (pro-erythroblasts).	5,78E-05	3	E2F1, NFIX, SP3	16	18,75	77	15958	0,49
		REACTOME_TRAF6_MEDIATED _INDUCTION_OF_NFKB_AND_MAP _KINASES_UPON_TLR7_8_OR_9 _ACTIVATION	Genes involved in TRAF6 mediated induction of NFkB and MAP kinases upon TLR7/8 or 9 activation	5,78E-05	3	MEF2C, CHUK, S100B	16	18,75	77	15958	0,49
		REACTOME_MYD88_MAL_CASCADE _INITIATED_ON_PLASMA_MEMBRANE	Genes involved in MyD88:Mal cascade initiated on plasma membrane	7,24E-05	3	MEF2C, CHUK, S100B	16	18,75	83	15958	0,53
		KEGG_PROSTATE_CANCER	Prostate cancer	8,91E-05	3	IGF1R, CHUK, E2F1	16	18,75	89	15958	0,56
		REACTOME_ACTIVATED_TLR4 _SIGNALLING	Genes involved in Activated TLR4 signalling	1,02E-04	3	MEF2C, CHUK, S100B	16	18,75	93	15958	0,59
		MCCLUNG_CREB1_TARGETS_UP	Genes up-regulated in the nucleus accumbens (a major reward center in the brain) 8 weeks after induction of CREB1	1,12E-04	3	MEF2C, FBXW7, STMN1	16	18,75	96	15958	0,61
		WESTON_VEGFA_TARGETS	Genes up-regulated in MMEC cells (myometrial endothelium) by VEGFA	1,50E-04	3	RHOB, LMO2, SP3	16	18,75	106	15958	0,67

Supplementary Figure 3 – Association of miR-146b-5p target genes with cancer pathways. Highlighted are experimentally validated target genes of miR-146b-5p (see Supplementary Table 3) within a zoomed representation of KEGG Pathways in Cancer.



Supplementary Table 4 – List of known TAL1 up-regulated genes that are potentially targeted by TAL1 down-regulated microRNAs, as in Figure 3.2.5a. Described function(s), putative role in cancer and corresponding relevant references are presented for each gene. For gene assessment details see supplemental methods.

Gene name	KRT1 (cytokeratin-1)	RAPGEF5	JAZF1	FAM64A
Function	Keratin 1 gene; upregulation of KRT1 could result in increased of drug resistance in nasopharyngeal carcinoma cell lines cell lines. The cytokeratin 1-Src complex was found to be associated with the molecular scaffolder RACK1 in neuroblastoma.	It activates Ras oncogene family members RAP1A and RAP1B. Guanine nucleotide exchange factor (GEF) for RAP1A, RAP2A and MRAS/M-Ras-GTP. Its association with MRAS inhibits Rap1 activation .	Potential transcription factor. A chromosomal aberration (Translocation t(7;17)(p15;q21) with SUZ12.) involving JAZF1 may be a cause of endometrial stromal tumors. The translocation generates the JAZF1-SUZ12 oncogene consisting of the N-terminus part of JAZF1 and the C-terminus part of SUZ12. It is frequently found in all cases of endometrial stromal tumors. Single nucleotide polymorphisms in this gene are associated with altered risk for type 2 diabetes and prostate cancer. 2 isoforms of the human protein are produced by alternative splicing.	family with sequence similarity 64, member A
Role in Cancer	Potential oncogene	Potential oncogene	Potential oncogene	Not Determined
References	1: Tang S, Huang W, Zhong M, Yin L, Jiang H, Hou S, Gan P, Yuan Y. Identification Keratin 1 as a cDDP-resistant protein in nasopharyngeal carcinoma cell lines. J Proteomics. 2012 Apr 18;75(8):2352-60. doi: 10.1016/j.jprot.2012.02.003. Epub 2012 Feb 12. PubMed PMID: 22348822. 2: Attallah AM, El-Far M, Abdel Malak CA, Zahran F, Farid K, Omran MM, Zagloul H, El-Deen MS. Evaluation of cytokeratin-1 in the diagnosis of hepatocellular carcinoma. Clin Chim Acta. 2011 Nov 20;412(23-24):2310-5. doi: 10.1016/j.cca.2011.08.029. Epub 2011 Sep 7. PubMed PMID: 21924253. 3: Chuang NN, Huang CC. Interaction of integrin beta1 with cytokeratin 1 in neuroblastoma NMB7 cells. Biochem Soc Trans. 2007 Nov;35(Pt 5):1292-4. Review. PubMed PMID: 17956333.	WEB reference: http://www.genecards.org/cgi-bin/carddisp.pl?gene=RAPGEF5	1: Amador-Ortiz C, Roma AA, Huettner PC, Becker N, Pfeifer JD. JAZF1 and JJAZ1 gene fusion in primary extrauterine endometrial stromal sarcoma. Hum Pathol. 2011 Jul;42(7):939-46. doi: 10.1016/j.humpath.2010.11.001. Epub 2011 Feb 11. PubMed PMID: 21316079. 2: Li H, Wang J, Mor G, Sklar J. A neoplastic gene fusion mimics trans-splicing of RNAs in normal human cells. Science. 2008 Sep 5;321(5894):1357-61. doi: 10.1126/science.1156725. PubMed PMID: 18772439. 3: Koontz JI, Soreng AL, Nucci M, Kuo FC, Pauwels P, van Den Berghe H, Dal Cin P, Fletcher JA, Sklar J. Frequent fusion of the JAZF1 and JJAZ1 genes in endometrial stromal tumors. Proc Natl Acad Sci U S A. 2001 May 22;98(11):6348-53. 4: Stevens VL, Ahn J, Sun J, Jacobs EJ, Moore SC, Patel AV, Berndt SI, Albanes D, Hayes RB. HNF1B and JAZF1 genes, diabetes, and prostate cancer risk. Prostate. 2010 May 1;70(6):601-7. doi:10.1002/pros.21094.	

Gene name	CD53	ARID3A (DRIL1)	SCML1	TLE4
Function	Leukocyte surface antigen of the as the tetraspanin family. In CD53- stimulated cells there is a significant reduction in caspase activation, as a reduction in the fragmentation of DNA. CD53-stimulated cells also have an increase in the level of bcl-X(L) and a reduction of bax protein, changing their ratio by 24-fold in the direction of survival. This survival signal appears to be mediated by activation of the AKT, as detected by its phosphorylation in Ser473 upon CD53 ligation.	AT-rich interaction domain DNA-binding transcription factor. DRIL1 renders primary murine fibroblasts unresponsive to RAS(V12)-induced anti-proliferative signalling by p19(ARF)/p53/p21(CIP1), as well as by p16(INK4a). DRIL1 induces the E2F1 target Cyclin E1, overexpression of which is sufficient to trigger escape from senescence.	Polycomb group protein.	Novel Groucho-transcription factor, regulates the repressive activity of PAX5 in human B-lymphocytes. Knockdown of TLE1 or TLE4 levels increased the rate of cell division of the AML1-ETO-expressing Kasumi-1 cell line, whereas forced expression of either TLE1 or TLE4 caused apoptosis and cell death.
Role in Cancer	Anti-apoptotic (in Jurkat cells)	Oncogene	Not Determined	Pro-apoptotic
References	1: Yunta M, Lazo PA. Apoptosis protection and survival signal by the CD53 tetraspanin antigen. Oncogene. 2003 Feb 27;22(8):1219-24. PubMed PMID: 12606948.	Peeper DS, Shvarts A, Brummelkamp T, Douma S, Koh EY, Daley GQ, Bernards R. A functional screen identifies hDRIL1 as an oncogene that rescues RAS-induced senescence. Nat Cell Biol. 2002 Feb;4(2):148-53.		Dayyani F, Wang J, Yeh JR, Ahn EY, Tobey E, Zhang DE, Bernstein ID, Peterson RT, Sweetser DA. Loss of TLE1 and TLE4 from the del(9q) commonly deleted region in AML cooperates with AML1-ETO to affect myeloid cell proliferation and survival. Blood. 2008 Apr 15;111(8):4338-47. doi: 10.1182/blood-2007-07-103291.

Gene name	ETS-1	CD28
Function	Member of the ETS family of transcription factors involved in stem cell development, cell senescence and death, and tumorigenesis. cortical thymic maturation arrest in T-ALLs that overexpress TLX1 or TLX3 is due to binding of TLX1/TLX3 to ETS1, leading to repression of T cell receptor α enhanceosome activity and blocked TCR-Jα rearrangement.	TCR co-stimulatory signal receptor for T cell activation. Involved in induction of cell proliferation and cytokine production and promotion of T-cell survival; CD28 co-stimulation directly controls T cell cycle progression by down-regulating the cdk inhibitor p27kip1, which actually integrates mitogenic MEK and P13K-dependent signals from both TCR and CD28.
Role in Cancer	Oncogene	Pro-proliferative
References	1: Wei W, et al. MicroRNA-1 and microRNA-499 downregulate the expression of the ets1 proto-oncogene in HepG2 cells. Oncol Rep.2012 Aug;28(2):701-6. doi: 10.3892/or.2012.1850. Epub 2012 Jun 1. 2: Pallai R, Bhaskar A, Sodi V, Rice LM. Ets1 and Elk1 transcription factors regulate cancerous inhibitor of protein phosphatase 2A expression in cervical and endometrial carcinoma cells. Transcription. 2012 Nov 1;3(6). [Epub ahead of print] PubMed PMID: 23117818. 3: Dadi S, et al. TLX homeodomain oncogenes mediate T cell maturation arrest in T-ALL via interaction with ETS1 and suppression of TCRα gene expression. Cancer Cell. 2012 Apr 17;21(4):563-76. doi: 10.1016/j.ccr.2012.02.013. PubMed PMID: 22516263. 4: Wang C, et al. Gambogic acid-loaded magnetic Fe(3)O(4) nanoparticles inhibit Panc-1 pancreatic cancer cell proliferation and migration by inactivating transcription factor ETS1. Int J Nanomedicine.2012;7:781-7. doi: 10.2147/IJN.S28509. Epub 2012 Feb 14. PubMed PMID: 22393285; 5: Shaikhibrahim Z, Wernert N. ETS transcription factors and prostate cancer: the role of the family prototype ETS-1 (review). Int J Oncol. 2012 Jun;40(6):1748-54. doi: 10.3892/ijo.2012.1380. Epub 2012 Feb 21. Review. PubMed PMID: 22366814. 6: Smith AM, et al. ETS1 transcriptional activity is increased in advanced prostate cancer and promotes the castrate-resistant phenotype. Carcinogenesis. 2012 Mar;33(3):572-80. doi: 10.1093/carcin/bgs007. Epub 2012 Jan 9. 7: Kato T, et al. ETS1 promotes chemoresistance and invasion of paclitaxel-resistant, hormone-refractory PC3 prostate cancer cells by up-regulating MDR1 and MMP9 expression. Biochem Biophys Res Commun. 2012 Jan 20;417(3):966-71. doi: 10.1016/j.bbrc.2011.12.047. Epub 2011 Dec 20. 8: Singh AK, Swarnalatha M, Kumar V. c-ETS1 facilitates G1/S-phase transition by up-regulating cyclin E and CDK2 genes and cooperates with hepatitis B virus X protein for their deregulation. J Biol Chem. 2011 Jun 24;286(25):21961-70. doi:10.1074/jbc.M111.238238. Epub 2011 Apr 22. PubMed PMID: 21515670; 9: Khanna A, et al. ETS1 mediates MEK1/2-dependent overexpression of cancerous inhibitor of protein phosphatase 2A (CIP2A) in human cancer cells. PLoS One. 2011 Mar 22;6(3):e17979. doi:10.1371/journal.pone.0017979. 10: Zhang Y, et al. miR-125b is methylated and functions as a tumor suppressor by regulating the ETS1 proto-oncogene in human invasive breast cancer. Cancer Res. 2011 May 15;71(10):3552-62. doi: 10.1158/0008-5472.CAN-10-2435.	1: Appleman LJ, van Puijenbroek AA, Shu KM, Nadler LM, Boussiotis VA. CD28 costimulation mediates down-regulation of p27kip1 and cell cycle progression by activation of the PI3K/PKB signaling pathway in primary human T cells. J Immunol. 2002 Mar 15;168(6):2729-36. 2: Takeda K, Harada Y, Watanabe R, Inutake Y, Ogawa S, Onuki K, Kagaya S, Tanabe K, Kishimoto H, Abe R. CD28 stimulation triggers NF-kappaB activation through the CARMA1-PKCtheta-Grb2/Gads axis. Int Immunol. 2008 Dec;20(12):1507-15. doi: 10.1093/intimm/dxn108. Epub 2008 Oct 1.

Gene name	UBE3C	MYB
Function	E3 ubiquitin-protein ligase	Myb proto-oncogene protein is a member of the MYB (myeloblastosis) family of transcription factors. C-MYB locus is involved in chromosomal translocation and genomic duplications in human T-cell acute leukemia.
Role in Cancer	Not Determined	Oncogene
References		1: Sanda T, Lawton LN, Barrasa MI, Fan ZP, Kohlhammer H, Gutierrez A, Ma W, Tatarek J, Ahn Y, Kelliher MA, Jamieson CH, Staudt LM, Young RA, Look AT. Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. Cancer Cell. 2012 Aug 14;22(2):209-21. doi: 10.1016/j.ccr.2012.06.007. PubMed PMID: 22897851; 2: Kawamata N, Zhang L, Ogawa S, Nannya Y, Dashti A, Lu D, Lim S, Schreck R, Koeffler HP. Double minute chromosomes containing MYB gene and NUP214-ABL1 fusion gene in T-cell leukemia detected by single nucleotide polymorphism DNA microarray and fluorescence in situ hybridization. Leuk Res. 2009 Apr;33(4):569-71. doi: 10.1016/j.leukres.2008.07.030. Epub 2008 Sep 16. 3: O'Neil J, Tchinda J, Gutierrez A, Moreau L, Maser RS, Wong KK, Li W, McKenna K, Liu XS, Feng B, Neuberg D, Silverman L, DeAngelo DJ, Kutok JL, Rothstein R, DePinho RA, Chin L, Lee C, Look AT. Alu elements mediate MYB gene tandem duplication in human T-ALL. J Exp Med. 2007 Dec 24;204(13):3059-66. Epub 2007 Dec 10. PubMed PMID: 18070937; PubMed Central PMCID: PMC2150982. 4: Clappier E, Cuccuini W, Kalota A, Crinquette A, Cayuela JM, Dik WA, Langerak AW, Montpellier B, Nadel B, Walrafen P, Delattre O, Aurias A, Leblanc T, Dombret H, Gewirtz AM, Baruchel A, Sigaux F, Soulier J. The C-MYB locus is involved in chromosomal translocation and genomic duplications in human T-cell acute leukemia (T-ALL), the translocation defining a new T-ALL subtype in very young children. Blood. 2007 Aug 15;110(4):1251-61. Epub 2007 Apr 23. PubMed PMID: 17452517. 5: Lahortiga I, De Keersmaecker K, Van Vlierberghe P, Graux C, Cauwelier B,nLambert F, Mentens N, Beverloo HB, Pieters R, Speleman F, Odero MD, Bauters M, Froyen G, Marynen P, Vandenberghe P, Wlodarska I, Meijerink JP, Cools J. Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia. Nat Genet. 2007 May;39(5):593-5. Epub 2007 Apr 15. PubMed PMID: 17435759. 6: Venturelli D, Mariano MT, Szczylik C, Valtieri M, Lange B, Crist W, Link M, Calabretta B. Down-regulated c-myb expression inhibits DNA synthesis of T-leukemia cells in most patients. Cancer Res. 1990 Nov 15;50(22):7371-5. PubMed PMID: 2224864. 7: Mavilio F, Sposi NM, Petrini M, Bottero L, Marinucci M, De Rossi G, Amadori S, Mandelli F, Peschle C. Expression of cellular oncogenes in primary cells from human acute leukemias. Proc Natl Acad Sci U S A. 1986 Jun;83(12):4394-8. PubMed PMID: 3520570; PubMed Central PMCID: PMC323739.

Gene name	ETV5	MED12L
Function	ETS oncogenic family member of transcription factors (like ETS-1). The proto-oncogenes ETV1, ETV4 and ETV5 includes the most frequently rearranged and overexpressed genes in prostate cancer; ETV5 is a target for the Ras/Raf-1/MAPK pathway and can also be activated through the protein kinase A. Is overexpressed in metastatic human breast cancer cells and mouse mammary tumors and might therefore play an important role in mammary oncogenesis; up-regulation of ETV5 in B-CLL and MCL suggests this gene as a new candidate for the pathomechanism of B-cell lymphomas.	The protein encoded by this gene is part of the Mediator complex, which is involved in transcriptional coactivation of nearly all RNA polymerase II-dependent genes. Exome sequencing identified recurrent MED12 mutations in prostate cancer; MED12 suppression results in activation of TGF-βR signaling, which is both necessary and sufficient for drug resistance.
Role in Cancer	Oncogene	Potencial tumor suppressor
References	1: T'sas F, et al. Expression of the Ets transcription factor Erm is regulated through a conventional PKC signaling pathway in the Molt4 lymphoblastic cell line. FEBS Lett. 2005 Jan 3;579(1):66-70. PubMed PMID: 15620692. 2: Planagumà J, et al. Up-regulation of ERM/ETV5 correlates with the degree of myometrial infiltration in endometrioid endometrial carcinoma. J Pathol. 2005 Dec;207(4):422-9. PubMed PMID: 16175655. 3: Monge M, et al. ERM/ETV5 up-regulation plays a role during myometrial infiltration through matrix metalloproteinase-2 activation in endometrial cancer. Cancer Res. 2007 Jul 15;67(14):6753-9. PubMed PMID: 17638886. 4: Helgeson BE, et al. Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. Cancer Res. 2008 Jan 1;68(1):73-80. doi: 10.1158/0008-5472.CAN-07-5352. PubMed PMID: 18172298. 5: Firlej V, et al. Reduced tumorigenesis in mouse mammary cancer cells following inhibition of Pea3- or Erm-dependent transcription. J Cell Sci. 2008 Oct 15;121(Pt 20):3393-402. doi: 10.1242/jcs.027201. Epub 2008 Sep 30. PubMed PMID: 18827017. 6: Charfi C, et al. Gene profiling of Graffi murine leukemia virus-induced lymphoid leukemias: identification of leukemia markers and Fmn2 as a potential oncogene. Blood. 2011 Feb 10;117(6):1899-910. doi: 10.1182/blood-2010-10-311001. Epub 2010 Dec 6. PubMed PMID: 21135260. 7: Yu H, Zhang Y, Ye L, Jiang WG. The FERM family proteins in cancer invasion and metastasis. Front Biosci. 2011 Jan 1;16:1536-50. Review. PubMed PMID: 21196246. 8: Llauradó M, et al. ETV5 transcription factor is overexpressed in ovarian cancer and regulates cell adhesion in ovarian cancer cells. Int J Cancer. 2012 Apr 1;130(7):1532-43. doi: 10.1002/ijc.26148. Epub 2011 Aug 12. PubMed PMID: 21520040. 9: Vitari AC, et al. COP1 is a tumour suppressor that causes degradation of ETS transcription factors. Nature. 2011 May 15;474(7351):403-6. doi:10.1038/nature10005. PubMed PMID: 21572435. 10: Oh S, Shin S, Janknecht R. ETV1, 4 and 5: An oncogenic subfamily of ETS transcription factors. Biochim Biophys Acta. 2012 Aug;1826(1):1-12. doi:10.1016/j.bbcan.2012.02.002. Epub 2012 Mar 8. PubMed PMID: 22425584; PubMed Central PMCID: PMC3362686. 11: Korz C, et al. Evidence for distinct pathomechanisms in B-cell chronic lymphocytic leukemia and mantle cell lymphoma by quantitative expression analysis of cell cycle and apoptosis-associated genes. Blood. 2002 Jun 15;99(12):4554-61.	1: Huang S, Hölzel M, Knijnenburg T, Schlicker A, Roepman P, McDermott U, Garnett M, Grernrum W, Sun C, Prahallad A, Groenendijk FH, Mittempergher L, Nijkamp W, Neefjes J, Salazar R, Ten Dijke P, Uramoto H, Tanaka F, Beijersbergen RL, Wessels LF, Bernards R. MED12 controls the response to multiple cancer drugs through regulation of TGF-β receptor signaling. Cell. 2012 Nov 21;151(5):937-50. doi: 10.1016/j.cell.2012.10.035. 2: Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, White TA, Stojanov P, Van Allen E, Stransky N, Nickerson E, Chae SS, Boysen G, Auclair D, Onofrio RC, Park K, Kitabayashi N, MacDonald TY, Sheikh K, Vuong T, Guiducci C, Cibulskis K, Sivachenko A, Carter SL, Saksena G, Voet D, Hussain WM, Ramos AH, Winckler W, Redman MC, Ardlie K, Tewari AK, Mosquera JM, Rupp N, Wild PJ, Moch H, Morrissey C, Nelson PS, Kantoff PW, Gabriel SB, Golub TR, Meyerson M, Lander ES, Getz G, Rubin MA, Garraway LA. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. Nat Genet. 2012 May 20;44(6):685-9. doi: 10.1038/ng.2279.

Gene name	C18ORF1	PDE3B	TSPAN2	RBL2
Function	Uncharacterized protein C18orf12	Cyclic nucleotide phosphodiesterase with a dual-specificity for the second messengers cAMP and cGMP; promotes survival of CLL and T-ALL cells, and associates with resistance to cisplatin in head & neck carcinoma; is positively regulated by Akt; is positively involved in insulin signaling	Tetraspanin 2. Cell-surface proteins that are characterized by the presence of four hydrophobic domains. The proteins mediate signal transduction events that play a role in the regulation of cell development, activation, growth and motility. Tetraspanins protect MT1-MMP from lysosomal degradation and support its delivery to the cell surface. MT1-MMMP is a metalloproteinase that supports tumor cell invasion through extracellular matrix barriers containing fibrin, collagen, fibronectin, and other proteins.	pRb2/p130 belongs to the retinoblastoma (RB) family of proteins. pRb2/p130 acts as a tumor suppressor and it is a potent inhibitor of E2F-mediated trans-activation.
Role in Cancer	Not Determined	Oncogene	Supports tumor cell invasion	Potencial tumor suppressor
References		1: Moon E, Lee R, Near R, Weintraub L, Wolda S, Lerner A. Inhibition of PDE3B augments PDE4 inhibitor-induced apoptosis in a subset of patients with chronic lymphocytic leukemia. Clin Cancer Res. 2002 Feb;8(2):589-95. PubMed PMID: 11839681. 2: Yamano Y, Uzawa K, Saito K, Nakashima D, Kasamatsu A, Koike H, Kouzu Y, Shinozuka K, Nakatani K, Negoro K, Fujita S, Tanzawa H. Identification of cisplatin-resistance related genes in head and neck squamous cell carcinoma. Int J Cancer. 2010 Jan 15;126(2):437-49. doi: 10.1002/ijc.24704. PubMed PMID:19569180. 3: Dong H, Zitt C, Auriga C, Hatzelmann A, Epstein PM. Inhibition of PDE3, PDE4 and PDE7 potentiates glucocorticoid-induced apoptosis and overcomes glucocorticoid resistance in CEM T leukemic cells. Biochem Pharmacol. 2010 Feb 1;79(3):321-9. doi:10.1016/j.bcp.2009.09.001. Epub 2009 Sep 6. PubMed PMID: 19737543.	Mol Biol Cell. 2009 Apr;20(7):2030-40. doi: 10.1091/mbc.E08-11-1149. Epub 2009 Feb 11. Tetraspanin proteins regulate membrane type-1 matrix metalloproteinase-dependent pericellular proteolysis. Lafleur MA, Xu D, Hemler ME. SourceDepartment of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, Boston, MA 02115, USA.	Li Y, Graham C, Lacy S, Duncan AM, Whyte P. The adenovirus E1A-associated 130-kD protein is encoded by a member of the retinoblastoma gene family and physically interacts with cyclins A and E. Genes Dev. 1993 Dec;7(12A):2366-77.

Gene name	CAMK2D	SPRED1	ZNRF1	WIBG
Function	Calcium/calmodulin-dependent protein kinase II delta	Spred-1 is a member of the Sprouty family of proteins and is phosphorylated by tyrosine kinase in response to several growth factors, inhibits growth-factor-mediated. Legius syndrome is caused by germline loss-of-function SPRED1 mutations, resulting in overactivation of the RAS-MAPK signal transduction cascade. Loss of function mutations appear to predispose to leukemia.	E3 ubiquitin-protein ligase that mediates the ubiquitination of AKT1 and GLUL, thereby playing a role in neuron cells differentiation	Key regulator of the exon junction complex (EJC); Interferes with nonsense-mediated mRNA decay and enhances translation of spliced mRNAs, probably by antagonizing EJC functions.
Role in Cancer	Not Determined	Potencial tumor suppressor	Not Determined	Not Determined
References		1: Batz C, Hasle H, Bergsträsser E, van den Heuvel-Eibrink MM, Zecca M, Niemeyer CM, Flotho C; European Working Group of Myelodysplastic Syndromes in Childhood (EWOG-MDS). Does SPRED1 contribute to leukemogenesis in juvenile myelomonocytic leukemia (JMML)? Blood. 2010 Mar 25;115(12):2557-8. doi: 10.1182/blood-2009-12-260901. 2: Brems H, Pasmant E, Van Minkelen R, Wimmer K, Upadhyaya M, Legius E, Messiaen L. Review and update of SPRED1 mutations causing Legius syndrome. Hum Mutat. 2012 Nov;33(11):1538-46. doi: 10.1002/humu.22152.	web reference: http://www.uniprot.org/uniprot/Q8ND25	web reference: http://www.uniprot.org/uniprot/Q9BRP8

Gene name	DOCK10	AKAP2	TFP1	PPP1R16B
Function	Guanine nucleotide exchange factor (GEF). Dock10 expression is upregulated in B-lymphocytes and Chronic Lymphocytic Leukemia (CLL) cells in response to the cytokine IL-4; Dock10 is overexpressed in some aggressive papillary thyroid carcinomas; Activated Cdc42 mediated by DOCK10 induces a mesenchymal-amoeboid transition and increases cell invasion of melanoma cells.	A kinase anchoring protein 2, binds to the regulatory subunit of protein kinase A and is found associated with the actin cytoskeleton; AKAP2 plays a key role in CTR-mediated oncogenic actions by targeting cyclic AMP-dependent protein kinase PKA to CTR within a localized sub-region of the tight junctions complex.	Transferrin pseudogene	Regulator of protein phosphatase 1; acts as a positive regulator of pulmonary endothelial cell barrier function. May be a downstream target for TGF-beta1 signaling cascade in endothelial cells; in a large-scale insertional mutagenesis in Eμ-c-myc mice model pp1r16b was identified at common insertion sites and is bona fide cellular oncogene.
Role in Cancer	Oncogenic	Potential oncogene	Not Determined	Oncogene
References	1: Yelo E, Bernardo MV, Gimeno L, Alcaraz-García MJ, Majado MJ, Parrado A. Dock10, a novel CZH protein selectively induced by interleukin-4 in human B lymphocytes. Mol Immunol. 2008 Jul;45(12):3411-8. doi: 10.1016/j.molimm.2008.04.003. Epub 2008 May 21. PubMed PMID: 18499258. 2: Gadea G, Sanz-Moreno V, Self A, Godi A, Marshall CJ. DOCK10-mediated Cdc42 activation is necessary for amoeboid invasion of melanoma cells. Curr Biol. 2008 Oct 14;18(19):1456-65. doi: 10.1016/j.cub.2008.08.053. Epub 2008 Oct 2. PubMed PMID: 18835169. 3: Humtsoe JO, Koya E, Pham E, Aramoto T, Zuo J, Ishikawa T, Kramer RH. Transcriptional profiling identifies upregulated genes following induction of epithelial-mesenchymal transition in squamous carcinoma cells. Exp Cell Res. 2012 Feb 15;318(4):379-90. doi: 10.1016/j.yexcr.2011.11.011. Epub 2011 Nov 29. PubMed PMID: 22154512. 4: Fluge Ø, Bruland O, Akslen LA, Lillehaug JR, Varhaug JE. Gene expression in poorly differentiated papillary thyroid carcinomas. Thyroid. 2006 Feb;16(2):161-75. PubMed PMID: 16676402.	web reference: http://www.labome.org/grant/r01/ca/calcitonin/in/calcitonin-in-prostate-growth-and-neplasia-7737843.html		Mendrysa SM, Akagi K, Roayaei J, Lien WH, Copeland NG, Jenkins NA, Eisenman RN. An Integrated Genetic-Genomic Approach for the Identification of Novel Cancer Loci in Mice Sensitized to c-Myc-Induced Apoptosis. Genes Cancer. 2010 May;1(5):465-479.

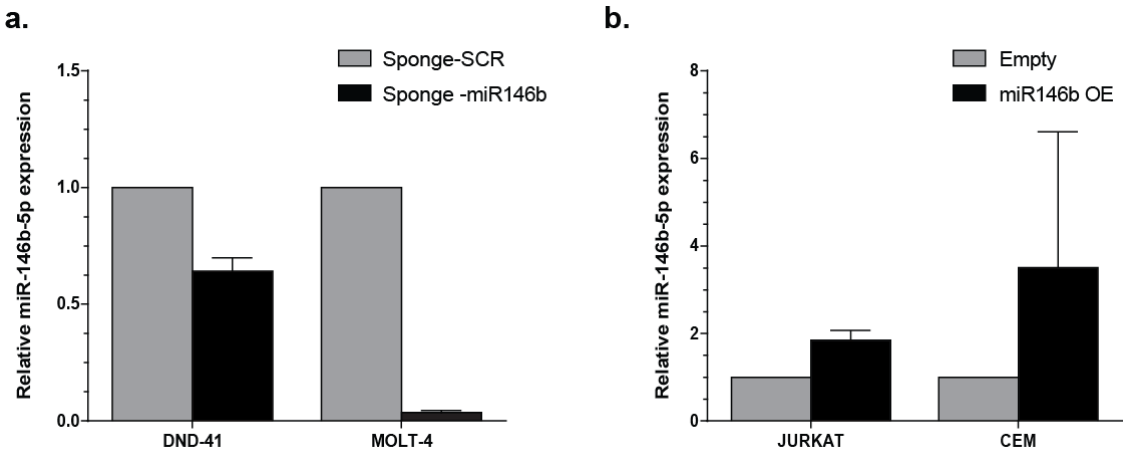
Gene name	TGFBR1	SPTBN1	LYSMD2	MGAT4
Function	Transforming growth factor beta receptor 1, involved in the regulation of cellular processes, including cell division, differentiation, motility, adhesion and death. A pro-oncogenic role for TGF-beta has been proposed. Once cells lose their sensitivity to TGF-beta1-mediated growth inhibition, autocrine TGF-beta signaling can promote tumorigenesis. Elevated levels of TGF-beta1 are often observed in advanced carcinomas, and have been correlated with increased tumor invasiveness and disease progression; TGFBR1*6A enhances the migration and invasion of MCF-7 breast cancer cells through RhoA activation; Oncogenic mutations in TGFBR1 were detected in several benign or malignant skin tumors.	Spectrin is an actin crosslinking and molecular scaffold protein that links the plasma membrane to the actin cytoskeleton; it contributes to contribute to platinum anticancer drug resistance in ovarian serous adenocarcinoma; is found as part of oncogenic fusions in atypical myeloproliferative disorder	LysM, putative peptidoglycan-binding, domain containing 2	Glycosyltransferase that regulates the formation of tri- and multiantennary branching structures in the Golgi apparatus; Oncogenic signaling in cancer cells up-regulates the transcription and activities of MGAT4. Oncogenic activation of PI3K and Erk/Ets increases Mgat4 expression and it's N-glycan products;
Role in Cancer	Potential oncogene	Involded in anticancer drug resistance	Not Determined	Potential oncogene
References	1: Arnault JP, Mateus C, Escudier B, Tomasic G, Wechsler J, Hollville E, Soria JC, Malka D, Sarasin A, Larcher M, André J, Kamsu-Kom N, Boussemart L, Lacroix L, Spatz A, Eggermont AM, Druillennec S, Vagner S, Eychène A, Dumaz N, Robert C. Skin tumors induced by sorafenib; paradoxic RAS-RAF pathway activation and oncogenic mutations of HRAS, TP53, and TGFBR1. Clin Cancer Res. 2012 Jan 1;18(1):263-72. doi: 10.1158/1078-0432.CCR-11-1344. Epub 2011 Nov 17. 2: Rosman DS, Phukan S, Huang CC, Pasche B. TGFBR1*6A enhances the migration and invasion of MCF-7 breast cancer cells through RhoA activation. Cancer Res. 2008 Mar 1;68(5):1319-28. doi: 10.1158/0008-5472.CAN-07-5424. 3: Bian Y, Terse A, Du J, Hall B, Molinolo A, Zhang P, Chen W, Flanders KC, Gutkind JS, Wakefield LM, Kulkarni AB. Progressive tumor formation in mice with conditional deletion of TGF-beta signaling in head and neck epithelia is associated with activation of the PI3K/Akt pathway. Cancer Res. 2009 Jul 15;69(14):5918-26. doi:10.1158/0008-5472.CAN-08-4623. Epub 2009 Jul 7.	1: Maeda O, Shibata K, Hosono S, Fujiwara S, Kajiyama H, Ino K, Nawa A, Tamakoshi K, Kikkawa F. Spectrin αII and βII tetramers contribute to platinum anticancer drug resistance in ovarian serous adenocarcinoma. Int J Cancer. 2012 Jan 1;130(1):113-21. doi: 10.1002/ijc.25983. Epub 2011 Apr 25. PubMed PMID: 21328338. 2: Gallagher G, Horsman DE, Tsang P, Forrest DL. Fusion of PRKG2 and SPTBN1 to the platelet-derived growth factor receptor beta gene (PDGFRB) in imatinib-responsive atypical myeloproliferative disorders. Cancer Genet Cytogenet. 2008 Feb;181(1):46-51. doi:10.1016/j.cancergencyto.2007.10.021. PubMed PMID: 18262053. 3: Grand FH, Iqbal S, Zhang L, Russell NH, Chase A, Cross NC. A constitutively active SPTBN1-FLT3 fusion in atypical chronic myeloid leukemia is sensitive to tyrosine kinase inhibitors and immunotherapy. Exp Hematol. 2007 Nov;35(11):1723-7. Epub 2007 Aug 30. PubMed PMID: 17764812.		1: Lau KS, Dennis JW. N-Glycans in cancer progression. Glycobiology. 2008 Oct;18(10):750-60. doi: 10.1093/glycob/cwn071.

Supplementary Table 5 – List of known TAL1 down-regulated genes that are potentially targeted by TAL1 up-regulated microRNAs, as in Figure 3.2.5b. Described function(s), putative role in cancer and corresponding relevant references are presented for each gene. For gene assessment details see supplemental methods.

Gene name	SKI	IKZF3	DOLPP1	SH2D3C
Function	Nuclear protooncoprotein; negatively regulates TGFb; A dual role of SKI has been observed in different malignancies. The human SKI gene is located at chromosome 1p36, a potential tumor suppressor locus that is frequently deleted in various human cancers including neuroblastoma, melanoma, colorectal carcinoma and leukemia.	AILOS Ikaros family of zinc-finger proteins; transcription factors involved in the regulation of lymphocyte development; Ikaros inactivation is a recurrent event in human T-ALL; it's also is also a major tumor suppressor in human B-ALL.	Dolichyl pyrophosphate (Dol-P-P) phosphatase, Required for efficient N-glycosylation	SH2 domain containing 3C, adaptor protein involved in cell migration.
Role in Cancer	Potencial tumor suppressor	Tumor suppressor	Not determined	Not determined
References	1:Shinagawa T, Nomura T, Colmenares C, Ohira M, Nakagawara A, Ishii S. Increased susceptibility to tumorigenesis of ski-deficient heterozygous mice. Oncogene. 2001 Dec 6;20(56):8100-8.. 2: Colmenares C, Heilstedt HA, Shaffer LG, Schwartz S, Berk M, Murray JC, Stavnezer E: Loss of the SKI proto-oncogene in individuals affected with 1p36 deletion syndrome is predicted by strain-dependent defects in Ski-/- mice.Nat Genet 2002, 30:106-109. 3: Wang P, Chen Z, Meng ZQ, Fan J, Luo JM, Liang W, Lin JH, Zhou ZH, Chen H, Wang K, Shen YH, Xu ZD, Liu LM. Dual role of Ski in pancreatic cancer cells: tumor-promoting versus metastasis-suppressive function. Carcinogenesis. 2009 Sep;30(9):1497-506. doi: 10.1093/carcin/bgp154. Epub 2009 Jun 22.	1: Philippe Kastner and Susan Chan. Role of Ikaros in T-cell acute lymphoblastic leukemia. World J Biol Chem. 2011 June 26; 2(6): 108–114. 2: Nakase K, Ishimaru F, Avitahl N, Dansako H, Matsuo K, Fujii K, Sezaki N, Nakayama H, Yano T, Fukuda S, Imajoh K, Takeuchi M, Miyata A, Hara M, Yasukawa M, Takahashi I, Taguchi H, Matsue K, Nakao S, Niho Y, Takenaka K, Shinagawa K, Ikeda K, Niiya K, Harada M. Dominant negative isoform of the Ikaros gene in patients with adult B-cell acute lymphoblastic leukemia. Cancer Res. 2000 Aug 1;60(15):4062-5.		

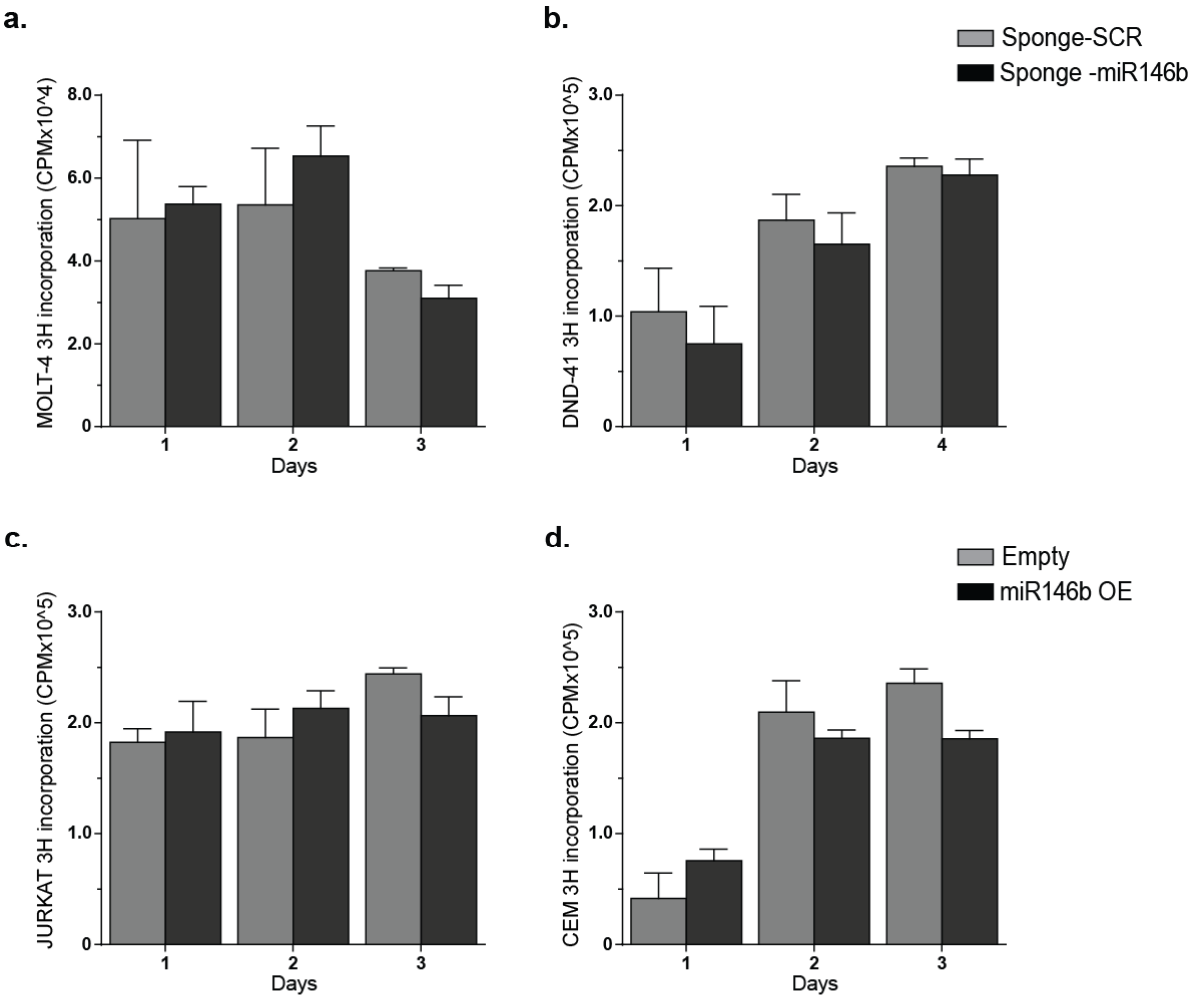
Gene name	SRGAP3	TOX	IGF1R
Function	Member of the Slit-Robo sub-family of Rho GTPase-activating proteins implicated in repulsive axon guidance and neuronal migration through Slit-Robo-mediated signal transduction; has tumor suppressor-like activity in HMECs, likely through its activity as a negative regulator of Rac1	Thymocyte selection-associated high mobility group box protein; it's required for T-Cell diferentiation; TAL1 binds to Tox genomic region in leukaemic T cells,and negatively regulates it's expression	The Insulin-like Growth Factor 1 Receptor is a transmembrane tyrosine kinase receptor that is activated by IGF-1 and IGF-2.
Role in Cancer	Tumor suppressor	Promotes T-Cell diferentiation	Oncogene
References	Lahoz A, Hall A. A tumor suppressor role for srGAP3 in mammary epithelial cells. Oncogene. 2012 Oct 29. doi: 10.1038/onc.2012.489	1: Aliahmad P, Kaye J (2008) Development of all CD4 T lineages requires nuclear factor TOX. J Exp Med 205: 245–256. 2: Palii CG, Perez-Iratxeta C, Yao Z, Cao Y, Dai F, Davison J, Atkins H, Allan D, Dilworth FJ, Gentleman R, Tapscott SJ, Brand M. Differential genomic targeting of the transcription factor TAL1 in alternate haematopoietic lineages. EMBO J. 2011 Feb 2;30(3):494-509. doi: 10.1038/emboj.2010.342. Epub 2010 Dec 21.	Medyouf H, Gusscott S, Wang H, Tseng JC, Wai C, Nemirovsky O, Trumpp A, Pflumio F, Carboni J, Gottardis M, Pollak M, Kung AL, Aster JC,Holzenberger M, Weng AP. High-level IGF1R expression is required for leukemia-initiating cell activity in T-ALL and is supported by Notch signaling. J Exp Med. 2011 Aug 29;208(9):1809-22. doi: 10.1084/jem.20110121. Epub 2011 Aug 1.

Gene name	LRP12	ID1	MX2	BTG2
Function	Low-density lipoprotein receptor; The level of this protein was found to be lower in tumor derived cell lines compared to normal cells; may act as tumor suppressor	Although it does not bind directly to DNA, by binding basic helix-loop-helix transcription factors through its HLH motif, ID1 may control tissue-specific genes related to cell growth, proliferation, differentiation and angiogenesis; ID1 is a common downstream target of oncogenic tyrosine kinases in leukemic cells;	Interferon-induced GTP-binding protein; is upregulated by interferon-alpha	TG family member 2 or NGF-inducible anti-proliferative protein PC3; has been shown to inhibit medulloblastoma, by inhibiting the proliferation and triggering the diffentiation of the precursors of cerebellar granule neurons. It is involved in the regulation of the G1/S transition of the cell cycle. Btg2 act as transcriptional cofactor of the Hoxb9 protein, and suggest that this interaction may mediate it's antiproliferative function.
Role in Cancer	Potencial tumor suppressor	Oncogene	Not determined	Tumor suppressor
References	web reference: http://refgene.com/gene/29967	1: Tam WF, Gu TL, Chen J, Lee BH, Bullinger L, Fröhling S, Wang A, Monti S, Golub TR, Gilliland DG. Id1 is a common downstream target of oncogenic tyrosine kinases in leukemic cells. Blood. 2008 Sep 1;112(5):1981-92. doi: 10.1182/blood-2007-07-103010. 2: Perk J, Iavarone A, Benezra R. Id family of helix-loop-helix proteins in cancer. Nat Rev Cancer. 2005 Aug;5(8):603-14.		1: Farioli-Vecchioli S, Tanori M, Micheli L, Mancuso M, Leonardi L, Saran A, Ciotti MT, Ferretti E, Gulino A, Pazzaglia S, Tirone F. Inhibition of medulloblastoma tumorigenesis by the antiproliferative and pro-differentiative gene PC3. FASEB J. 2007 Jul;21(9):2215-25. Epub 2007 Mar 19. 2: Prévôt D, Voeltzel T, Birot AM, Morel AP, Rostan MC, Magaud JP, Corbo L. The leukemia-associated protein Btg1 and the p53-regulated protein Btg2 interact with the homeoprotein Hoxb9 and enhance its transcriptional activation. J Biol Chem. 2000 Jan 7;275(1):147-53



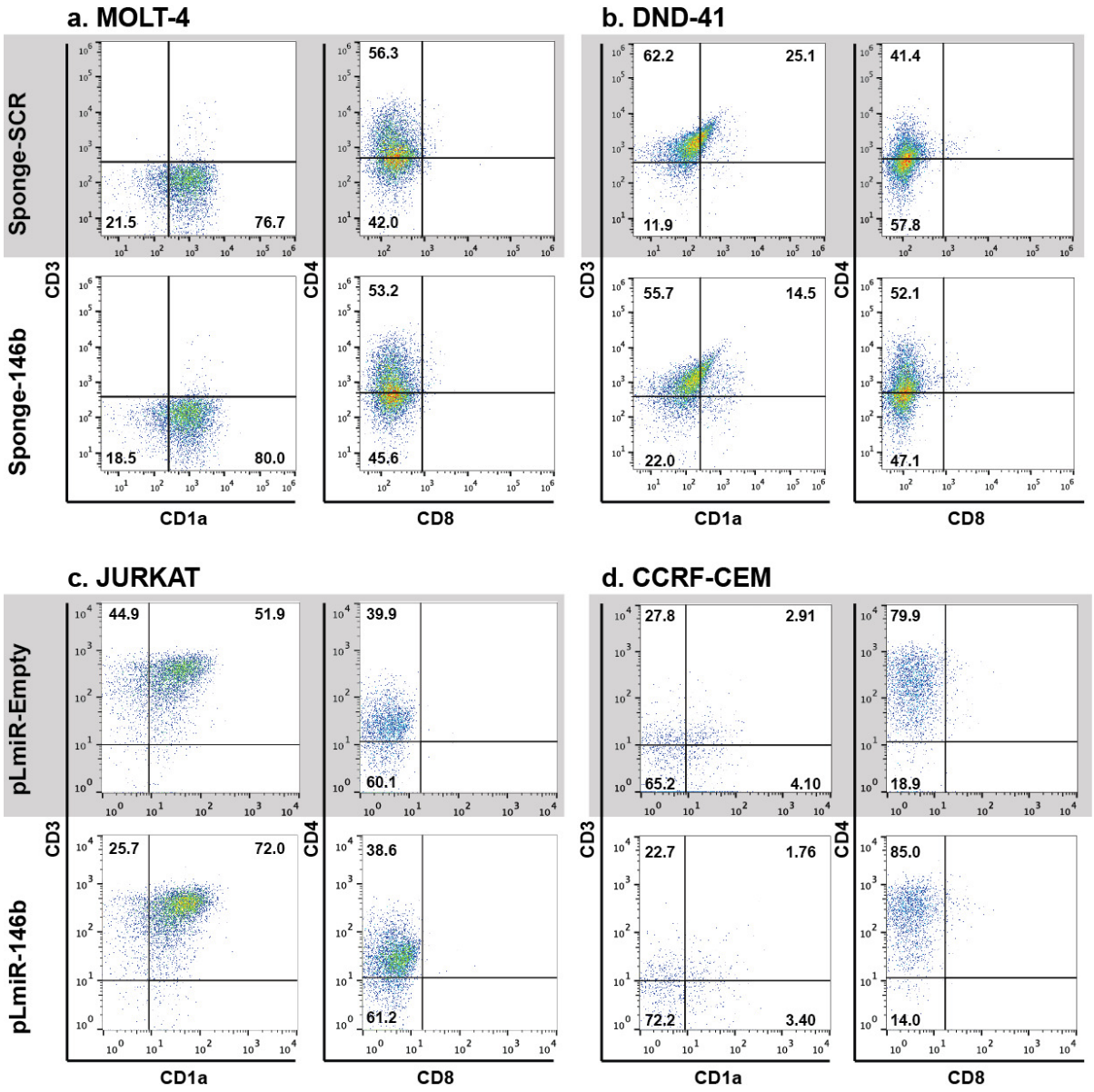
Supplementary Figure 4 – miR-146b-5p expression in transduced T-ALL cell lines.

a) Cell lines transduced with sponge-vector express lower levels of mir-146b-5p. TAL1-negative T-ALL cell lines (DND-41 and MOLT-4) where stable transduced to inhibit human miR-146b-5p. **b)** Cell lines transduced with pLemir- vector over-express (OE) miR-146b-5p. TAL1-positive T-ALL cell lines (JURKAT and CCRF-CEM) were stably transduced with a lentiviral vector to ectopically express the human pre-miR-146b. miR-146b-5p levels were accessed by qRT-PCR and normalized to *SNORD38*. Values indicate the mean ± lower and upper limit of three technical replicates relatively to the mock transduction (SCR or Empty).



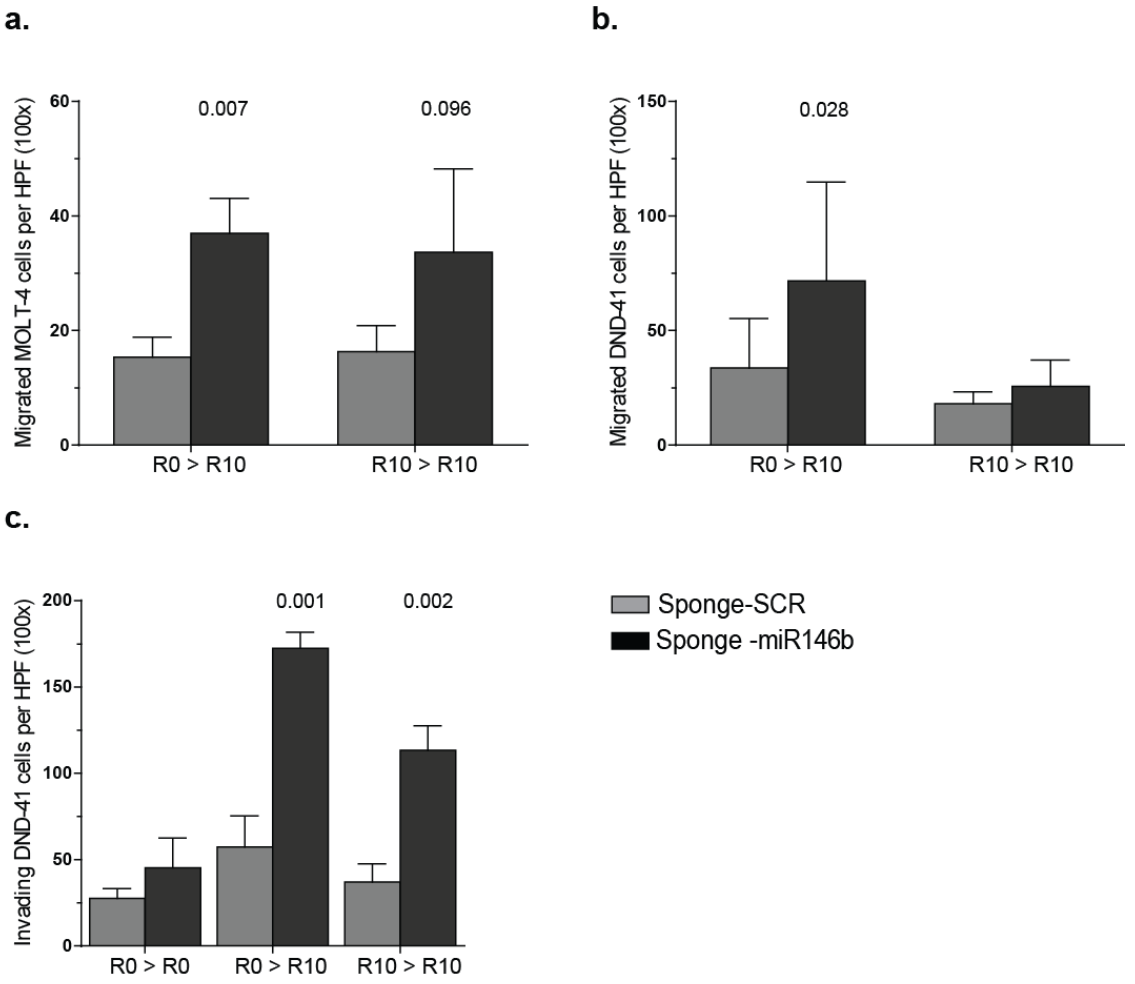
Supplementary Figure 5 – ^{3}H-Thymidine incorporation assays upon modulation of miR-146b

Proliferation of T-ALL cell lines was assessed by ^{3}H thymidine incorporation determined at several time points. MOLT-4 **a)** and DND-41 **b)** cell lines expressing lower levels of miR-146b-5p (Sponge-146b) are compared to the mock transduced cell lines (Sponge-SCR). JURKAT **c)** and CCRF-CEM **d)** cell lines ectopically expressing miR-146b (146b-OE) are compared to the mock transduced cell lines (Empty). The bars graphs represent the mean (+/- SEM) of three technical replicates.



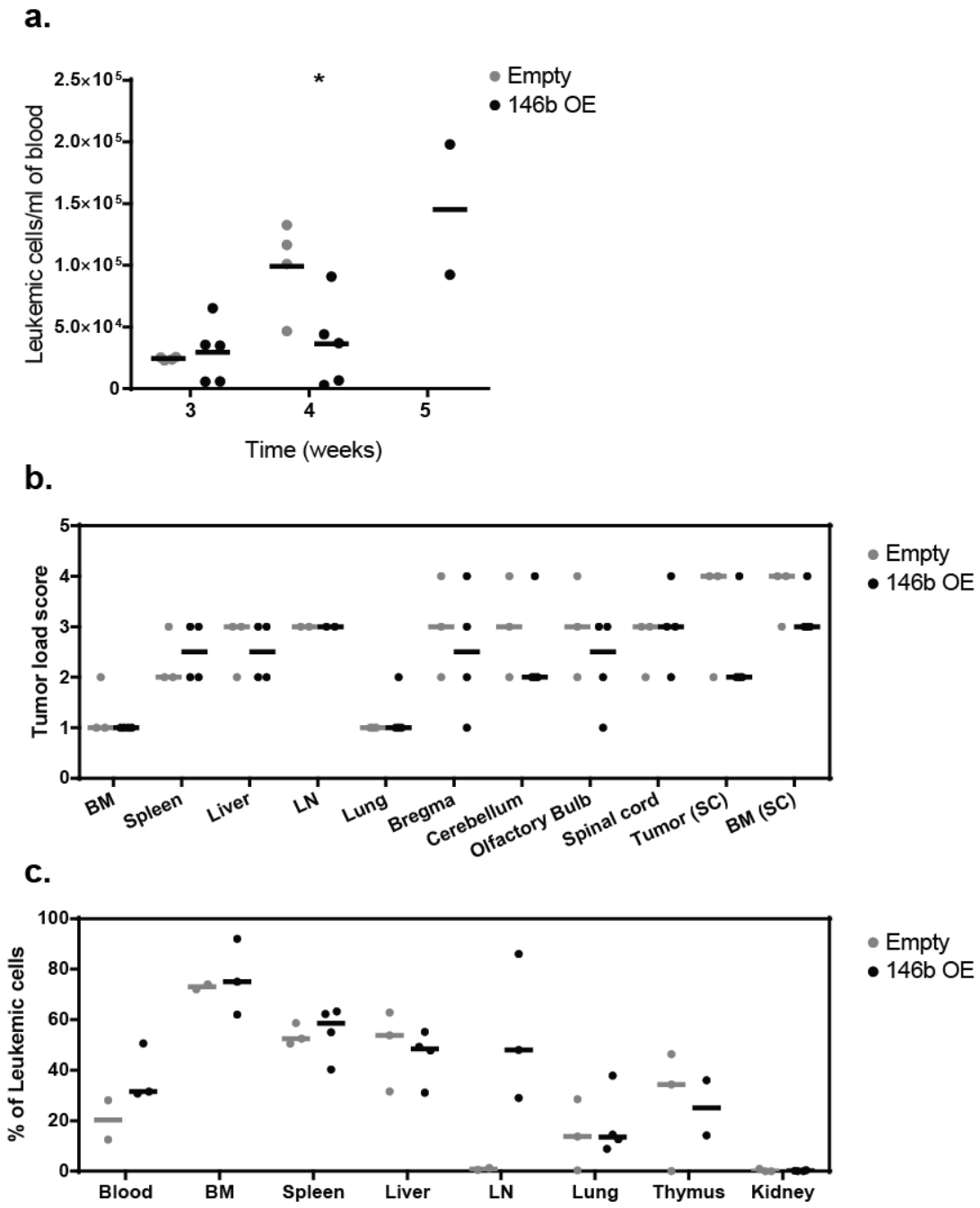
Supplementary Figure 6 – Phenotype of transduced T-ALL cell lines.

The expression of surface markers CD1a, CD3, CD4 and CD8 was accessed by flow cytometry (FACS) in the transduced T-ALL cell lines in order to determine their immunological phenotype. The cell lines presented were stably transduced for more than two months. Here are shown the FACS plots of cell lines transduced with sponge-146b or the mock vector, **a)** MOLT-4 and **b)** DND-41, and also the cell lines transduced with pLmiR-146b and correspondent control vector, **c)** JURKAT and **d)** CCRF-CEM.



Supplementary Figure 7 – Migration and invasion assays upon down-regulation of miR-146b-5p levels in T-ALL cell lines.

The migration capacities of cells with down-regulation of miR-146b-5p expression **a)** MOLT-4 and **b)** DND-41 were assessed through transwell migration assays. The invasion capacities of DND-41 **c)** were assessed by matrigel coated transwell assays. In both kinds of experiments serum was used as chemoattractant. A total of 100e3 cells were plated on the upper chamber of the transwell in culture medium (RPMI) either in the absence of serum (R0) or in the presence of 10% serum containing RPMI medium (R10). The number of migrated or invading cells per high-power field (HPF ×100) was determined by the average number of cells counted in five no-overlapping microscope fields of the same transwell. The number of migrated or invading cells was counted in three transwells per condition in each experiment. The graphs depict the average number of cells per HPF counted in at least three independent migration experiments of MOLT-4 **a)** and DND-41 **b)** cells expressing lower levels of miR-146b-5p (sponge-146b) and of mock transduced cells (sponge-SCR). These represent the values used to calculate the migration index of Figure 3.3.4a and 4c. The graph **c)** depicts another representative experiment of Invasion by showing the average number of DND-41 cells that migrated in three matrigel coated transwells. Error bars represent SD and values depicted represent the p-values calculated using a two-tailed Student's t-test. The symbol > represents the direction of the migration, ie, R0>R10 means cells migrated from RPMI with 0% serum to RPMI with 10% serum.



Supplementary Figure 8 - MiR-146b in the development of T-ALL *in vivo*.

a) Absolute number of human leukemic cells per ml of mice peripheral blood determined weekly by flow cytometry evaluation of RFP+ cells. The lines represent the mean for each time point. p-value calculated using a two-tailed Student's t-test (*p<0.05). **b)** Tumor load score determined by hematoxilin/eosin staining of histological sections of several organs. **c)** Flow cytometry analysis of single-cell suspensions of several organs. The lines in **b)** and **c)** represent the median value for each organ. BM – bone-marrow; LN – lymph nodes; SC – spinal cord.

REFERENCES

1. Huret, J.L., et al., *Atlas of genetics and cytogenetics in oncology and haematology in 2013.* Nucleic Acids Res, 2013. **41**(Database issue): p. D920-4.
2. Orkin, S.H. and L.I. Zon, *Hematopoiesis: an evolving paradigm for stem cell biology.* Cell, 2008. **132**(4): p. 631-44.
3. Wiemels, J., *Perspectives on the causes of childhood leukemia.* Chem Biol Interact, 2012. **196**(3): p. 59-67.
4. Greaves, M., *Infection, immune responses and the aetiology of childhood leukaemia.* Nat Rev Cancer, 2006. **6**(3): p. 193-203.
5. Kaatsch, P. and A. Mergenthaler, *Incidence, time trends and regional variation of childhood leukaemia in Germany and Europe.* Radiat Prot Dosimetry, 2008. **132**(2): p. 107-13.
6. Pui, C.H., M.V. Relling, and J.R. Downing, *Acute lymphoblastic leukemia.* N Engl J Med, 2004. **350**(15): p. 1535-48.
7. Oeffinger, K.C., et al., *Chronic Health Conditions in Adult Survivors of Childhood Cancer.* New England Journal of Medicine, 2006. **355**(15): p. 1572-1582.
8. Pui, C.H. and W.E. Evans, *A 50-year journey to cure childhood acute lymphoblastic leukemia.* Semin Hematol, 2013. **50**(3): p. 185-96.
9. Pui, C.H., et al., *Biology, risk stratification, and therapy of pediatric acute leukemias: an update.* J Clin Oncol, 2011. **29**(5): p. 551-65.
10. Mori, H., et al., *Chromosome translocations and covert leukemic clones are generated during normal fetal development.* Proc Natl Acad Sci U S A, 2002. **99**(12): p. 8242-7.
11. Mullighan, C.G., et al., *Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia.* Science, 2008. **322**(5906): p. 1377-80.
12. Lo Nigro, L., *Biology of childhood acute lymphoblastic leukemia.* J Pediatr Hematol Oncol, 2013. **35**(4): p. 245-52.
13. Mullighan, C.G., et al., *Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia.* Nature, 2007. **446**(7137): p. 758-64.
14. Stanulla, M. and M. Schrappe, *Treatment of childhood acute lymphoblastic leukemia.* Semin Hematol, 2009. **46**(1): p. 52-63.
15. DeAngelo, D.J., *The treatment of adolescents and young adults with acute lymphoblastic leukemia.* Hematology Am Soc Hematol Educ Program, 2005: p. 123-30.
16. Goldberg, J.M., et al., *Childhood T-cell acute lymphoblastic leukemia: the Dana-Farber Cancer Institute acute lymphoblastic leukemia consortium experience.* J Clin Oncol, 2003. **21**(19): p. 3616-22.
17. Graux, C., et al., *Cytogenetics and molecular genetics of T-cell acute lymphoblastic leukemia: from thymocyte to lymphoblast.* Leukemia, 2006. **20**(9): p. 1496-510.
18. Bene, M.C., et al., *Proposals for the immunological classification of acute leukemias. European Group for the Immunological Characterization of Leukemias (EGIL).* Leukemia, 1995. **9**(10): p. 1783-6.
19. Szczepański, T., V.H.J. van der Velden, and J.J.M. van Dongen, *Classification systems for acute and chronic leukaemias.* Best Practice & Research Clinical Haematology, 2003. **16**(4): p. 561-582.
20. Akashi, K., et al., *Lymphoid precursors.* Current Opinion in Immunology, 2000. **12**(2): p. 144-150.
21. Orkin, S.H., *Diversification of haematopoietic stem cells to specific lineages.* Nat Rev Genet, 2000. **1**(1): p. 57-64.
22. Aifantis, I., E. Raetz, and S. Buonamici, *Molecular pathogenesis of T-cell leukaemia and lymphoma.* Nat Rev Immunol, 2008. **8**(5): p. 380-90.
23. Martins, V.C., et al., *Cell competition is a tumour suppressor mechanism in the thymus.* Nature, 2014. **509**(7501): p. 465-70.
24. Staal, F.J., et al., *Transcriptional control of t lymphocyte differentiation.* Stem Cells, 2001. **19**(3): p. 165-79.

25. Kruisbeek, A.M., et al., *Branching out to gain control: how the pre-TCR is linked to multiple functions.* Immunology Today, 2000. **21**(12): p. 637-644.
26. Bommhardt, U., et al., *Molecular and cellular mechanisms of T cell development.* Cell Mol Life Sci, 2004. **61**(3): p. 263-80.
27. Plum, J.D.S.M.L.G.V.B.V.B., *Interleukin-7 is a critical growth factor in early human T-cell development.* Blood, 1996. **88**(11): p. 4239-4245.
28. Jaleco, A.C., et al., *Differential effects of Notch ligands Delta-1 and Jagged-1 in human lymphoid differentiation.* J Exp Med, 2001. **194**(7): p. 991-1002.
29. Petrie, H.T. and J.C. Zuniga-Pflucker, *Zoned out: functional mapping of stromal signaling microenvironments in the thymus.* Annu Rev Immunol, 2007. **25**: p. 649-79.
30. Ferrando, A.A., et al., *Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia.* Cancer Cell, 2002. **1**(1): p. 75-87.
31. Homminga, I., et al., *Integrated Transcript and Genome Analyses Reveal NKX2-1 and MEF2C as Potential Oncogenes in T Cell Acute Lymphoblastic Leukemia.* Cancer Cell, 2011. **19**(4): p. 484-497.
32. Zhang, J., et al., *The genetic basis of early T-cell precursor acute lymphoblastic leukaemia.* Nature, 2012. **481**(7380): p. 157-63.
33. Coustan-Smith, E., et al., *Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia.* The Lancet Oncology, 2009. **10**(2): p. 147-156.
34. Asnafi, V., et al., *Analysis of TCR, pT alpha, and RAG-1 in T-acute lymphoblastic leukemias improves understanding of early human T-lymphoid lineage commitment.* Blood, 2003. **101**(7): p. 2693-703.
35. Cauwelier, B., et al., *Molecular cytogenetic study of 126 unselected T-ALL cases reveals high incidence of TCRbeta locus rearrangements and putative new T-cell oncogenes.* Leukemia, 2006. **20**(7): p. 1238-44.
36. Atak, Z.K., et al., *Comprehensive analysis of transcriptome variation uncovers known and novel driver events in T-cell acute lymphoblastic leukemia.* PLoS Genet, 2013. **9**(12): p. e1003997.
37. Le Noir, S., et al., *Extensive molecular mapping of TCRalpha/delta- and TCRbeta-involved chromosomal translocations reveals distinct mechanisms of oncogene activation in T-ALL.* Blood, 2012. **120**(16): p. 3298-309.
38. Hayette, S., et al., *Recurrent involvement of the MLL gene in adult T-lineage acute lymphoblastic leukemia.* Blood, 2002. **99**(12): p. 4647-9.
39. Asnafi, V.R.-W.I.D.N.B.C.L.D.C.C.G.R.L.-P.M.D.E.B.A.T.X.M.E., *CALM-AF10 is a common fusion transcript in T-ALL and is specific to the TCRγδ lineage.* Blood, 2003. **102**(3): p. 1000-1006.
40. Graux, C., et al., *Fusion of NUP214 to ABL1 on amplified episomes in T-cell acute lymphoblastic leukemia.* Nat Genet, 2004. **36**(10): p. 1084-9.
41. Van Vlierberghe, P.v.G.M.B.H.B.L.C.H.T.B.-G.J.P.M.v.W.E.R.V.A.J.P.K.W.A., *The cryptic chromosomal deletion del(11)(p12p13) as a new activation mechanism of LMO2 in pediatric T-cell acute lymphoblastic leukemia.* Blood, 2006. **108**(10): p. 3520-3529.
42. Lahortiga, I., et al., *Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia.* Nat Genet, 2007. **39**(5): p. 593-5.
43. Hebert, J.C.J.M.B.J.S.F., *Candidate tumor-suppressor genes MTS1 (p16INK4A) and MTS2 (p15INK4B) display frequent homozygous deletions in primary cells from T- but not from B-cell lineage acute lymphoblastic leukemias [see comments].* Blood, 1994. **84**(12): p. 4038-4044.
44. Weng, A.P., et al., *Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia.* Science, 2004. **306**(5694): p. 269-71.
45. Zenatti, P.P., et al., *Oncogenic IL7R gain-of-function mutations in childhood T-cell acute lymphoblastic leukemia.* Nat Genet, 2011. **43**(10): p. 932-9.
46. Bains, T., et al., *Newly described activating JAK3 mutations in T-cell acute lymphoblastic leukemia.* Leukemia, 2012. **26**(9): p. 2144-6.

47. De Keersmaecker, K., et al., *Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia.* Nat Genet, 2013. **45**(2): p. 186-90.
48. Ntziachristos, P., et al., *Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia.* Nat Med, 2012. **18**(2): p. 298-301.
49. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.
50. Sherr, C.J., *The INK4a/ARF network in tumour suppression.* Nat Rev Mol Cell Biol, 2001. **2**(10): p. 731-7.
51. Omura-Minamisawa, M., et al., *Universal Inactivation of Both p16 and p15 but not Downstream Components Is an Essential Event in the Pathogenesis of T-Cell Acute Lymphoblastic Leukemia.* Clinical Cancer Research, 2000. **6**(4): p. 1219-1228.
52. Ramakers-van Woerden, N.L., et al., *In vitro drug resistance and prognostic impact of p16INK4A/P15INK4B deletions in childhood T-cell acute lymphoblastic leukaemia.* Br J Haematol, 2001. **112**(3): p. 680-90.
53. Clappier, E., et al., *Cyclin D2 dysregulation by chromosomal translocations to TCR loci in T-cell acute lymphoblastic leukemias.* Leukemia, 2006. **20**(1): p. 82-6.
54. Cardoso, B.A., et al., *Aberrant signaling in T-cell acute lymphoblastic leukemia: biological and therapeutic implications.* Braz J Med Biol Res, 2008. **41**(5): p. 344-50.
55. Barata, J.T., et al., *Activation of PI3K is indispensable for interleukin 7-mediated viability, proliferation, glucose use, and growth of T cell acute lymphoblastic leukemia cells.* J Exp Med, 2004. **200**(5): p. 659-69.
56. Van Limbergen, H., et al., *Molecular cytogenetic and clinical findings in ETV6/ABL1-positive leukemia.* Genes Chromosomes Cancer, 2001. **30**(3): p. 274-82.
57. De Keersmaecker, K., et al., *Fusion of EML1 to ABL1 in T-cell acute lymphoblastic leukemia with cryptic t(9;14)(q34;q32).* Blood, 2005. **105**(12): p. 4849-52.
58. Burnett, R.C., et al., *Molecular analysis of the T-cell acute lymphoblastic leukemia-associated t(1;7)(p34;q34) that fuses LCK and TCRB.* Blood, 1994. **84**(4): p. 1232-6.
59. Carron, C., et al., *TEL-JAK2 transgenic mice develop T-cell leukemia.* Blood, 2000. **95**(12): p. 3891-9.
60. Flex, E., et al., *Somatically acquired JAK1 mutations in adult acute lymphoblastic leukemia.* J Exp Med, 2008. **205**(4): p. 751-8.
61. Jeong, E.G., et al., *Somatic mutations of JAK1 and JAK3 in acute leukemias and solid cancers.* Clin Cancer Res, 2008. **14**(12): p. 3716-21.
62. Shochat, C., et al., *Gain-of-function mutations in interleukin-7 receptor-alpha (IL7R) in childhood acute lymphoblastic leukemias.* J Exp Med, 2011. **208**(5): p. 901-8.
63. von Lintig, F.C., et al., *Ras Activation in Normal White Blood Cells and Childhood Acute Lymphoblastic Leukemia.* Clinical Cancer Research, 2000. **6**(5): p. 1804-1810.
64. Yokota, S., et al., *Mutational analysis of the N-ras gene in acute lymphoblastic leukemia: a study of 125 Japanese pediatric cases.* Int J Hematol, 1998. **67**(4): p. 379-87.
65. Ciofani, M. and J.C. Zuniga-Pflucker, *Notch promotes survival of pre-T cells at the beta-selection checkpoint by regulating cellular metabolism.* Nat Immunol, 2005. **6**(9): p. 881-8.
66. Palomero, T., et al., *NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth.* Proc Natl Acad Sci U S A, 2006. **103**(48): p. 18261-6.
67. Weng, A.P., et al., *c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma.* Genes Dev, 2006. **20**(15): p. 2096-109.
68. Sharma, V.M., et al., *Notch1 contributes to mouse T-cell leukemia by directly inducing the expression of c-myc.* Mol Cell Biol, 2006. **26**(21): p. 8022-31.
69. Vilimas, T., et al., *Targeting the NF-kappaB signaling pathway in Notch1-induced T-cell leukemia.* Nat Med, 2007. **13**(1): p. 70-7.

70. Palomero, T., et al., *Mutational loss of PTEN induces resistance to NOTCH1 inhibition in T-cell leukemia.* Nat Med, 2007. **13**(10): p. 1203-10.
71. Gonzalez-Garcia, S., et al., *CSL-MAML-dependent Notch1 signaling controls T lineage-specific IL-7R{alpha} gene expression in early human thymopoiesis and leukemia.* J Exp Med, 2009. **206**(4): p. 779-91.
72. O'Neil, J., et al., *FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors.* J Exp Med, 2007. **204**(8): p. 1813-24.
73. Thompson, B.J., et al., *The SCFFBW7 ubiquitin ligase complex as a tumor suppressor in T cell leukemia.* J Exp Med, 2007. **204**(8): p. 1825-35.
74. Cantley, L.C., *The phosphoinositide 3-kinase pathway.* Science, 2002. **296**(5573): p. 1655-7.
75. Gutierrez, A.S.T.G.R.C.A.S.L.A.Y.D.S.N.D.M.L.A.W.S.S.L.R.Z.J.P.A.C.L., *High frequency of PTEN, PI3K, and AKT abnormalities in T-cell acute lymphoblastic leukemia.* Blood, 2009. **114**(3): p. 647-650.
76. Silva, A., et al., *PTEN posttranslational inactivation and hyperactivation of the PI3K/Akt pathway sustain primary T cell leukemia viability.* J Clin Invest, 2008. **118**(11): p. 3762-74.
77. Avellino, R., et al., *Rapamycin stimulates apoptosis of childhood acute lymphoblastic leukemia cells.* Blood, 2005. **106**(4): p. 1400-6.
78. Uddin, S., et al., *Inhibition of phosphatidylinositol 3'-kinase induces preferentially killing of PTEN-null T leukemias through AKT pathway.* Biochemical and Biophysical Research Communications, 2004. **320**(3): p. 932-938.
79. Van Vlierberghe, P., et al., *PHF6 mutations in T-cell acute lymphoblastic leukemia.* Nat Genet, 2010. **42**(4): p. 338-342.
80. Ferrando, A.A.H.S.P.T.H.M.H.T.F.E.A.L.A.T., *Biallelic transcriptional activation of oncogenic transcription factors in T-cell acute lymphoblastic leukemia.* Blood, 2003. **103**(5): p. 1909-1911.
81. Jones, S., *An overview of the basic helix-loop-helix proteins.* Genome Biol, 2004. **5**(6): p. 226.
82. Tremblay, M., et al., *Regulation of pT alpha gene expression by a dosage of E2A, HEB, and SCL.* J Biol Chem, 2003. **278**(15): p. 12680-7.
83. Bain, G., et al., *E2A deficiency leads to abnormalities in alphabeta T-cell development and to rapid development of T-cell lymphomas.* Mol Cell Biol, 1997. **17**(8): p. 4782-91.
84. Engel, I. and C. Murre, *E2A proteins enforce a proliferation checkpoint in developing thymocytes.* EMBO J, 2004. **23**(1): p. 202-11.
85. Engel, I. and C. Murre, *Ectopic expression of E47 or E12 promotes the death of E2A-deficient lymphomas.* Proc Natl Acad Sci U S A, 1999. **96**(3): p. 996-1001.
86. Kim, D., X.C. Peng, and X.H. Sun, *Massive apoptosis of thymocytes in T-cell-deficient Id1 transgenic mice.* Mol Cell Biol, 1999. **19**(12): p. 8240-53.
87. Baer, R., *TAL1, TAL2 and LYL1: a family of basic helix-loop-helix proteins implicated in T cell acute leukaemia.* Semin Cancer Biol, 1993. **4**(6): p. 341-7.
88. Mellentin, J.D., S.D. Smith, and M.L. Cleary, *lyl-1, a novel gene altered by chromosomal translocation in T cell leukemia, codes for a protein with a helix-loop-helix DNA binding motif.* Cell, 1989. **58**(1): p. 77-83.
89. Zhong, Y., et al., *Overexpression of a transcription factor LYL1 induces T- and B-cell lymphoma in mice.* Oncogene, 2007. **26**(48): p. 6937-47.
90. Xia, Y., et al., *Products of the TAL2 oncogene in leukemic T cells: bHLH phosphoproteins with DNA-binding activity.* Oncogene, 1994. **9**(5): p. 1437-46.
91. Wang, J., et al., *The t(14;21)(q11.2;q22) chromosomal translocation associated with T-cell acute lymphoblastic leukemia activates the BHLHB1 gene.* Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3497-502.

92. Larson, R.C., et al., *Protein dimerization between Lmo2 (Rbtn2) and Tal1 alters thymocyte development and potentiates T cell tumorigenesis in transgenic mice.* EMBO J, 1996. **15**(5): p. 1021-7.
93. Wadman, I.A., et al., *The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins.* EMBO J, 1997. **16**(11): p. 3145-57.
94. Pike-Overzet, K., et al., *Ectopic retroviral expression of LMO2, but not IL2Rgamma, blocks human T-cell development from CD34+ cells: implications for leukemogenesis in gene therapy.* Leukemia, 2007. **21**(4): p. 754-63.
95. Larson, R.C., et al., *The oncogenic LIM protein Rbtn2 causes thymic developmental aberrations that precede malignancy in transgenic mice.* Oncogene, 1995. **11**(5): p. 853-62.
96. Fisch, P., et al., *T-cell acute lymphoblastic lymphoma induced in transgenic mice by the RBTN1 and RBTN2 LIM-domain genes.* Oncogene, 1992. **7**(12): p. 2389-97.
97. Aplan, P.D., et al., *An scl gene product lacking the transactivation domain induces bony abnormalities and cooperates with LMO1 to generate T-cell malignancies in transgenic mice.* EMBO J, 1997. **16**(9): p. 2408-19.
98. Hacein-Bey-Abina, S., et al., *LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1.* Science, 2003. **302**(5644): p. 415-9.
99. Taghon, T., et al., *Homeobox gene expression profile in human hematopoietic multipotent stem cells and T-cell progenitors: implications for human T-cell development.* Leukemia, 2003. **17**(6): p. 1157-63.
100. Soulier, J., et al., *HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL).* Blood, 2005. **106**(1): p. 274-86.
101. Bernard, O.A., et al., *A new recurrent and specific cryptic translocation, t(5;14)(q35;q32), is associated with expression of the Hox11L2 gene in T acute lymphoblastic leukemia.* Leukemia, 2001. **15**(10): p. 1495-504.
102. Su, X.Y., et al., *Various types of rearrangements target TLX3 locus in T-cell acute lymphoblastic leukemia.* Genes, Chromosomes and Cancer, 2004. **41**(3): p. 243-249.
103. Kusy, S., et al., *NKX3.1 is a direct TAL1 target gene that mediates proliferation of TAL1-expressing human T cell acute lymphoblastic leukemia.* J Exp Med, 2010. **207**(10): p. 2141-56.
104. Bash, R.O.H.S.T.C.F.C.W.M.A.M.S.R.G.B.R., *Does activation of the TAL1 gene occur in a majority of patients with T- cell acute lymphoblastic leukemia? A pediatric oncology group study.* Blood, 1995. **86**(2): p. 666-676.
105. Chen, Q., et al., *The tal gene undergoes chromosome translocation in T cell leukemia and potentially encodes a helix-loop-helix protein.* EMBO J, 1990. **9**(2): p. 415-24.
106. Aplan, P.D., S.C. Raimondi, and I.R. Kirsch, *Disruption of the SCL gene by a t(1;3) translocation in a patient with T cell acute lymphoblastic leukemia.* J Exp Med, 1992. **176**(5): p. 1303-10.
107. Janssen, J.W., et al., *SIL-TAL1 deletion in T-cell acute lymphoblastic leukemia.* Leukemia, 1993. **7**(8): p. 1204-10.
108. Brown, L., et al., *Site-specific recombination of the tal-1 gene is a common occurrence in human T cell leukemia.* EMBO J, 1990. **9**(10): p. 3343-51.
109. van Grotel, M., et al., *The outcome of molecular-cytogenetic subgroups in pediatric T-cell acute lymphoblastic leukemia: a retrospective study of patients treated according to DCOG or COALL protocols.* Haematologica, 2006. **91**(9): p. 1212-21.
110. van Grotel, M., et al., *Prognostic significance of molecular-cytogenetic abnormalities in pediatric T-ALL is not explained by immunophenotypic differences.* Leukemia, 2008. **22**(1): p. 124-31.
111. Elwood, N.J. and C.G. Begley, *Reconstitution of mice with bone marrow cells expressing the SCL gene is insufficient to cause leukemia.* Cell Growth Differ, 1995. **6**(1): p. 19-25.

112. Robb, L., et al., *scl, a gene frequently activated in human T cell leukaemia, does not induce lymphomas in transgenic mice.* Oncogene, 1995. **10**(1): p. 205-9.
113. Kelliher, M.A., D.C. Seldin, and P. Leder, *Tal-1 induces T cell acute lymphoblastic leukemia accelerated by casein kinase IIalpha.* EMBO J, 1996. **15**(19): p. 5160-6.
114. Condorelli, G.L., et al., *T-cell-directed TAL-1 expression induces T-cell malignancies in transgenic mice.* Cancer Res, 1996. **56**(22): p. 5113-9.
115. Curtis, D.J., et al., *The CD2-scl transgene alters the phenotype and frequency of T-lymphomas in N-ras transgenic or p53 deficient mice.* Oncogene, 1997. **15**(24): p. 2975-83.
116. Shank-Calvo, J.A., et al., *p16Ink4a or p19Arf loss contributes to Tal1-induced leukemogenesis in mice.* Oncogene, 2006. **25**(21): p. 3023-31.
117. O'Neil, J., et al., *TAL1/SCL induces leukemia by inhibiting the transcriptional activity of E47/HEB.* Cancer Cell, 2004. **5**(6): p. 587-596.
118. Herblot, S., et al., *SCL and LMO1 alter thymocyte differentiation: inhibition of E2A-HEB function and pre-T alpha chain expression.* Nat Immunol, 2000. **1**(2): p. 138-44.
119. O'Neil, J., et al., *The DNA binding activity of TAL-1 is not required to induce leukemia/lymphoma in mice.* Oncogene, 2001. **20**(29): p. 3897-905.
120. Cardoso, B.A., et al., *TAL1/SCL is downregulated upon histone deacetylase inhibition in T-cell acute lymphoblastic leukemia cells.* Leukemia, 2011. **25**(10): p. 1578-86.
121. Wen, J., et al., *Tal1/SCL binding to pericentromeric DNA represses transcription.* J Biol Chem, 2005. **280**(13): p. 12956-66.
122. Hu, X., et al., *LSD1-mediated epigenetic modification is required for TAL1 function and hematopoiesis.* Proc Natl Acad Sci U S A, 2009. **106**(25): p. 10141-6.
123. Huang, S., et al., *P/CAF-mediated acetylation regulates the function of the basic helix-loop-helix transcription factor TAL1/SCL.* EMBO J, 2000. **19**(24): p. 6792-803.
124. Palomero, T.O.D.T.O. and J.F.A.A.M.A.N.D.S.W.S.S.L.R.S.L.W.L.X.S.Y.R.A.L.A.T. Neil, *Transcriptional regulatory networks downstream of TAL1/SCL in T-cell acute lymphoblastic leukemia.* Blood, 2006. **108**(3): p. 986-992.
125. Porcher, C., et al., *The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages.* Cell, 1996. **86**(1): p. 47-57.
126. Shivdasani, R.A., E.L. Mayer, and S.H. Orkin, *Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL.* Nature, 1995. **373**(6513): p. 432-4.
127. Porcher, C., et al., *Specification of hematopoietic and vascular development by the bHLH transcription factor SCL without direct DNA binding.* Development, 1999. **126**(20): p. 4603-15.
128. Visvader, J.E., Y. Fujiwara, and S.H. Orkin, *Unsuspected role for the T-cell leukemia protein SCL/tal-1 in vascular development.* Genes Dev, 1998. **12**(4): p. 473-479.
129. Mikkola, H.K.A., et al., *Haematopoietic stem cells retain long-term repopulating activity and multipotency in the absence of stem-cell leukaemia SCL/tal-1 gene.* Nature, 2003. **421**(6922): p. 547-551.
130. Lacombe, J.H.S.R.-S.S.H.A.B.S.I.N.N.S.G.H.T., *Scl regulates the quiescence and the long-term competence of hematopoietic stem cells.* Blood, 2009. **115**(4): p. 792-803.
131. Curtis, D.J., et al., *SCL is required for normal function of short-term repopulating hematopoietic stem cells.* Blood, 2004. **103**(9): p. 3342-8.
132. Lazrak, M., et al., *The bHLH TAL-1/SCL regulates endothelial cell migration and morphogenesis.* J Cell Sci, 2004. **117**(Pt 7): p. 1161-71.
133. Zhang, Y., et al., *SCL expression at critical points in human hematopoietic lineage commitment.* Stem Cells, 2005. **23**(6): p. 852-60.
134. Brunet de la Grange, P., et al., *Low SCL/TAL1 expression reveals its major role in adult hematopoietic myeloid progenitors and stem cells.* Blood, 2006. **108**(9): p. 2998-3004.

135. Lahlil, R., et al., *SCL assembles a multifactorial complex that determines glycophorin A expression.* Mol Cell Biol, 2004. **24**(4): p. 1439-52.
136. Xu, Z., et al., *Identification of a TAL1 target gene reveals a positive role for the LIM domain-binding protein Ldb1 in erythroid gene expression and differentiation.* Mol Cell Biol, 2003. **23**(21): p. 7585-99.
137. Bradley, C.K., et al., *The essential haematopoietic transcription factor Scl is also critical for neuronal development.* Eur J Neurosci, 2006. **23**(7): p. 1677-89.
138. Hsu, H.L., et al., *Preferred sequences for DNA recognition by the TAL1 helix-loop-helix proteins.* Mol Cell Biol, 1994. **14**(2): p. 1256-65.
139. Ono, Y., N. Fukuhara, and O. Yoshie, *TAL1 and LIM-only proteins synergistically induce retinaldehyde dehydrogenase 2 expression in T-cell acute lymphoblastic leukemia by acting as cofactors for GATA3.* Mol Cell Biol, 1998. **18**(12): p. 6939-50.
140. Lecuyer, E., et al., *The SCL complex regulates c-kit expression in hematopoietic cells through functional interaction with Sp1.* Blood, 2002. **100**(7): p. 2430-40.
141. Soler, E., et al., *The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation.* Genes Dev, 2010. **24**(3): p. 277-89.
142. Kassouf, M.T., et al., *Genome-wide identification of TAL1's functional targets: insights into its mechanisms of action in primary erythroid cells.* Genome Res, 2010. **20**(8): p. 1064-83.
143. Palii, C.G., et al., *Differential genomic targeting of the transcription factor TAL1 in alternate haematopoietic lineages.* EMBO J, 2011. **30**(3): p. 494-509.
144. Sanda, T., et al., *Core Transcriptional Regulatory Circuit Controlled by the TAL1 Complex in Human T Cell Acute Lymphoblastic Leukemia.* Cancer Cell, 2012. **22**(2): p. 209-221.
145. Bernard, M., et al., *Helix-loop-helix (E2-5, HEB, TAL1 and Id1) protein interaction with the TCRalphadelta enhancers.* Int Immunol, 1998. **10**(10): p. 1539-49.
146. Hansson, A., et al., *The basic helix–loop–helix transcription factor TAL1/SCL inhibits the expression of the p16INK4A and pTα genes.* Biochemical and Biophysical Research Communications, 2003. **312**(4): p. 1073-1081.
147. Ono, Y., N. Fukuhara, and O. Yoshie, *Transcriptional activity of TAL1 in T cell acute lymphoblastic leukemia (T-ALL) requires RBTN1 or -2 and induces TALLA1, a highly specific tumor marker of T-ALL.* J Biol Chem, 1997. **272**(7): p. 4576-81.
148. Mavrakis, K.J., et al., *Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia.* Nat Cell Biol, 2010. **12**(4): p. 372-9.
149. Thoms, J.A., et al., *ERG promotes T-acute lymphoblastic leukemia and is transcriptionally regulated in leukemic cells by a stem cell enhancer.* Blood, 2011. **117**(26): p. 7079-89.
150. Chang, P.Y., et al., *NFKB1 is a direct target of the TAL1 oncoprotein in human T leukemia cells.* Cancer Res, 2006. **66**(12): p. 6008-13.
151. Aplan, P.D., et al., *The SCL gene is formed from a transcriptionally complex locus.* Mol Cell Biol, 1990. **10**(12): p. 6426-35.
152. Bockamp, E.O., et al., *Lineage-restricted regulation of the murine SCL/TAL-1 promoter.* Blood, 1995. **86**(4): p. 1502-14.
153. Bernard, O., et al., *A third tal-1 promoter is specifically used in human T cell leukemias.* J Exp Med, 1992. **176**(4): p. 919-25.
154. Patel, B., et al., *Aberrant TAL1 activation is mediated by an interchromosomal interaction in human T-cell acute lymphoblastic leukemia.* Leukemia, 2014. **28**(2): p. 349-61.
155. Ong, C.T. and V.G. Corces, *Enhancer function: new insights into the regulation of tissue-specific gene expression.* Nat Rev Genet, 2011. **12**(4): p. 283-93.
156. Sanchez, M., et al., *An SCL 3' enhancer targets developing endothelium together with embryonic and adult haematopoietic progenitors.* Development, 1999. **126**(17): p. 3891-904.

157. Gottgens, B., et al., *The scl +18/19 stem cell enhancer is not required for hematopoiesis: identification of a 5' bifunctional hematopoietic-endothelial enhancer bound by Fli-1 and Elf-1.* Mol Cell Biol, 2004. **24**(5): p. 1870-83.
158. Ogilvy, S., et al., *The SCL +40 enhancer targets the midbrain together with primitive and definitive hematopoiesis and is regulated by SCL and GATA proteins.* Mol Cell Biol, 2007. **27**(20): p. 7206-19.
159. Gottgens, B., et al., *Analysis of vertebrate SCL loci identifies conserved enhancers.* Nat Biotechnol, 2000. **18**(2): p. 181-6.
160. Courtes, C., et al., *Erythroid-specific inhibition of the tal-1 intragenic promoter is due to binding of a repressor to a novel silencer.* J Biol Chem, 2000. **275**(2): p. 949-58.
161. Le Clech, M., et al., *PU.1/Spi-1 binds to the human TAL-1 silencer to mediate its activity.* J Mol Biol, 2006. **355**(1): p. 9-19.
162. Dhami, P., et al., *Genomic approaches uncover increasing complexities in the regulatory landscape at the human SCL (TAL1) locus.* PLoS One, 2010. **5**(2): p. e9059.
163. Zhou, Y., et al., *Chromatin looping defines expression of TAL1, its flanking genes, and regulation in T-ALL.* Blood, 2013. **122**(26): p. 4199-209.
164. Orom, U.A., et al., *Long noncoding RNAs with enhancer-like function in human cells.* Cell, 2010. **143**(1): p. 46-58.
165. Lai, F., et al., *Activating RNAs associate with Mediator to enhance chromatin architecture and transcription.* Nature, 2013. **494**(7438): p. 497-501.
166. Mattick, J.S., *The genetic signatures of noncoding RNAs.* PLoS Genet, 2009. **5**(4): p. e1000459.
167. Ling, H., M. Fabbri, and G.A. Calin, *MicroRNAs and other non-coding RNAs as targets for anticancer drug development.* Nat Rev Drug Discov, 2013. **12**(11): p. 847-65.
168. Kowalczyk, M.S., D.R. Higgs, and T.R. Gingeras, *Molecular biology: RNA discrimination.* Nature, 2012. **482**(7385): p. 310-311.
169. Harrow, J., et al., *GENCODE: The reference human genome annotation for The ENCODE Project.* Genome Res, 2012. **22**(9): p. 1760-1774.
170. *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. **489**(7414): p. 57-74.
171. Lim, L.P., et al., *Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs.* Nature, 2005. **433**(7027): p. 769-773.
172. Griffiths-Jones, S., et al., *miRBase: tools for microRNA genomics.* Nucleic Acids Research, 2008. **36**(suppl 1): p. D154-D158.
173. Lewis, B.P., C.B. Burge, and D.P. Bartel, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.* Cell, 2005. **120**(1): p. 15-20.
174. Kozomara, A. and S. Griffiths-Jones, *miRBase: integrating microRNA annotation and deep-sequencing data.* Nucleic Acids Res, 2011. **39**(Database issue): p. D152-7.
175. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* Cell, 1993. **75**(5): p. 843-854.
176. Lagos-Quintana, M., et al., *Identification of Novel Genes Coding for Small Expressed RNAs.* Science, 2001. **294**(5543): p. 853-858.
177. Lau, N.C., et al., *An Abundant Class of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis elegans.* Science, 2001. **294**(5543): p. 858-862.
178. Krol, J., I. Loedige, and W. Filipowicz, *The widespread regulation of microRNA biogenesis, function and decay.* Nat Rev Genet, 2010. **11**(9): p. 597-610.
179. Bartel, D.P., *MicroRNAs: Genomics, Biogenesis, Mechanism, and Function.* Cell, 2004. **116**(2): p. 281-297.
180. Peters, L. and G. Meister, *Argonaute proteins: mediators of RNA silencing.* Mol Cell, 2007. **26**(5): p. 611-23.

181. Rodriguez, A., et al., *Identification of mammalian microRNA host genes and transcription units.* Genome Res, 2004. **14**(10A): p. 1902-10.
182. Altuvia, Y., et al., *Clustering and conservation patterns of human microRNAs.* Nucleic Acids Res, 2005. **33**(8): p. 2697-706.
183. Baskerville, S. and D.P. Bartel, *Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes.* RNA, 2005. **11**(3): p. 241-7.
184. Poliseno, L., et al., *Identification of the miR-106b~25 microRNA cluster as a proto-oncogenic PTEN-targeting intron that cooperates with its host gene MCM7 in transformation.* Sci Signal, 2010. **3**(117): p. ra29.
185. Small, E.M., et al., *MicroRNA-218 Regulates Vascular Patterning by Modulation of Slit-Robo Signaling.* Circulation Research, 2010. **107**(11): p. 1336-1344.
186. Berezikov, E., et al., *Mammalian mirtron genes.* Mol Cell, 2007. **28**(2): p. 328-36.
187. Ender, C., et al., *A human snoRNA with microRNA-like functions.* Mol Cell, 2008. **32**(4): p. 519-28.
188. Maute, R.L., et al., *tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma.* Proceedings of the National Academy of Sciences, 2013. **110**(4): p. 1404-1409.
189. Grimson, A., et al., *MicroRNA targeting specificity in mammals: determinants beyond seed pairing.* Mol Cell, 2007. **27**(1): p. 91-105.
190. Rajewsky, N., *microRNA target predictions in animals.* Nat Genet, 2006.
191. Friedman, R.C., et al., *Most mammalian mRNAs are conserved targets of microRNAs.* Genome Res, 2009. **19**(1): p. 92-105.
192. Selbach, M., et al., *Widespread changes in protein synthesis induced by microRNAs.* Nature, 2008. **455**(7209): p. 58-63.
193. Baek, D., et al., *The impact of microRNAs on protein output.* Nature, 2008. **455**(7209): p. 64-71.
194. Bartel, D.P., *MicroRNAs: Target Recognition and Regulatory Functions.* Cell, 2009. **136**(2): p. 215-233.
195. Yekta, S., I.h. Shih, and D.P. Bartel, *MicroRNA-Directed Cleavage of HOXB8 mRNA.* Science, 2004. **304**(5670): p. 594-596.
196. Kanellopoulou, C. and S. Monticelli, *A role for microRNAs in the development of the immune system and in the pathogenesis of cancer.* Semin Cancer Biol, 2008. **18**(2): p. 79-88.
197. Uhlmann, S., et al., *Global microRNA level regulation of EGFR-driven cell-cycle protein network in breast cancer*. Vol. 8. 2012.
198. Pillai, R.S., S.N. Bhattacharyya, and W. Filipowicz, *Repression of protein synthesis by miRNAs: how many mechanisms?* Trends Cell Biol, 2007. **17**(3): p. 118-26.
199. Wu, L., J. Fan, and J.G. Belasco, *MicroRNAs direct rapid deadenylation of mRNA.* Proc Natl Acad Sci U S A, 2006. **103**(11): p. 4034-9.
200. Jing, Q., et al., *Involvement of microRNA in AU-rich element-mediated mRNA instability.* Cell, 2005. **120**(5): p. 623-34.
201. Bagga, S., et al., *Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation.* Cell, 2005. **122**(4): p. 553-63.
202. Beilharz, T.H., et al., *microRNA-mediated messenger RNA deadenylation contributes to translational repression in mammalian cells.* PLoS One, 2009. **4**(8): p. e6783.
203. Valencia-Sanchez, M.A., et al., *Control of translation and mRNA degradation by miRNAs and siRNAs.* Genes & Development, 2006. **20**(5): p. 515-524.
204. Nishikura, K., *Editor meets silencer: crosstalk between RNA editing and RNA interference.* Nat Rev Mol Cell Biol, 2006. **7**(12): p. 919-31.
205. Kawahara, Y., et al., *Redirection of silencing targets by adenosine-to-inosine editing of miRNAs.* Science, 2007. **315**(5815): p. 1137-40.

206. Yang, W., et al., *Modulation of microRNA processing and expression through RNA editing by ADAR deaminases.* Nat Struct Mol Biol, 2006. **13**(1): p. 13-21.
207. Bernstein, E., et al., *Dicer is essential for mouse development.* Nat Genet, 2003. **35**(3): p. 215-7.
208. Wienholds, E., et al., *The microRNA-producing enzyme Dicer1 is essential for zebrafish development.* Nat Genet, 2003. **35**(3): p. 217-8.
209. Wang, Y., et al., *DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal.* Nat Genet, 2007. **39**(3): p. 380-5.
210. Calin, G.A., et al., *Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia.* Proceedings of the National Academy of Sciences, 2002. **99**(24): p. 15524-15529.
211. Fabbri, M., C.M. Croce, and G.A. Calin, *MicroRNAs in the ontogeny of leukemias and lymphomas.* Leukemia & Lymphoma, 2009. **50**(2): p. 160-170.
212. Landgraf, P., et al., *A mammalian microRNA expression atlas based on small RNA library sequencing.* Cell, 2007. **129**(7): p. 1401-14.
213. Chen, C.Z., et al., *MicroRNAs modulate hematopoietic lineage differentiation.* Science, 2004. **303**(5654): p. 83-6.
214. Georgantas, R.W., et al., *CD34+ hematopoietic stem-progenitor cell microRNA expression and function: A circuit diagram of differentiation control.* Proceedings of the National Academy of Sciences, 2007. **104**(8): p. 2750-2755.
215. Li, Q.-J., et al., *miR-181a Is an Intrinsic Modulator of T Cell Sensitivity and Selection.* Cell, 2007. **129**(1): p. 147-161.
216. Cobb, B.S., et al., *A role for Dicer in immune regulation.* The Journal of Experimental Medicine, 2006. **203**(11): p. 2519-2527.
217. Muljo, S.A., et al., *Aberrant T cell differentiation in the absence of Dicer.* The Journal of Experimental Medicine, 2005. **202**(2): p. 261-269.
218. Cobb, B.S., et al., *T cell lineage choice and differentiation in the absence of the RNase III enzyme Dicer.* The Journal of Experimental Medicine, 2005. **201**(9): p. 1367-1373.
219. Ghisi, M., et al., *Modulation of microRNA expression in human T-cell development: targeting of NOTCH3 by miR-150.* Blood, 2011. **117**(26): p. 7053-62.
220. Cimmino, A., et al., *miR-15 and miR-16 induce apoptosis by targeting BCL2.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(39): p. 13944-13949.
221. Calin, G.A., et al., *Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(9): p. 2999-3004.
222. Schotte, D., et al., *Expression of miR-196b is not exclusively MLL-driven but is especially linked to activation of HOXA genes in pediatric acute lymphoblastic leukemia.* Haematologica, 2010. **95**(10): p. 1675-82.
223. Lu, J., et al., *MicroRNA expression profiles classify human cancers.* Nature, 2005. **435**(7043): p. 834-838.
224. Jiang, J., et al., *Real-time expression profiling of microRNA precursors in human cancer cell lines.* Nucleic Acids Res, 2005. **33**(17): p. 5394-403.
225. Mayr, C. and D.P. Bartel, *Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells.* Cell, 2009. **138**(4): p. 673-84.
226. He, L., et al., *A microRNA polycistron as a potential human oncogene.* Nature, 2005. **435**(7043): p. 828-833.
227. Costinean, S., et al., *Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice.* Proc Natl Acad Sci U S A, 2006. **103**(18): p. 7024-9.

228. Chang, T.-C., et al., *Widespread microRNA repression by Myc contributes to tumorigenesis.* Nat Genet, 2008. **40**(1): p. 43-50.
229. O'Donnell, K.A., et al., *c-Myc-regulated microRNAs modulate E2F1 expression.* Nature, 2005. **435**(7043): p. 839-843.
230. Correia, N.C., et al., *The multiple layers of non-genetic regulation of PTEN tumour suppressor activity.* European Journal of Cancer, 2014. **50**(1): p. 216-225.
231. Meng, F., et al., *MicroRNA-21 Regulates Expression of the PTEN Tumor Suppressor Gene in Human Hepatocellular Cancer.* Gastroenterology, 2007. **133**(2): p. 647-658.
232. Lou, Y., et al., *MicroRNA-21 promotes the cell proliferation, invasion and migration abilities in ovarian epithelial carcinomas through inhibiting the expression of PTEN protein.* Int J Mol Med, 2010. **26**(6): p. 819-27.
233. Zhang, J.G., et al., *MicroRNA-21 (miR-21) represses tumor suppressor PTEN and promotes growth and invasion in non-small cell lung cancer (NSCLC).* Clin Chim Acta, 2010. **411**(11-12): p. 846-52.
234. Darido, C., et al., *Targeting of the tumor suppressor GRHL3 by a miR-21-dependent proto-oncogenic network results in PTEN loss and tumorigenesis.* Cancer cell, 2011. **20**(5): p. 635-48.
235. Yang, H., et al., *MicroRNA expression profiling in human ovarian cancer: miR-214 induces cell survival and cisplatin resistance by targeting PTEN.* Cancer Res, 2008. **68**(2): p. 425-33.
236. Huse, J.T., et al., *The PTEN-regulating microRNA miR-26a is amplified in high-grade glioma and facilitates gliomagenesis in vivo.* Genes Dev, 2009. **23**(11): p. 1327-37.
237. Liu, B., et al., *MiR-26a enhances metastasis potential of lung cancer cells via AKT pathway by targeting PTEN.* Biochim Biophys Acta, 2012. **1822**(11): p. 1692-704.
238. Garofalo, M., et al., *miR-221&222 regulate TRAIL resistance and enhance tumorigenicity through PTEN and TIMP3 downregulation.* Cancer cell, 2009. **16**(6): p. 498-509.
239. Chun-Zhi, Z., et al., *MicroRNA-221 and microRNA-222 regulate gastric carcinoma cell proliferation and radioresistance by targeting PTEN.* BMC Cancer, 2010. **10**: p. 367.
240. Wang, C., et al., *miR-29b regulates migration of human breast cancer cells.* Mol Cell Biochem, 2011. **352**(1-2): p. 197-207.
241. Shi, W., et al., *MicroRNA-301 mediates proliferation and invasion in human breast cancer.* Cancer Res, 2011. **71**(8): p. 2926-37.
242. Wu, Z., et al., *Upregulation of miR-153 promotes cell proliferation via downregulation of the PTEN tumor suppressor gene in human prostate cancer.* Prostate, 2013. **73**(6): p. 596-604.
243. He, L., et al., *A microRNA component of the p53 tumour suppressor network.* Nature, 2007. **447**(7148): p. 1130-1134.
244. Malumbres, M., *miRNAs versus oncogenes: the power of social networking.* Molecular Systems Biology, 2012. **8**(1).
245. Felli, N., et al., *MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation.* Proc Natl Acad Sci U S A, 2005. **102**(50): p. 18081-6.
246. Salmena, L., et al., *A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language?* Cell, 2011. **146**(3): p. 353-8.
247. Poliseno, L., et al., *A coding-independent function of gene and pseudogene mRNAs regulates tumour biology.* Nature, 2010. **465**(7301): p. 1033-8.
248. Tay, Y., et al., *Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs.* Cell, 2011. **147**(2): p. 344-57.
249. Fabbri, M., C.M. Croce, and G.A. Calin, *MicroRNAs in the ontogeny of leukemias and lymphomas.* Leuk Lymphoma, 2009. **50**(2): p. 160-70.
250. Schotte, D., et al., *Identification of new microRNA genes and aberrant microRNA profiles in childhood acute lymphoblastic leukemia.* Leukemia, 2009. **23**(2): p. 313-22.

251. de Oliveira, J.C., et al., *Differential MiRNA expression in childhood acute lymphoblastic leukemia and association with clinical and biological features.* Leukemia Research, 2012. **36**(3): p. 293-298.
252. Schotte, D., et al., *Discovery of new microRNAs by small RNAome deep sequencing in childhood acute lymphoblastic leukemia.* Leukemia, 2011. **25**(9): p. 1389-1399.
253. Zhang, H., et al., *MicroRNA patterns associated with clinical prognostic parameters and CNS relapse prediction in pediatric acute leukemia.* PLoS One, 2009. **4**(11): p. e7826.
254. Han, B.W., et al., *A set of miRNAs that involve in the pathways of drug resistance and leukemic stem-cell differentiation is associated with the risk of relapse and glucocorticoid response in childhood ALL.* Hum Mol Genet, 2011. **20**(24): p. 4903-15.
255. Schotte, D., et al., *MicroRNAs characterize genetic diversity and drug resistance in pediatric acute lymphoblastic leukemia.* Haematologica, 2011.
256. Mi, S., et al., *MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia.* Proceedings of the National Academy of Sciences, 2007. **104**(50): p. 19971-19976.
257. Fulci, V., et al., *Characterization of B- and T-lineage acute lymphoblastic leukemia by integrated analysis of MicroRNA and mRNA expression profiles.* Genes, Chromosomes and Cancer, 2009. **48**(12): p. 1069-1082.
258. Xiao, C., et al., *Lymphoproliferative disease and autoimmunity in mice with increased miR-17-92 expression in lymphocytes.* Nat Immunol, 2008. **9**(4): p. 405-14.
259. Rao, E., et al., *The miRNA-17 approximately 92 cluster mediates chemoresistance and enhances tumor growth in mantle cell lymphoma via PI3K/AKT pathway activation.* Leukemia, 2012. **26**(5): p. 1064-72.
260. Bousquet, M., et al., *MicroRNA miR-125b causes leukemia.* Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21558-63.
261. Mavrakis, K.J., et al., *A cooperative microRNA-tumor suppressor gene network in acute T-cell lymphoblastic leukemia (T-ALL).* Nat Genet, 2011. **43**(7): p. 673-678.
262. Coskun, E., et al., *MicroRNA profiling reveals aberrant microRNA expression in adult ETP-ALL and functional studies implicate a role for miR-222 in acute leukemia.* Leuk Res, 2013. **37**(6): p. 647-56.
263. Gimenes-Teixeira, H.L., et al., *Increased expression of miR-221 is associated with shorter overall survival in T-cell acute lymphoid leukemia.* Exp Hematol Oncol, 2013. **2**(1): p. 10.
264. Mavrakis, K.J., et al., *Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia.* Nat Cell Biol, 2010. **12**(4): p. 372-379.
265. Nagel, S., et al., *Activation of miR-17-92 by NK-like homeodomain proteins suppresses apoptosis via reduction of E2F1 in T-cell acute lymphoblastic leukemia.* Leuk Lymphoma, 2009. **50**(1): p. 101-8.
266. Mets, E., et al., *MicroRNA-128-3p is a novel oncomiR targeting PHF6 in T-cell acute lymphoblastic leukemia.* Haematologica, 2014.
267. Real, P.J., et al., *Gamma-secretase inhibitors reverse glucocorticoid resistance in T cell acute lymphoblastic leukemia.* Nat Med, 2009. **15**(1): p. 50-8.
268. Fragoso, R., et al., *Modulating the strength and threshold of NOTCH oncogenic signals by mir-181a-1/b-1.* PLoS Genet, 2012. **8**(8): p. e1002855.
269. Li, X., et al., *Repression of tumor suppressor miR-451 is essential for NOTCH1-induced oncogenesis in T-ALL.* The Journal of Experimental Medicine, 2011. **208**(4): p. 663-675.
270. Kumar, V., et al., *Notch and NF-kB signaling pathways regulate miR-223/FBXW7 axis in T-cell acute lymphoblastic leukemia.* Leukemia, 2014.
271. Lv, M., et al., *An oncogenic role of miR-142-3p in human T-cell acute lymphoblastic leukemia (T-ALL) by targeting glucocorticoid receptor-[alpha] and cAMP/PKA pathways.* Leukemia, 2012. **26**(4): p. 769-777.

272. Coskun, E., et al., *The role of microRNA-196a and microRNA-196b as ERG regulators in acute myeloid leukemia and acute T-lymphoblastic leukemia.* Leukemia Research, 2011. **35**(2): p. 208-213.
273. Bhatia, S., D. Kaul, and N. Varma, *Functional genomics of tumor suppressor miR-196b in T-cell acute lymphoblastic leukemia.* Mol Cell Biochem, 2011. **346**(1-2): p. 103-16.
274. Schotte, D., R. Pieters, and M.L. Den Boer, *MicroRNAs in acute leukemia: from biological players to clinical contributors.* Leukemia, 2012. **26**(1): p. 1-12.
275. Choong, M.L., H.H. Yang, and I. McNiece, *MicroRNA expression profiling during human cord blood-derived CD34 cell erythropoiesis.* Experimental Hematology, 2007. **35**(4): p. 551-564.
276. Garzon, R., et al., *MicroRNA fingerprints during human megakaryocytopoiesis.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(13): p. 5078-5083.
277. Griffiths-Jones, S., et al., *miRBase: tools for microRNA genomics.* Nucleic Acids Res, 2008. **36**(Database issue): p. D154-8.
278. Krek, A., et al., *Combinatorial microRNA target predictions.* Nat Genet, 2005. **37**(5): p. 495-500.
279. Betel, D., et al., *Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites.* Genome Biol, 2010. **11**(8): p. R90.
280. John, B., et al., *Human MicroRNA targets.* PLoS Biol, 2004. **2**(11): p. e363.
281. Betel, D., et al., *The microRNA.org resource: targets and expression.* Nucleic Acids Research, 2008. **36**(suppl 1): p. D149-D153.
282. Maragkakis, M., et al., *Accurate microRNA target prediction correlates with protein repression levels.* BMC Bioinformatics, 2009. **10**: p. 295.
283. Wang, X., *miRDB: a microRNA target prediction and functional annotation database with a wiki interface.* RNA, 2008. **14**(6): p. 1012-7.
284. Yang, J.-H., et al., *starBase: a database for exploring microRNA–mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data.* Nucleic Acids Research, 2011. **39**(suppl 1): p. D202-D209.
285. Vlachos, I.S., et al., *DIANA miRPath v.2.0: investigating the combinatorial effect of microRNAs in pathways.* Nucleic Acids Res, 2012. **40**(Web Server issue): p. W498-504.
286. Bueno, M.J., et al., *Multiple E2F-induced microRNAs prevent replicative stress in response to mitogenic signaling.* Mol Cell Biol, 2010. **30**(12): p. 2983-95.
287. Voorhoeve, P.M., et al., *A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors.* Cell, 2006. **124**(6): p. 1169-81.
288. Barata, J.T., et al., *Interleukin-7 promotes survival and cell cycle progression of T-cell acute lymphoblastic leukemia cells by down-regulating the cyclin-dependent kinase inhibitor p27(kip1).* Blood, 2001. **98**(5): p. 1524-31.
289. Zufferey, R., et al., *Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery.* J Virol, 1998. **72**(12): p. 9873-80.
290. Mestdagh, P., et al., *A novel and universal method for microRNA RT-qPCR data normalization.* Genome Biol, 2009. **10**(6): p. R64.
291. Gould, J., et al., *Comparative gene marker selection suite.* Bioinformatics, 2006. **22**(15): p. 1924-5.
292. Lee, T.I., S.E. Johnstone, and R.A. Young, *Chromatin immunoprecipitation and microarray-based analysis of protein location.* Nat Protoc, 2006. **1**(2): p. 729-48.
293. Shirdel, E.A., et al., *NAViGaTing the micronome--using multiple microRNA prediction databases to identify signalling pathway-associated microRNAs.* PLoS One, 2011. **6**(2): p. e17429.
294. Brown, K.R., et al., *NAViGaTOR: Network Analysis, Visualization and Graphing Toronto.* Bioinformatics, 2009. **25**(24): p. 3327-9.
295. Sanda, T., et al., *Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia.* Cancer Cell, 2012. **22**(2): p. 209-21.

296. Lee, M.S., et al., *Gene expression profiles during human CD4+ T cell differentiation.* Int Immunol, 2004. **16**(8): p. 1109-24.
297. Patnaik, S.K., et al., *Overexpression of the lung cancer-prognostic miR-146b microRNAs has a minimal and negative effect on the malignant phenotype of A549 lung cancer cells.* PLoS One, 2011. **6**(7): p. e22379.
298. Hammell, M., et al., *mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts.* Nat Methods, 2008. **5**(9): p. 813-9.
299. Kuhn, D.E., et al., *Experimental validation of miRNA targets.* Methods, 2008. **44**(1): p. 47-54.
300. Mansour, M.R., et al., *The TAL1 complex targets the FBXW7 tumor suppressor by activating miR-223 in human T cell acute lymphoblastic leukemia.* J Exp Med, 2013. **210**(8): p. 1545-57.
301. Clappier, E., et al., *Clonal selection in xenografted human T cell acute lymphoblastic leukemia recapitulates gain of malignancy at relapse.* J Exp Med, 2011. **208**(4): p. 653-61.
302. Mets, E., et al., *MicroRNA-193b-3p acts as a tumor suppressor by targeting the MYB oncogene in T-cell acute lymphoblastic leukemia.* Leukemia, 2014.
303. Katakowski, M., et al., *MiR-146b-5p suppresses EGFR expression and reduces in vitro migration and invasion of glioma.* Cancer Invest, 2010. **28**(10): p. 1024-30.
304. Lin, F., et al., *Inhibitory effects of miR-146b-5p on cell migration and invasion of pancreatic cancer by targeting MMP16.* J Huazhong Univ Sci Technolog Med Sci, 2011. **31**(4): p. 509-14.
305. Hurst, D.R., et al., *Breast cancer metastasis suppressor 1 up-regulates miR-146, which suppresses breast cancer metastasis.* Cancer Res, 2009. **69**(4): p. 1279-83.
306. Silva, A., et al., *IL-7 contributes to the progression of human T-cell acute lymphoblastic leukemias.* Cancer Res, 2011. **71**(14): p. 4780-9.
307. Schotte, D., et al., *MicroRNA characterize genetic diversity and drug resistance in pediatric acute lymphoblastic leukemia.* Haematologica, 2011. **96**(5): p. 703-11.
308. Lecuyer, E. and T. Hoang, *SCL: from the origin of hematopoiesis to stem cells and leukemia.* Exp Hematol, 2004. **32**(1): p. 11-24.
309. Choong, M.L., H.H. Yang, and I. McNiece, *MicroRNA expression profiling during human cord blood-derived CD34 cell erythropoiesis.* Exp Hematol, 2007. **35**(4): p. 551-64.
310. Garzon, R., et al., *MicroRNA fingerprints during human megakaryocytopoiesis.* Proc Natl Acad Sci U S A, 2006. **103**(13): p. 5078-83.
311. Felli, N., et al., *MicroRNA 223-dependent expression of LMO2 regulates normal erythropoiesis.* Haematologica, 2009. **94**(4): p. 479-86.
312. Wang, D., et al., *OncomiRDB: a database for the experimentally verified oncogenic and tumor-suppressive microRNAs.* Bioinformatics, 2014. **30**(15): p. 2237-8.
313. Varambally, S., et al., *Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer.* Science, 2008. **322**(5908): p. 1695-9.
314. Friedman, J.M., et al., *The putative tumor suppressor microRNA-101 modulates the cancer epigenome by repressing the polycomb group protein EZH2.* Cancer Res, 2009. **69**(6): p. 2623-9.
315. Wang, H.J., et al., *MicroRNA-101 is down-regulated in gastric cancer and involved in cell migration and invasion.* Eur J Cancer, 2010. **46**(12): p. 2295-303.
316. Carvalho, J., et al., *Lack of microRNA-101 causes E-cadherin functional deregulation through EZH2 up-regulation in intestinal gastric cancer.* J Pathol, 2012. **228**(1): p. 31-44.
317. Sakurai, T., et al., *The enhancer of zeste homolog 2 (EZH2), a potential therapeutic target, is regulated by miR-101 in renal cancer cells.* Biochem Biophys Res Commun, 2012. **422**(4): p. 607-14.
318. Qazi, A.M., et al., *Restoration of E-cadherin expression in pancreatic ductal adenocarcinoma treated with microRNA-101.* Surgery, 2012. **152**(4): p. 704-11; discussion 711-3.
319. Luo, C., et al., *MiR-101 inhibits melanoma cell invasion and proliferation by targeting MITF and EZH2.* Cancer Lett, 2013. **341**(2): p. 240-7.

320. Hao, Y., et al., *Enforced expression of miR-101 inhibits prostate cancer cell growth by modulating the COX-2 pathway in vivo.* Cancer Prev Res (Phila), 2011. **4**(7): p. 1073-83.
321. He, X.P., et al., *Downregulation of miR-101 in gastric cancer correlates with cyclooxygenase-2 overexpression and tumor growth.* FEBS J, 2012. **279**(22): p. 4201-12.
322. Robertus, J.L., et al., *MiRNA profiling in B non-Hodgkin lymphoma: a MYC-related miRNA profile characterizes Burkitt lymphoma.* Br J Haematol, 2010. **149**(6): p. 896-9.
323. Liu, X., et al., *miRNA mediated up-regulation of cochaperone p23 acts as an anti-apoptotic factor in childhood acute lymphoblastic leukemia.* Leuk Res, 2012. **36**(9): p. 1098-104.
324. Sasaki, D., et al., *Overexpression of Enhancer of zeste homolog 2 with trimethylation of lysine 27 on histone H3 in adult T-cell leukemia/lymphoma as a target for epigenetic therapy.* Haematologica, 2011. **96**(5): p. 712-9.
325. Song, B., et al., *Mechanism of chemoresistance mediated by miR-140 in human osteosarcoma and colon cancer cells.* Oncogene, 2009. **28**(46): p. 4065-74.
326. Zhang, Y., et al., *Estrogen receptor alpha signaling regulates breast tumor-initiating cells by down-regulating miR-140 which targets the transcription factor SOX2.* J Biol Chem, 2012. **287**(49): p. 41514-22.
327. Yang, H., et al., *MicroRNA-140-5p suppresses tumor growth and metastasis by targeting transforming growth factor beta receptor 1 and fibroblast growth factor 9 in hepatocellular carcinoma.* Hepatology, 2013. **58**(1): p. 205-17.
328. Yuan, Y., et al., *miR-140 suppresses tumor growth and metastasis of non-small cell lung cancer by targeting insulin-like growth factor 1 receptor.* PLoS One, 2013. **8**(9): p. e73604.
329. Li, Q., et al., *Downregulation of miR-140 promotes cancer stem cell formation in basal-like early stage breast cancer.* Oncogene, 2014. **33**(20): p. 2589-600.
330. Tsuno, S., et al., *Hsa-miR-520d induces hepatoma cells to form normal liver tissues via a stemness-mediated process.* Sci Rep, 2014. **4**: p. 3852.
331. Kim, T.H., et al., *Deregulation of miR-519a, 153, and 485-5p and its clinicopathological relevance in ovarian epithelial tumours.* Histopathology, 2010. **57**(5): p. 734-43.
332. Anaya-Ruiz, M., C. Bandala, and J.L. Perez-Santos, *miR-485 acts as a tumor suppressor by inhibiting cell growth and migration in breast carcinoma T47D cells.* Asian Pac J Cancer Prev, 2013. **14**(6): p. 3757-60.
333. He, N., et al., *miR-485-5p binding site SNP rs8752 in HPGD gene is associated with breast cancer risk.* PLoS One, 2014. **9**(7): p. e102093.
334. Li, Q.Q., et al., *Involvement of NF-kappaB/miR-448 regulatory feedback loop in chemotherapy-induced epithelial-mesenchymal transition of breast cancer cells.* Cell Death Differ, 2011. **18**(1): p. 16-25.
335. Yuan, W., et al., *Downregulation of microRNA-142 by proto-oncogene LMO2 and its co-factors.* Leukemia, 2007. **22**(5): p. 1067-1071.
336. Sun, W., et al., *miR-223 and miR-142 attenuate hematopoietic cell proliferation, and miR-223 positively regulates miR-142 through LMO2 isoforms and CEBP-beta.* Cell Res, 2010. **20**(10): p. 1158-69.
337. Fukao, T., et al., *An evolutionarily conserved mechanism for microRNA-223 expression revealed by microRNA gene profiling.* Cell, 2007. **129**(3): p. 617-31.
338. Johnnidis, J.B., et al., *Regulation of progenitor cell proliferation and granulocyte function by microRNA-223.* Nature, 2008. **451**(7182): p. 1125-9.
339. Pulikkan, J.A., et al., *Cell-cycle regulator E2F1 and microRNA-223 comprise an autoregulatory negative feedback loop in acute myeloid leukemia.* Blood, 2010. **115**(9): p. 1768-78.
340. Agatheeswaran, S., et al., *BCR-ABL mediated repression of miR-223 results in the activation of MEF2C and PTBP2 in chronic myeloid leukemia.* Leukemia, 2012.

341. Mavrakis, K.J., et al., *A cooperative microRNA-tumor suppressor gene network in acute T-cell lymphoblastic leukemia (T-ALL).* Nat Genet, 2011. **43**(7): p. 673-8.
342. Chiaretti, S., et al., *Gene expression profiling identifies a subset of adult T-cell acute lymphoblastic leukemia with myeloid-like gene features and over-expression of miR-223.* Haematologica, 2010. **95**(7): p. 1114-21.
343. Xu, Y., et al., *MicroRNA-223 regulates cyclin E activity by modulating expression of F-box and WD-40 domain protein 7.* J Biol Chem, 2010. **285**(45): p. 34439-46.
344. Inuzuka, H., et al., *SCF(FBW7) regulates cellular apoptosis by targeting MCL1 for ubiquitylation and destruction.* Nature, 2011. **471**(7336): p. 104-9.
345. Yada, M., et al., *Phosphorylation-dependent degradation of c-Myc is mediated by the F-box protein Fbw7.* EMBO J, 2004. **23**(10): p. 2116-25.
346. Strohmaier, H., et al., *Human F-box protein hCdc4 targets cyclin E for proteolysis and is mutated in a breast cancer cell line.* Nature, 2001. **413**(6853): p. 316-22.
347. Cassavaugh, J.M., et al., *Negative regulation of HIF-1alpha by an FBW7-mediated degradation pathway during hypoxia.* J Cell Biochem, 2011. **112**(12): p. 3882-90.
348. Kitagawa, K., et al., *GSK3 regulates the expressions of human and mouse c-Myb via different mechanisms.* Cell Div, 2010. **5**: p. 27.
349. Homminga, I., et al., *Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia.* Cancer Cell, 2011. **19**(4): p. 484-97.
350. Arora, S., et al., *MicroRNA-328 is associated with (non-small) cell lung cancer (NSCLC) brain metastasis and mediates NSCLC migration.* Int J Cancer, 2011. **129**(11): p. 2621-31.
351. Lee, S.H., et al., *E2F1-directed activation of Bcl-2 is correlated with lactoferrin-induced apoptosis in Jurkat leukemia T lymphocytes.* Biometals, 2010. **23**(3): p. 507-14.
352. Su, J.L., et al., *Resveratrol induces FasL-related apoptosis through Cdc42 activation of ASK1/JNK-dependent signaling pathway in human leukemia HL-60 cells.* Carcinogenesis, 2005. **26**(1): p. 1-10.
353. Yunta, M. and P.A. Lazo, *Apoptosis protection and survival signal by the CD53 tetraspanin antigen.* Oncogene, 2003. **22**(8): p. 1219-24.
354. Dong, H., et al., *Inhibition of PDE3, PDE4 and PDE7 potentiates glucocorticoid-induced apoptosis and overcomes glucocorticoid resistance in CEM T leukemic cells.* Biochem Pharmacol, 2010. **79**(3): p. 321-9.
355. Dadi, S., et al., *TLX homeodomain oncogenes mediate T cell maturation arrest in T-ALL via interaction with ETS1 and suppression of TCRalpha gene expression.* Cancer cell, 2012. **21**(4): p. 563-76.
356. Du, B., et al., *MicroRNA-545 suppresses cell proliferation by targeting cyclin D1 and CDK4 in lung cancer cells.* PLoS One, 2014. **9**(2): p. e88022.
357. Katakowski, M., et al., *Exosomes from marrow stromal cells expressing miR-146b inhibit glioma growth.* Cancer Lett, 2013. **335**(1): p. 201-4.
358. Xiang, M., et al., *STAT3 induction of miR-146b forms a feedback loop to inhibit the NF-kappaB to IL-6 signaling axis and STAT3-driven cancer phenotypes.* Sci Signal, 2014. **7**(310): p. ra11.
359. Al-Khalaf, H.H. and A. Aboussekhra, *MicroRNA-141 and MicroRNA-146b-5p Inhibit the pro-Metastatic Mesenchymal Characteristics through the RNA Binding Protein AUF1 Targeting the Transcription Factor ZEB1 and the Protein Kinase AKT.* J Biol Chem, 2014.
360. Xia, H., et al., *microRNA-146b inhibits glioma cell migration and invasion by targeting MMPs.* Brain Res, 2009. **1269**: p. 158-65.
361. Li, J., et al., *Transcriptional regulation of miR-146b by C/EBPbeta LAP2 in esophageal cancer cells.* Biochem Biophys Res Commun, 2014. **446**(1): p. 267-71.
362. Wu, P.Y., et al., *Low expression of microRNA-146b-5p and microRNA-320d predicts poor outcome of large B-cell lymphoma treated with cyclophosphamide, doxorubicin, vincristine, and prednisone.* Hum Pathol, 2014. **45**(8): p. 1664-73.

APPENDIX I

Leukemia (2013), 1–4

www.nature.com/leu

LETTER TO THE EDITOR

Novel TAL1 targets beyond protein-coding genes: identification of TAL1-regulated microRNAs in T-cell acute lymphoblastic leukemia

Leukemia advance online publication, 26 March 2013;
doi:10.1038/leu.2013.63

The basic helix-loop-helix transcription factor TAL1 is aberrantly expressed in a majority of T-cell acute lymphoblastic leukemia (T-ALL) cases characterized by arrested development in the thymic late cortical stage.[1,2] Although TAL1 is a bona fide T-cell oncogene,[3] with known direct targets in T-ALL,[4] the aberrant transcriptional circuitry responsible for thymocyte transformation is not yet fully understood. MicroRNAs are small, non-coding RNAs that function as endogenous post-transcriptional repressors of protein-coding genes by binding to target sites in the 3′-UTR of messenger RNAs.[5] Aberrant expression of these molecules has been reported in several hematological malignancies, and microRNA expression signatures delineate ALL subgroups[6] and can be interpreted in light of their variation during hematopoiesis.[7] Individual microRNAs and networks have been implicated in T-ALL,[8] but the mechanisms responsible for altered microRNA expression in this malignancy remain poorly explored. Here, we report the identification of novel, non-protein-coding TAL1 target genes, implicating microRNA genes as part of the transcriptional network downstream of TAL1 that may be putatively involved in its oncogenic properties.

To identify a TAL1-dependent microRNA gene expression profile, we ectopically expressed TAL1 in the TAL1-negative T-ALL cell line P12 and performed low-density array analysis (see Supplementary Data online for materials and methods). From 204 detected microRNAs (out of 372 analyzed), we identified eight whose expression changed significantly upon TAL1 overexpression (Figure 1a and Supplementary Tables 1 and 2). Subsequent validation was performed by quantitative PCR analysis of each microRNA after enforcing or silencing the expression of TAL1. This allowed us to confirm the expected TAL1-mediated regulation for five microRNAs: namely, miR-135a, miR-223 and miR-330-3p as being upregulated by TAL1; and miR-146b-5p and miR-545 as being downregulated (Figures 1b–f). The three remaining microRNAs were excluded from subsequent analyses, as we stringently considered only those genes to be validated whose expression was regulated in the predicted manner upon both TAL1 overexpression and silencing (data not shown).

Next, we evaluated whether the validated microRNAs were direct targets of TAL1 in T-ALL cells. For this purpose, we scrutinized publicly available TAL1 ChIP-seq data (GEO accession number GSE29181) for two T-ALL cell lines (JURKAT and CCRF-CEM) and two primary T-ALL samples[4] for the presence of TAL1-binding peaks up to 10 kb upstream of the transcription start site of each microRNA gene. We identified one peak in a putative promoter region for miR-146b (Supplementary Figure 1a), suggesting that this gene may be a transcriptional target of TAL1. Furthermore, two peaks were observed upstream of miR-223 transcription start site (Supplementary Figure 1b). To confirm these findings, we performed TAL1 ChIP-quantitative PCR in JURKAT and CCRF-CEM cells using primers designed for the genomic areas covered by the two peaks in the miR-223 locus. We confirmed that there is more than twofold enrichment, as compared with a mock ChIP performed against fibrillarin, in the amplified area within 3.5 kb upstream of the miR-223 transcription start site (Figure 1g). These results indicate that miR-223 is a direct target of TAL1 in T-ALL. Interestingly, TAL1 appears to bind to a previously described region containing a conserved proximal genomic element with possible binding sites for the transcription factor C/EBP.[9] We did not find evidence from the available TAL1 ChIP-seq data for direct binding of TAL1 to the remaining microRNA genes, suggesting that miR-135a, miR-330-3p and miR-545 are indirectly regulated by TAL1, at least in the T-ALL cells analyzed.

Interestingly, analysis of microRNA gene expression profiles in different T-ALL subsets[8] revealed that TAL/LMO primary samples (integrating Sil-Tal1 + and LMO + cases, which frequently express high TAL1 levels) display higher levels of miR-223 ($P = 0.035$) and tend to express lower levels of miR-146b-5p ($P = 0.092$) than other T-ALL cases (Supplementary Figure 2). In line with these observations, miR-223 appears to follow the same pattern of expression along normal human thymocyte development as TAL1,[10] with high levels in $CD34^+$ T-cell precursors and sharp downregulation in more differentiated subsets (Supplementary Figure 3a). A similar pattern was observed for miR-135a (Supplementary Figure 3b), in agreement with the notion that TAL1 positively regulates both genes. In contrast, miR-146b-5p is clearly upregulated in the double-positive to single-positive transition and is amongst the most upregulated microRNAs in mature, single-positive thymocytes.[11] The fact that miR-146b-5p levels associate with thymocyte maturation (Supplementary Figure 4) is in agreement with a model whereby TAL1 overexpression during leukemogenesis inhibits miR-146b-5p and promotes T-cell developmental arrest.

Data from the analysis of congruent putative interactions between known TAL1-regulated protein-coding genes and the validated microRNA genes are in line with the notion that the latter could be part of downstream networks collaborating in TAL1-mediated leukemogenesis (Figure 2). Indeed, most TAL1-upregulated genes that have 3′-UTRs predicted as targets for the TAL1-downregulated miR-146b-5p and -miR-545 have a known or putative oncogenic function (Figure 2a). For example, CD53 was shown to protect JURKAT cells from apoptosis, PDE3B appears to be involved in glucocorticoid resistance in CEM cells and ETS-1 participates in the T-cell maturation arrest mediated by *TLX* genes in T-ALL (see Supplementary Table 3 for details and references). Interestingly, the T-ALL-associated oncogene *MYB* was recently shown to be a direct TAL1 target forming a feed-forward loop involved in the TAL1-dependent leukemogenic program.[4] Our bioinformatics analyses now raise the possibility that TAL1 may reinforce MYB upregulation by inhibiting the expression of miR-545. Also of note, three of the four genes (*KRT1*, *Rapgef5*, *JAZF1*) with predicted 3′-UTR seed sequences for both miR-146b-5p and miR-545 are associated with protumoral functions (Supplementary Table 3). In sharp contrast, the TAL1-downregulated genes that are predicted targets for miR-135a, miR-223 and miR-330-3p display a

Accepted article preview online 1 March 2013

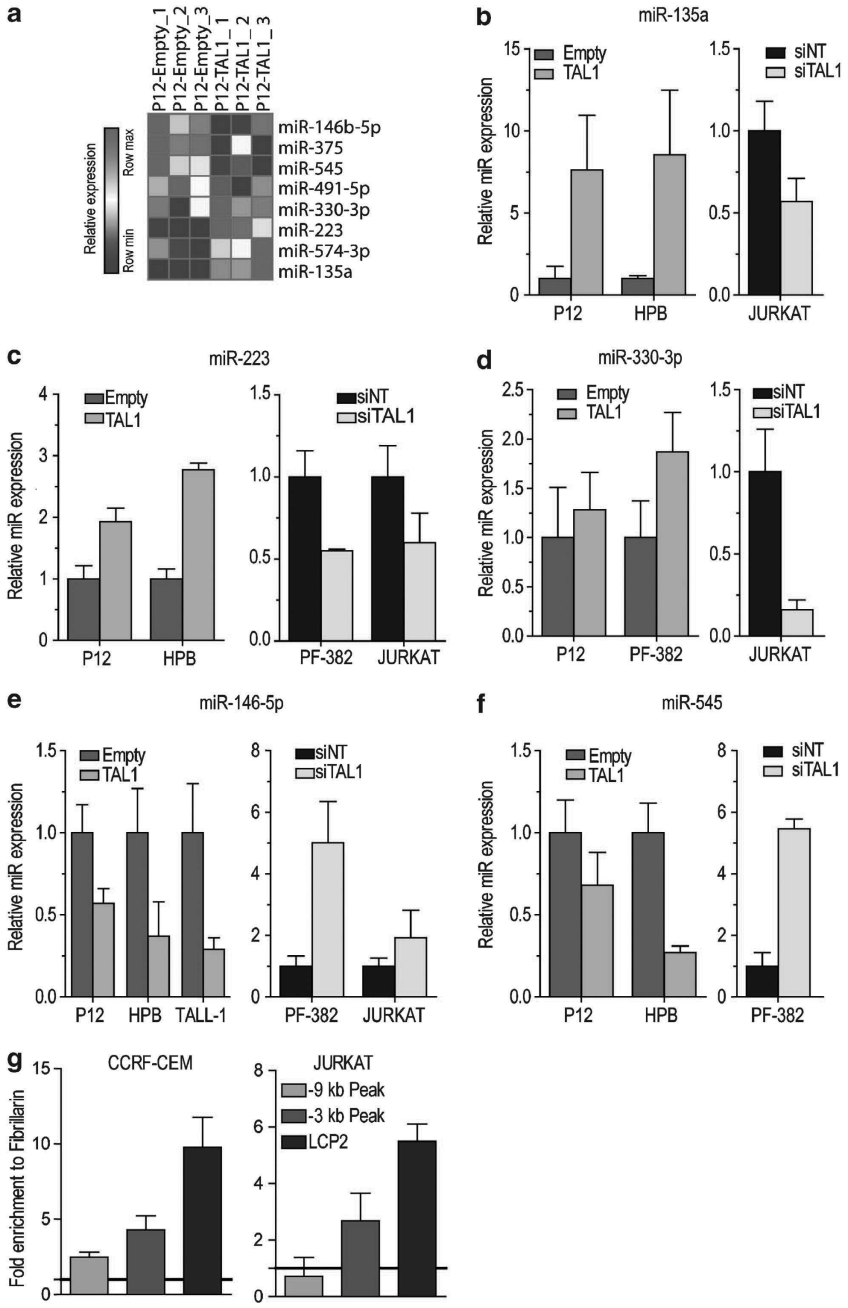


Figure 1. Identification of TAL1-regulated microRNA genes. (**a**) Heat map of differentially expressed microRNAs upon TAL1 overexpression. MicroRNAs were hierarchically clustered (rows, microRNAs; columns, experiments). See Supplementary Table 2 for fold-difference values. Levels greater than or less than the mean are shown in shades of red or blue, respectively. (**b–f**) Quantitative PCR (qPCR) validation of microRNA expression modulation by TAL1. Relative expression of hsa-miR-135a (**b**), hsa-miR-223 (**c**), hsa-miR-330-3p (**d**), hsa-miR-146b-5p (**e**), and hsa-miR-545 (**f**) normalized to SNORD38B in T-ALL cell lines with overexpression (left) or knockdown of TAL1 (right). The bars represent the mean ± s.d. of three independent replicates. siNT—non-targeting siRNA (**g**) TAL1 ChIP-qPCR in T-ALL cell lines. The occupancy by TAL1 of the genomic regions 9.2 and 3.5 kb upstream of miR-223 transcription start site was analyzed by ChIP-qPCR in JURKAT and CCRF-CEM cells. The promoter region of LCP2 was used as positive control for TAL1 binding, and a random intergenic region was used as negative control. TAL1 binding is expressed as the fold enrichment relative to a mock ChIP performed against fibrillarin. The error bars represent the 95% CI of the fold enrichment. The horizontal line denotes the fold enrichment detection for the negative control.

clear abundance in (putative) tumor suppressors or in genes whose functions are compatible with antitumoral effects (Figure 2b and Supplementary Table 4). This is evident, for instance, in the case of the four genes potentially regulated by two microRNAs, in which only one has an oncogenic role (*IGF1R*) and three likely have tumor-suppressive functions (*SRGAP3*, *TOX*, *LRP12*).

Circumscription of our analysis to validate target genes of each microRNA, followed by gene set enrichment analysis showed an enrichment in biological processes related to inflammation (e.g., NF-kB signaling pathway and IL1/IL1R signaling pathway) and cancer (e.g., pathways in cancer), as detailed in Supplementary Table 5 and Supplementary Table 6. Interestingly, the validated targets for the TAL1-downregulated gene of miR-146-5p include *IRAK1*, *TRAF6* and *NFKB1* (all of which are involved in chronic inflammation), as well as the oncogene *KIT* (Supplementary Table 5 and Supplementary Figure 5). In contrast,

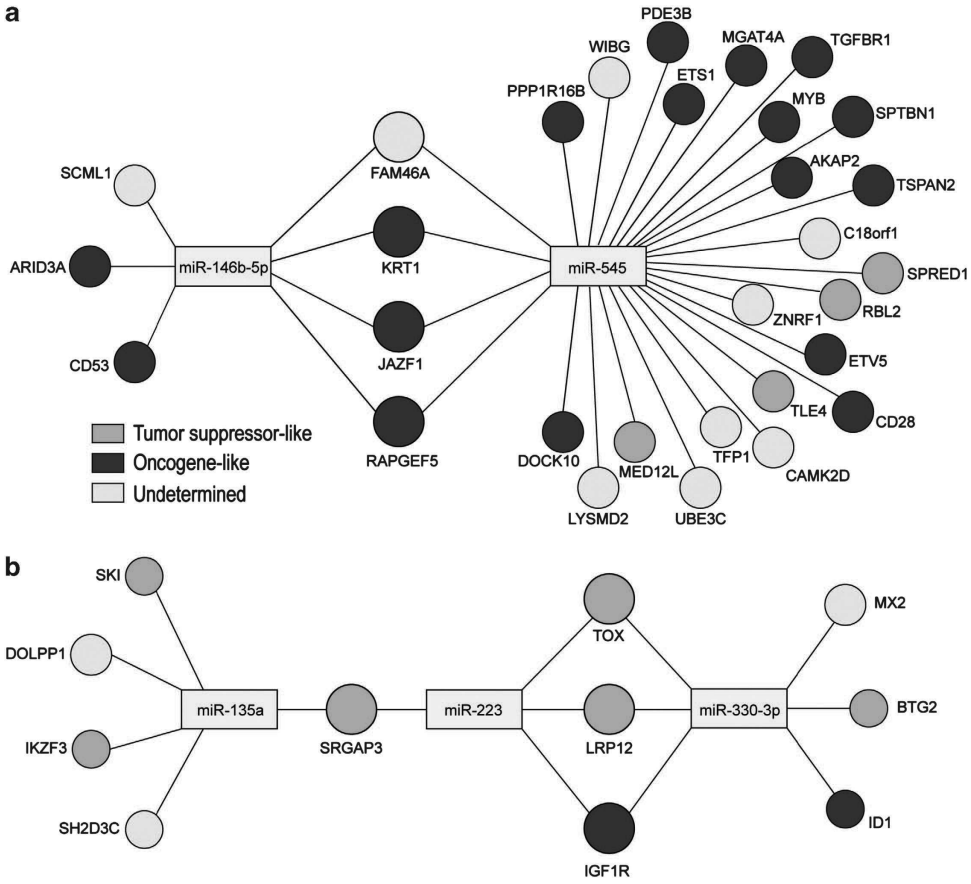


Figure 2. Potential participation of the newly identified TAL1 microRNA target genes in TAL1-mediated leukemogenic pathways. Cross-examination of congruent TAL1-regulated protein-coding and miRNA genes was performed as described in Supplementary Methods. (**a**) Downregulated miRNAs and their predicted target genes previously shown to be upregulated by TAL1. (**b**) Upregulated miRNAs and their predicted target genes previously shown to be downregulated by TAL1. Genes are color-coded according to their reported function in the context of cancer, as detailed in Supplementary Tables 3 and 4. 'Tumor suppressor-like': bona fide or putative tumor suppressors or genes that have proapoptotic, antiproliferative or prodifferentiating roles; 'Oncogene-like': bona fide or putative oncogenes or genes that have antiapoptotic or proliferative roles. 'Undetermined': genes with undetermined function or whose role in cancer remains unknown.

miR-330-3p, upregulated by TAL1, reportedly targets *E2F1* and *CDC42*, both of which are described to promote apoptosis in different cell types, including Jurkat cells. MiR-223 is a myeloid-specific microRNA essential for normal neutrophil maturation, and responsible for granulocyte differentiation and negative regulation of progenitor proliferation via MEF2C downregulation.[12] In agreement, miR-223 functions as a tumor suppressor in acute myeloid leukemia,[13] and appears to be repressed in chronic myeloid leukemia, allowing for the expression of MEF2C.[14] In contrast, miR-223 is frequently overexpressed in T-ALL, cooperating with NOTCH1 to accelerate the onset of disease in a Notch-induced leukemia mouse model. This effect was proposed to be due, at least in part, to inhibition of FBXW7, a negative regulator of Notch signaling.[8] However, FBXW7 targets the degradation of other oncogenic proteins, such as c-Myc and mTOR, and the expression of miR-223 is significantly elevated in TAL1-positive T-ALL cases (Supplementary Figure 2), suggesting that the oncogenic function of this microRNA may extend beyond mere collaboration in Notch-induced leukemia. Moreover, the proleukemic role of mir-223 may be also achieved by downregulating targets such as E2F1, FOXO1, RHOB or EPB41L3, which have been associated with induction of apoptosis and/or have tumor-suppressive roles (Supplementary Table 5). Interestingly, the intriguing possibility that miR-223 may potentially act downstream of TAL1 to negatively regulate MEF2C, recently identified as an oncogene in T-ALL,[15] would be in line with the observations that TAL1 and MEF2C tend to segregate, defining two discrete T-ALL subsets.[15]

In summary, our studies identify and validate for the first time a small set of TAL1-regulated microRNA genes whose role may be important in the context of hematopoiesis and T-cell leukemogenesis.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This study was supported by grants from Fundação para a Ciência e a Tecnologia (FCT; PTDC/BIM-ONC/1548/2012) and Liga Portuguesa Contra o Cancro (Terry Fox Award) to JTB; and FWO Vlaanderen (G056413N) and UGent (GOA 01G01910W) to FS. NCC and KD have PhD fellowships from FCT and IWT Vlaanderen, respectively. PR has a post-doctoral grant from FWO Vlaanderen. We would like to thank Steve Lefever for primer design.

NC Correia[1], K Durinck[2], AP Leite[1], M Ongenaert[2], P Rondou[2], F Speleman[2], FJ Enguita[1] and JT Barata[1]
[1]*Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisbon, Portugal and*
[2]*Center for Medical Genetics, Department of Pediatrics and Genetics, Ghent University, Ghent, Belgium*
E-mail: joao_barata@fm.ul.pt

REFERENCES

1 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.
2 Cardoso BA, de Almeida SF, Laranjeira AB, Carmo-Fonseca M, Yunes JA, Coffer PJ *et al.* TAL1/SCL is downregulated upon histone deacetylase inhibition in T-cell acute lymphoblastic leukemia cells. *Leukemia* 2011; **25**: 1578–1586.
3 O'Neil J, Shank J, Cusson N, Murre C, Kelliher M. TAL1/SCL induces leukemia by inhibiting the transcriptional activity of E47/HEB. *Cancer Cell* 2004; **5**: 587–596.
4 Sanda T, Lawton LN, Barrasa MI, Fan ZP, Kohlhammer H, Gutierrez A *et al.* Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. *Cancer Cell* 2012; **22**: 209–221.
5 Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004; **116**: 281–297.
6 Schotte D, Pieters R, Den Boer ML. MicroRNAs in acute leukemia: from biological players to clinical contributors. *Leukemia* 2012; **26**: 1–12.
7 Fabbri M, Croce CM, Calin GA. MicroRNAs in the ontogeny of leukemias and lymphomas. *Leuk Lymphoma* 2009; **50**: 160–170.
8 Mavrakis KJ, Van Der Meulen J, Wolfe AL, Liu X, Mets E, Taghon T *et al.* A cooperative microRNA-tumor suppressor gene network in acute T-cell lymphoblastic leukemia (T-ALL). *Nat Genet* 2011; **43**: 673–678.
9 Fukao T, Fukuda Y, Kiga K, Sharif J, Hino K, Enomoto Y *et al.* An evolutionarily conserved mechanism for microRNA-223 expression revealed by microRNA gene profiling. *Cell* 2007; **129**: 617–631.
10 Pike-Overzet K, de Ridder D, Weerkamp F, Baert MR, Verstegen MM, Brugman MH *et al.* Ectopic retroviral expression of LMO2, but not IL2Rgamma, blocks human T-cell development from CD34 + cells: implications for leukemogenesis in gene therapy. *Leukemia* 2007; **21**: 754–763.
11 Ghisi M, Corradin A, Basso K, Frasson C, Serafin V, Mukherjee S *et al.* Modulation of microRNA expression in human T-cell development: targeting of NOTCH3 by miR-150. *Blood* 2011; **117**: 7053–7062.
12 Johnnidis JB, Harris MH, Wheeler RT, Stehling-Sun S, Lam MH, Kirak O *et al.* Regulation of progenitor cell proliferation and granulocyte function by microRNA-223. *Nature* 2008; **451**: 1125–1129.
13 Pulikkan JA, Dengler V, Peramangalam PS, Peer Zada AA, Muller-Tidow C, Bohlander SK *et al.* Cell-cycle regulator E2F1 and microRNA-223 comprise an autoregulatory negative feedback loop in acute myeloid leukemia. *Blood* 2010; **115**: 1768–1778.
14 Agatheeswaran S, Singh S, Biswas S, Biswas G, Pattnayak NC, Chakraborty S. BCR-ABL mediated repression of miR-223 results in the activation of MEF2C and PTBP2 in chronic myeloid leukemia. *Leukemia* 2012; e-pub ahead of print 21 December 2012. doi:10.1038/leu.2012.339.
15 Homminga I, Pieters R, Langerak AW, de Rooi JJ, Stubbs A, Verstegen M *et al.* Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia. *Cancer Cell* 2011; **19**: 484–497.

Supplementary Information accompanies this paper on the Leukemia website (http://www.nature.com/leu)