Tesis Doctoral

Biological and functional characterization of the fusion gene *NUP98-HOXA9* in Acute myeloid leukemia: New therapeutic approaches

ANA DEL RIO MACHIN
Madrid 2014

Departamento de Bioquímica

Facultad de Medicina

Universidad Autónoma de Madrid - UAM

UAM
UNIVERSIDAD AUTONOMA
DE MADRID

Biological and functional characterization of the fusion gene *NUP98-HOXA9* in Acute myeloid leukemia: New therapeutic approaches

Tesis Doctoral presentada por **Ana del Río Machín**

Licenciada en Biología y Licenciada en Bioquímica por la Universidad de Salamanca (USAL), Máster en Biomedicina Molecular por la Universidad Autónoma de Madrid (UAM)

Director:
Juan Cruz Cigudosa García

Grupo de Citogenética Molecular
Programa de Genética del Cáncer Humano
Centro Nacional de Investigaciones Oncológicas (CNIO)

Dr. Juan Cruz Cigudosa García, Jefe del Grupo de Citogenética Molecular del Programa de Genética del Cáncer Humano del Centro Nacional de Investigaciones Oncológicas (CNIO), como director de la Tesis,

CERTIFICA:

Que Doña Ana del Río Machín ha realizado el presente trabajo: "**Biological and functional characterization of the fusion gene *NUP98-HOXA9* in Acute myeloid leukemia: New therapeutic approaches**" y que a su juicio reúne plenamente todos los requisitos necesarios para optar al grado de Doctor, a cuyos efectos será presentado en la Universidad Autónoma de Madrid. El trabajo ha sido realizado bajo su dirección, autorizando su presentación ante el Tribunal Calificador.

Y para que conste, se extiende el presente certificado.

Madrid, Octubre 2014

VºBº del Director

Juan Cruz Cigudosa García

CONTENTS

PREFACE

Initially, this doctoral thesis had the major and ambiguous aim of gaining a deeper knowledge of the genetic and molecular mechanisms that underlie the leukemogenesis process. We explored and carried investigations on three types of hematological malignancies. Since the work done in two of them has been already published and communicated in meetings, we have chosen for this PhD dissertation, and its corresponding public defence, the more scientifically challenging project entitled "Biological and functional characterization of the fusion gene NUP98-HOXA9 in Acute myeloid leukemia: New therapeutic approaches". A brief description of the other two is presented below:

1. ***Abrogation of RUNX1 gene expression in de novo myelodysplastic syndrome with t(4;21)(q21;q22)***

(This work was published in *Haematologica* in April 2012, Appendix II)

We describe the molecular characterization of a new t(4;21)(q21;q22) in a de novo myelodysplastic syndrome that resulted in the deletion of the *RUNX1* gene. We demonstrated by quantitative real-time RT-PCR an almost complete depletion of the expression of the *RUNX1* gene in our t(4;21) case, compared with CD34 + cells, that was independent of mutation or DNA methylation. We explored and confirmed the fact that this abrogation also prevents transactivation of *RUNX1* target genes, shedding some light into the genetic origin of the thrombocytopenia and the myelodysplastic features observed in our patients, and certainly mimicking what has been observed in the presence of the RUNX1/ETO fusion protein.

2. *The downregulation of specific miRNAs in hyperdiploid multiple myeloma mimics the effect of the most frequent IgH translocations observed in the non-hyperdiploid subtype*.

(This work was published in *Leukemia* in April 2013; Appendix II)

We compare the micro RNA (miRNA) expression profiles of the two big groups of multiple myeloma (MM): non-hyperdiploid (nh-MM) group, highly enriched for IgH translocations, and hyperdiploid (h-MM), typically characterized by trisomies of some odd-numbered chromosomes. We found that target genes of the most differentially expressed miRNAs are directly involved in the pathogenesis of MM; specifically, the inhibition of *hsa-miR-425*, *hsa-miR-152* and *hsa- miR-24*, which are all downregulated in h-MM, leads to the overexpression of *CCND1*, *TACC3*, *MAFB*, *FGFR3* and *MYC*, which are the also the oncogenes upregulated by the most frequent IgH chromosomal translocations occurring in nh-MM. Importantly, we validate these results in primary cases of h-MM. These data establish the importance of miRNA deregulation in the context of MM, thereby opening up the potential for future therapeutic approaches based on this molecular mechanism.

RESUMEN/ABSTRACT

RESUMEN

La translocación cromosómica t(7;11)(p15;p15), que origina la proteína de fusión NUP98-HOXA9 (NH), aparece como evento primario en el 1% de los pacientes con Leucemia Mieloide Aguda (LMA) y está asociada con una forma muy agresiva de leucemia (mal pronóstico y baja supervivencia global). Aunque se ha sugerido un papel como factor de transcripción oncogénico, se desconoce el mecanismo molecular mediante el cual NH tiene estos efectos. Además, no existe una terapia específica para estos pacientes, que son tratados con regímenes de quimioterapia estándar muy ineficaces. Por ello, nuestro objetivo ha sido generar modelos celulares humanos que expresen la proteína quimera de forma constitutiva, en los que poder estudiar su papel en la transformación leucémica, así como ampliar la búsqueda de nuevas dianas terapéuticas que permitan tratar de forma dirigida a pacientes que presentan esta translocación cromosómica. Así, hemos identificado por primera vez los sitios de unión al ADN de NH, la mayoría de los cuales son regiones génicas *potenciadoras* que permiten regular la expresión de genes implicados en el desarrollo de la LMA. En concreto, hemos demostrado que NH induce directamente la sobreexpresión del complejo MEIS1-HOXA9-PBX3, un elemento clave en la biología de leucemias originadas por otros reordenamientos cromosómicos. Para confirmar la importancia de este complejo también en este tipo de LMA, demostramos que el péptido HXR9, un inhibidor de la interacción HOXA9-PBX3, es capaz de matar selectivamente a las células que expresan NH. Además, observamos que, tanto en un modelo celular humano que expresa NH generado a partir de progenitores hematopoyéticos (hHP-NH) como en pacientes con la translocación cromosómica, la proteína de fusión es capaz de activar y reprimir la expresión de sus genes diana. De hecho, demostramos que la interacción de NH con p300 (un activador transcripcional) y con HDAC1 (un inhibidor transcripcional) puede explicar este papel activador-represor de la proteína de fusión. Apoyando esta hipótesis, encontramos un efecto inhibitorio dramático de la viabilidad de hHP-NH con el tratamiento con LBH589, un inhibidor de HDACs, lo que nos lleva a considerar su potencial terapéutico en esta leucemia. Finalmente, en un intento de comprender la contribución de las dos partes que forman la proteína de fusión en la regulación de la expresión génica, observamos que un tercio de los genes diana de NH son comunes a HOXA9. Así mismo, de forma sorprendente, hemos descubierto que también NUP98 parecía funcionar como un factor de transcripción implicado en la regulación de la hematopoiesis. El nuevo papel descubierto para esta nucleoporina ayudará a comprender mejor cómo funciona NH, así como otras fusiones leucémicas en las que también interviene NUP98. En resumen, estos resultados nos han permitido identificar algunos de los mecanismos patogénicos más relevantes inducidos por NH e identificar nuevas dianas terapéuticas que han sido estudiadas en los modelos celulares, con resultados muy prometedores.

ABSTRACT

The chromosomal translocation t(7;11)(p15, p15), which results in the oncogenic fusion protein NUP98-HOXA9 (NH), is a rare but recurrent oncogenic event in AML that is associated with very poor prognosis and short overall survival. The molecular processes triggered by NH are poorly understood, even though a potential role as an aberrant transcription factor has been proposed. Moreover, there is no specific therapy for these patients; they are treated with ineffective standard chemotherapy regimens. Therefore, our main objective has been to generate human cellular models that constitutively express the chimera protein, to study its role in leukemic transformation and expand the search for new therapeutic targets that enable the treatment of the patients with this chromosomal translocation in a targeted manner. Using these models, we have described for the first time the DNA binding sites of NH, most of which are enhancers that regulate the expression of genes involved in the development of AML. In particular, we have shown that the direct overexpression of the complex MEIS1-HOXA9-PBX3, a key element in the onset of leukemia that is driven by other chromosome rearrangements in AML, is also one of the actionable pathogenic mechanisms induced by NH. We have demonstrated that the peptide HXR9, an inhibitor of the interaction HOXA9-PBX3, was indeed able of selectively kill cells expressing NH. Furthermore, we have observed that, in a model of human hematopoietic progenitors expressing NH, and also in patients with the chromosomal translocation, the fusion protein was able to activate and repress the expression of their target genes. We have also shown that the interaction of NH with p300 (transcriptional activator) and HDAC1 (transcriptional inhibitor), could explain this activator-repressor role of the fusion protein. Supporting this hypothesis, we have found a dramatic inhibitory effect on hHP-NH viability after the treatment with LBH589 (an HDAC inhibitor) which allowed us to consider its therapeutic potential for patients carrying the translocation. Finally, in an attempt to understand the contribution of each of the two NH moieties to the regulation of gene expression, we found that a third of the target genes of NH are common to HOXA9 wt. In addition, ChIP-seq results for NUP98 wt revealed the potential role of this nucleoporin in regulating hematopoietic differentiation. This newly role of NUP98 is likely to contribute to a better understanding of how NH, as well as other leukemic fusion proteins in which NUP98 is involved, works.

To summarize, these results have allowed us to characterize some of the most relevant pathogenic mechanisms induced by NH and to identify new therapeutic targets that have been studied in our model, yielding very promising results.

ABBREVIATIONS

Ac - Acetylation

AES - Amino-terminal Enhancer of Split protein

AML - Acute Myeloid Leukemia

ANWL - Ala-Asn-Trp-Leu motif

APL - acute promyelocytic leukemia

Ara-C - Cytarabine (arabinofuranosyl cytidin)

AUL - acute undifferentiated leukemia

B-ALL - B-cell acute lymphoblastic leukemia

BFU-E - Burst Forming Unit-erythroid

CA-AML - Cytogenetically abnormal AML

CBF - Core binding factor

CFU - Colony Forming Unit

CFU-G - Colony Forming Unit-granulocyte

CFU-GEMM - Colony Forming Unit-granulocyte, erythrocyte, macrophage, megakaryocyte)

CFU-GM - Colony Forming Unit-granulocyte, macrophage

CFU-M - Colony Forming Unit-macrophage

ChIP - Chromatin Immunoprecipitation

ChIP-seq - Chromatin Immunoprecipitation Sequencing

CML - Chronic Myeloid Leukemia

CML-BC - chronic myelogenous leukemia in blast crisis

Co-IP - Co-immunoprecipitation

CR - Complete Remission

eRNA - Enhancer RNA

FDA - US Food and Drug Administration

FDR - False Discovery Rate

FG repeats - Phe-X-Phe-Gly or Gly-Leu-Phe- Gly amino acid residues

GBD - GLEBS binding domain

GLEBS - Gle2-binding sequence

GLFG - Gly-Leu-Phe- Gly

GSEA - Gene Set Enrichment Analysis

H3 - Histone 3

HATs - Histone acetyltransferases

HD - Homeodomain

HDACi - HDAC inhibitor

HDACs - Histone deacetylases

hHP - Human Hematopoietic Progenitor

IPA - Ingenuity Pathways Analysis

K - Lysine

LC50 - Median Lethal Dose

LOH - Loss of Heterozygosity

MD - MEIS domain

MDS - Myelodysplastic Syndrome

Me - methylation

NH - NUP98-HOXA9

NPC - Nuclear Pore Complex

Ph - Philadelphia Chromosome

PM - PBX motif

PolII - Polymerase II

qPCR - Quantitative PCR

R9 - 9 Arginine residues

RBD - RNA binding domain

T-ALL - T-cell Acute Lymphoblastic Leukemia

t-AML - therapy-related acute myeloid leukemia

t-MDS - therapy-related myelodysplastic syndrome

TSS - Transcription Start Site

WBC - White Blood Count

wt - Wild Type

TABLES AND FIGURES

TABLES

Table 1: Standardized correlation of cytogenetic and molecular genetic data with prognosis data in AML

Table 2: Types of *NUP98-HOXA9* transcripts

Table 3: Nucleoporin gene rearrangements in hematologic malignancies

Table 4: qChIP primers

Table 5: Sybr Green qPCR primers

Table 6: Target miRNAs of NUP98-HOXA9 identified in the ChIP-seq

Table 7: Identification of DNA binding motifs in NH target regions using MEME-ChIP.

Table 8: Identification of known transcription factor binding motifs in NH target regions using oOPOSSUM

Table S1: NUP98-HOXA9 target genes identified by ChIP-seq (see data on CD-room)

Table S2: Co-occurrences of HOXA9-MEIS1-PBX binding motif in the DNA target sequences of NH (FIMO results) (see data on CD-room)

Table S3: HOXA9 target genes identified by ChIP-seq (see data on CD-room)

Table S4: NUP98 target genes identified by ChIP-seq (see data on CD-room)

Table S5: Expression arrays in cellular models (see data on CD-room)

Table S6: Expression arrays in patients (see data on CD-room)

Table S7: Interacting-proteins indentify by Mass spectrometry assay

FIGURES

Figure 1: Cellular origin of Acute Myeloid Leukemia

Figure 2: Integrated model for the origin and evolution of the oncogenic events in AML

Figure 3: t(7;11)(p15;p15) chromosomal translocation

Figure 4: The nuclear pore complex (NPC)

Figure 5: Schematic representation of the NUP98 protein

Figure 6: Clustered HOX genes

Figure 7: Schematic representation of the HOXA9 protein

Figure 8: Schematic representation of NUP98, HOXA9 and NUP98-HOXA9 proteins

Figure 9: Fusion point sequence in each type of transcript

Figure 10: Subcellular localization of NUP98-HOXA9

Figure 11: Structure of enhancers

Figure 12: Overview of the bioinformatics tools used for the motif analysis

Figure 13: NUP98-HOXA9 human cellular models construction

Figure 14: Identification of NUP98-HOXA9 genomic binding regions

Figure 15: Identification of *HOXA9* and *NUP98 wt* genomic binding regions

Figure 16: Nup98 plays a role in hematopoiesis

Figure 17: NUP98-HOXA9 target genes

INTRODUCTION

1 ACUTE MYELOID LEUKEMIA

Acute Myeloid Leukemia (AML) is a highly malignant hematopoietic tumor that is characterized by the uncontrolled proliferation, increased survival and impaired differentiation of the hematopoietic myeloid progenitor cells. These abnormal immature leukemic cells, known as blasts, accumulate in blood and bone marrow. The presence of 20% or more of blasts in bone marrow is the main criteria for the diagnosis of AML[1]. This causes the disruption of normal hematopoiesis, which results in fatigue, infection, bleeding and multiorgan failure[2]. (Figure 1A and 1B)

AML represents approximately 25% of all leukemia diagnosed in adults with a median age range that goes from 66 to 71 years[3]. On average, the age-standardized incidence rate for AML is 2.95 cases per 100 000 in Europe, and 4.6 cases in USA. However, the most worrying aspect of this disease is its high mortality rate, with a 5-year Overall Survival (OS) rate of 40% to 45% (OS in older patients still remains poor at < 10% after 5 years)[4]. Thus, almost 70% of patients on average succumb to the disease, representing more than 9,000 deaths each year alone in the United States[5].

A

B
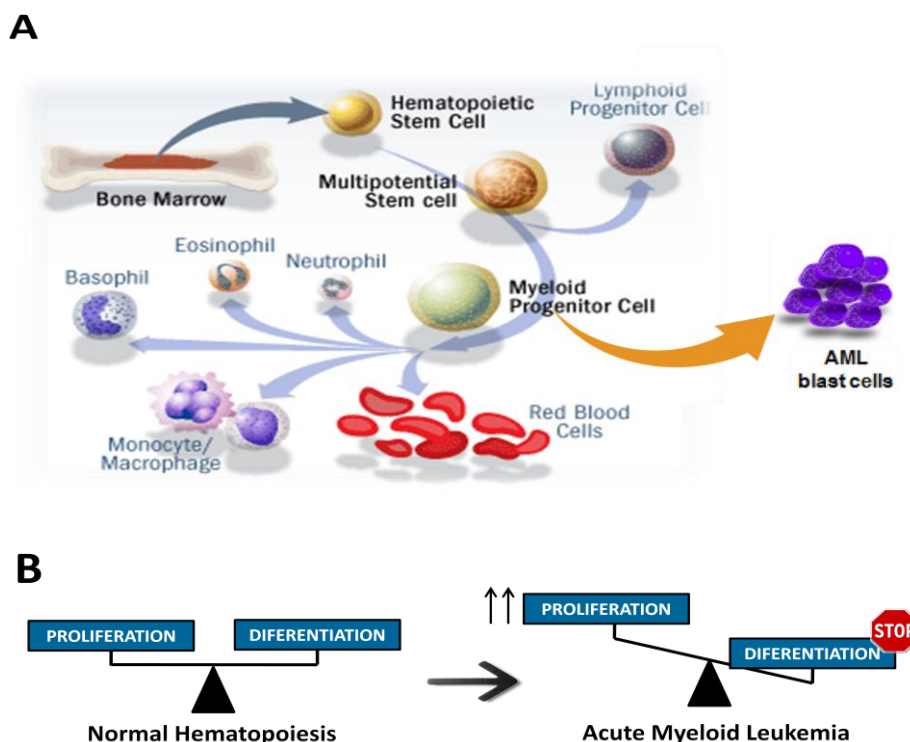
Figure 1: Cellular origin of Acute Myeloid Leukemia (A) Shows the differentiation of normal lymphoid and myeloid lineage from hematopoietic stem cells. Yellow arrow points at the abnormal undifferentiated leukemic blast cells. (Modified from *How stem cells work* by Stephanie Watson) (B) Shows how in AML, the normal hematopoietic balance between proliferation and differentiation is disrupted and an increase in proliferation with a blockage of differentiation is observed.

1.1 Genetic origin and evolution of AML

The production of blood is a tightly controlled process in which transcription factors and chromatin remodeling genes are necessary for the orchestrated regulation of gene expression that defines the phenotype of a given blood cell[6].

1.1.1 The "Two hits hypothesis"

For many years, the accepted model of leukemogenesis was the "two hit hypothesis". This model suggests that mutations of a single gene are rarely sufficient to cause the development of AML and that instead, two different types of genetic mutation are required for malignant transformation of a myeloid precursor: Class I and class II. Class I mutations lead to increased proliferation, survival, or evasion of apoptosis, and include mutations conferring constitutive activity to tyrosine kinases or deregulation of downstream signaling molecules, such as BCR-ABL or mutations in *FLT-3*, *c-KIT*, and *RAS*. Class II mutations impair differentiation or enhance self-renewal of hematopoietic progenitors and include the translocations associated with the core-binding factor (CBF) or mutations in genes that are involved in transcriptional regulation, such as *p300, CBP, MOX, TIF2*, and *MLL*[7,8]. However, recent research highlighting the presence of novel mutations in genes that are related to epigenetic control of the genome and modifications to the epigenome itself found in the AML patients suggests that Class I and Class II mutations are only one part of a more complex picture. Furthermore, that some other mutations occurring in AML are not regarded as belonging to Class I or Class II also indicates that the "two-hit" theory is an oversimplification[9]. Moreover, there is evidence that there is also a temporal component to leukemogenesis: mutations have to occur at a particular point in cell development, and in a particular order, to allow for leukemic transformation[7]. Therefore, the models for the development of AML are becoming increasingly complex.

1.1.2 New model for AML development

An example of a model that best fits the actual scenario of AML would be the one proposed by Welch *et al*[10] (Figure 2). In the first step of the leukemogenesis, a driver mutation (recurrent mutations with translational consequences or recurrent chromosomal translocations that result in leukemic fusion genes) occurs within the context of a Human Hematopoietic Progenitor (hHP) that already contains hundreds of random benign mutations that have accumulated over time. This mutation confers a proliferative advantage to the cell, allowing the formation of a clone that

expands, carrying along all of the random background mutations within its genome (passengers). Eventually, an additional driver mutation (cooperating mutation) occurs within a cell in an expanding clone and it becomes the leukemic "founding" clone that is detected at the diagnosis. Thus, these cells contain only a few drivers but many passengers. Each progression event yields a small cluster of new mutations, of which only one or two may be relevant for clonal or subclonal outgrowth.

In summary, the development of AML is associated with an accumulation of acquired genetic alterations and epigenetic changes in hematopoietic progenitor cells that alter the normal cellular mechanisms. AML constitutes an exceptional biological model of cooperative effects between genetic and epigenetic alterations on the transformation, progression and phenotype of a clonal neoplasia[11].
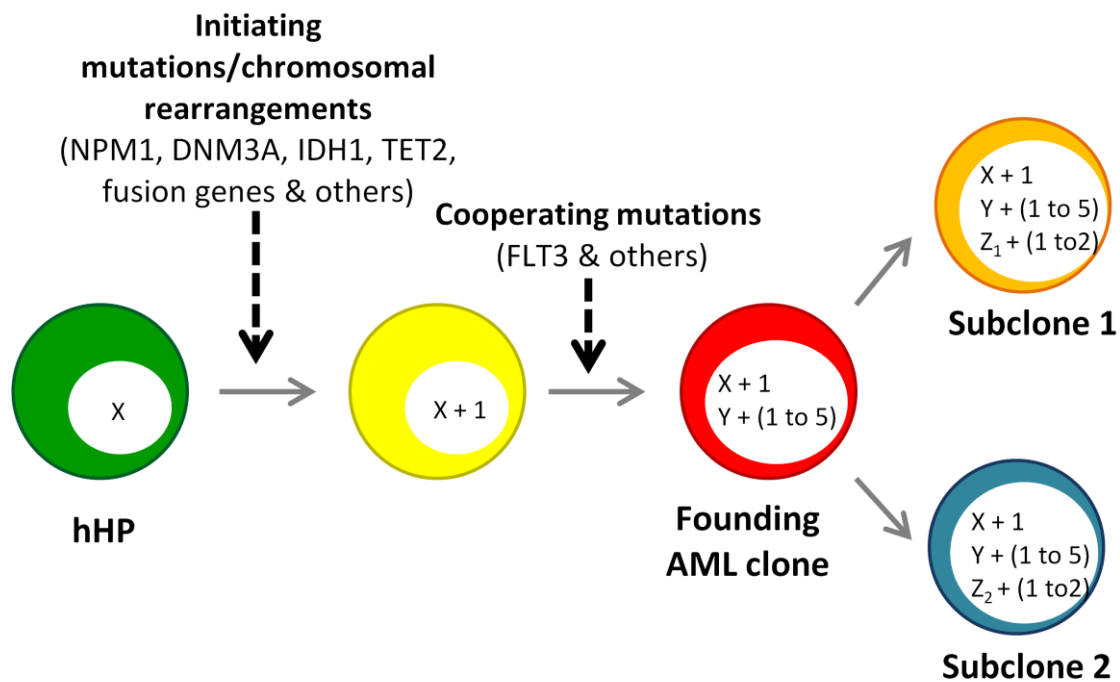
Figure 2: Integrated model for the origin and evolution of the oncogenic events in AML. Hematopoietic progenitors (hHP, shown in green) are long-lived cells that accumulate random benign background mutations as a function of age (X). The initiating mutations, which are drivers, provide an advantage for the affected cell and clonal expansion (shown in yellow). Likewise, 1–5 of cooperating mutations (Y) contribute to progression in most cases of AML, giving the expanding clone an additional advantage. Clonal outgrowth of cells with appropriate progression events results in AML, which is dominated by the "founding" AML clone, depicted in red. Subclones arise from the founding AML clone by acquiring a small number of additional mutations that confer an advantage to that cell, along with any additional background mutations (Z) that may have occurred in the interim. Therefore, *X* represents the age-dependent passenger mutations pre-existing in HSPC; *Y*, the passenger mutations gained between initiating and cooperating mutations; and *Z*, the passenger mutations gained during progression to subclone (Modified from Welch *et al.*, *Cell*, 2012).

1.2 AML prognostic factors

Although there are unifying cellular characteristics, AML remains a heterogeneous group of diseases in terms of clinical presentation, genetic alterations, prognosis and response to treatment. Classifying the different types of AML and especially identifying prognostic factors in this disease remains one of the main needs for the clinicians nowadays.

Prognostic factors may be subdivided into those related to patient characteristics and general health condition and those related to characteristics particular to the AML clone. Patient-related factors include age, as a clear adverse prognostic factor, and comorbidities, as an important aspect to be evaluated before the treatment of the patient. The AML-related factors are the white blood count (WBC), existence of prior myelodysplastic syndrome (MDS), previous cytotoxic therapy for another disorder and cytogenetic and molecular genetic changes in the leukemic cells at diagnosis. Other factors have been also included in this group, but the prognostic effect is observed with variable consistency among studies[12].

Numerous recurrent cytogenetic aberrations have been identified in AML cases that are not only diagnostic markers for specific subtypes but also constitute independent prognostic factors for response to therapy and for overall survival[5,11] In particular, recurrent chromosomal rearrangements that result in the production of fusion proteins are common initiation events in 40% of AML cases[13]. The European LeukemiaNet has proposed a standardized reporting system that correlates genetic abnormalities with clinical outcome, including data from cytogenetic analysis and from mutation analysis of *NPM1, CEBPA* and *FLT3* genes.(Table 1) This classification divides AML patients into 4 risk groups with a high variability between them (e.g. complete remission observed in 80-90% of the patients in the favorable genetic group but only in 30% of the patients in the adverse group[4]).

Different next-generation sequencing analysis of AML genomes has recently shown hundreds of separate genetic lesions within individual cases, which represents an important advance in the study of this disease[14]. Moreover, the progress in genomics technology has proven valuable for the discovery of novel leukemia subgroups and prognostic signatures. However, to date, conventional cytogenetics analysis is a mandatory, irreplaceable and the strongest component in the diagnosis and prognosis evaluation of a patient with suspected AML[12].

Table 1: Standardized correlation of cytogenetic and molecular genetic data with prognosis data in AML (Modified from Döhner *et al*[12])

GENETIC GROUP	SUBSETS
Favorable	t(8;21(q22;q22); *RUNX1-RUNX1T1* inv(16)(p13.1;q22) or t(16;16)(p13.1;q22); *CBF-MYH11* Mutated *NPM1* without *FLT3-ITD* (normal karyotype) Mutated *CEBPA* (normal karyotype)
Intermediate I	Mutated *NPM1* and *FLT3-ITD* (normal karyotype) Wild-type NPM1 and *FLT3-ITD* (normal karyotype) Wild-type *NPM1* without FLT3-ITD (normal karyotype) All AMLs with normal karyotype except those included in favorable group
Intermediate II	t(9;11)(p22;q23); *MLLT3-MLL* Cytogenetic abnormalities not classified as favorable or adverse
Adverse	inv(3)(q21q26.2) or t(3;3)(q21;q26.2); *RPN-EVI1* t(6;9)(p23;q34); *DEK-NUP214* t(v;11)(v;q23); *MLL* rearranged -5 or del(5q); -7; abnl(17p); complex karyotype

1.3 Treatment

1.3.1 Standard chemotherapy protocols

1.3.1.1 Intensive induction therapy

Patients diagnosed with AML, considered suitable for intensive chemotherapy, are treated with a combination of 3 days of anthracycline and 7 days of cytarabine ("3 + 7")[4,12]. Anthracycline (eg. Idarubicin) inserts itself into the DNA and prevents that DNA from unwinding by interfering with the enzyme topoisomerase II[15]. Cytarabine (arabinofuranosyl cytidin or ara-C) is a pyrimidine nucleoside analogue that inhibits the synthesis of DNA, by blocking DNA polymerase[16]. Complete remission (CR) can be achieved in 65% to 75% of younger patients (≤ 60 years) and in approximately 40% to 60% of older patients (> 60 years).

For AML patients who could not tolerate the intensive treatment, the therapeutic options are reduced and include best supportive care with hydroxyurea, low-dose cytarabine and hypomethylating agents such as decitabine or azacitidine. CR can be achieved in 10% to 30% of patients treated with this low-dose therapy[4].

1.3.1.2 Consolidation phase

In patients who achieve a CR after induction chemotherapy, some post-remission therapy is required to prevent relapse. This treatment is based on genetic and molecular features and can range from high-dose cytarabine to allogeneic hematopoietic stem cell transplantation[4,12].

The risk of relapse following consolidation phase is highly variable among the different prognostic groups of AML, but remains high for all cases (≈ 30% in the favorable group)[17]. Indeed, in the majority of patients with AML who achieve CR, the leukemia will recur within 3 years after diagnosis. In general, the prognosis of patients after relapse is poor and treatment options are unsatisfactory. Therefore, this scenario raises the urgent need for improved therapies for the disease.

1.3.2   Future perspectives

The association of specific cytogenetic subtypes of AML with altered molecular pathways has led to the introduction of specifically targeted therapeutics (e.g. retinoic acid in *PML-RARA* positive cases or selective *FLT3* inhibitors in cases with *FLT3* internal tandem duplication). However, these discoveries have not translated into significant advancements in survival for the majority of patients with AML that are still treated with the current standard chemotherapy regimens and most of them ultimately relapse[17]. Therefore, the identification of novel therapeutic targets for the treatment of this disease is of great importance. For this reason, it is necessary to understand completely the underlying biology for each subtype and explore whether the leukemic transformation could be mediated by common or overlapping genetic programs downstream[5]. Identifying genetic pathways that cooperate in the leukemogenesis will facilitate the understanding of the molecular mechanisms at play and the search for common critical therapeutic targets.

1.3.2.1   HDAC inhibitors:

It has been demonstrated that the epigenetic regulation of gene expression is crucial to the onset and progression of cancer. Changes in the structure of nucleosomes have a direct effect in the remodeling of chromatin between relatively "open" and "closed" forms that allow or prevent gene expression. Histones, the main structural component of the nucleosomes, belong to the diverse group of proteins regulated by acetylation, a common and reversible epigenetic modification[18]. Acetylation is regulated by the opposing activities of histone acetyltransferases (HATs) and histone deacetylases (HDACs). HATs transfer acetyl groups to amino-terminal lysine residues in histones, which results in a local chromatin expansion and an increased accessibility of transcription factors to DNA, whereas HDACs catalyze the removal of acetyl groups, leading to chromatin condensation and transcriptional repression[19].

Altered acetylation of histone proteins has been identified as a hallmark of cancer development[20] and several studies have widely shown that HDACs are promising targets for

therapeutic interventions intended to reverse aberrant epigenetics states associated with cancer. Indeed, HDAC inhibitors (HDACi), a novel class of chemotherapeutic agents, have been found to have anticancer activities with remarkable tumor specificity in hematological malignancies. Some HDACi have been already approved by the US Food and Drug Administration (FDA), such as varionstat or romidepsin, for the treatment of cutaneous T-cell lymphoma[21,22]. Others such as Panobinostat (LB589), a potent oral pan-HDACi in Phase I/II clinical trial, holds promise as a combination therapy for a broad range of cytopenias, including some subtypes of AML[18].

2   NUP98-HOXA9 FUSION PROTEIN

The chromosomal translocation t(7;11)(p15;p15), that results in the oncogenic fusion protein NUP98-HOXA9, was first described in 1976[23], but it was not until 1996 when the genes involved in the rearrangement were identified[24] (Figure 3). It is a rare but recurrent event that occurs predominantly in patients with *de novo* AML. It must be noted that it has also been observed in some cases with trilineage myelodisplasic syndrome (MDS) and chronic myelomonocytic leukemia[25,26]. In some occasions it appears in a secondary leukemia after an exposure to Topoisomerase II inhibitors[27]. It is associated with a very poor prognosis and a low degree of overall survival[14]. It appears more frequently in Asian countries, in females and in younger people, compared to the other M2-subtype cases. Moreover, the patients are highly refractory to intensive treatment, including allogeneic stem cell transplantation[14]. Because of such special features that characterize the patients with this chromosomal translocation, it has been considered as an independent entity[14]. However, despite the large severity of the leukemia induced by this translocation, the oncogenic events underlying its malignancy are poorly understood.
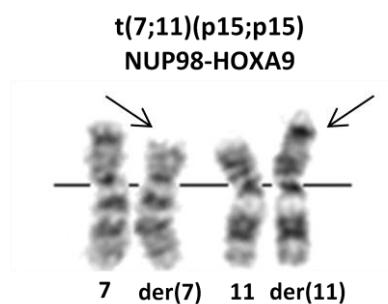
Figure 3: t(7;11)(p15;p15) chromosomal translocation. Normal and derivative chromosomes that result from the translocation.

2.1   Genes involved in the fusion protein

2.1.1   NUP98

The 98 kDa nucleoporin (NUP98) is a component of the Nuclear Pore Complex (NPC), a large route of transport between the nucleus and the cytoplasm, allowing small ions and polypeptides to pass through by diffusion and larger macromolecules (mRNA and proteins > 40 kDa) by multiprotein structure embedded in and traversing the nuclear membrane. It consists of approximately 30 different proteins, many of which are present in multiple copies, arranged in a distinguishing octagonal symmetry around a central transport channel[28]. The NPC provides a

bidirectional active transport mediated via carrier proteins and transport factors collectively called karyopherins (e.g. importins, exportins, and transportin)[8]. A still small but accumulating body of data suggests that these structures and, in particular, Nup proteins exposed to the nucleoplasmic face of the NPC, also play important roles in modulating chromatin structure and gene expression[28-30].

The *NUP98* gene maps to chromosome 11p15.4, is 122 kb long and is located 3.6 Mb from the telomere on the short arm of chromosome 11. It is composed of 33 exons and encodes 2 alternatively spliced mRNA variants: *NUP98* and *NUP98-NUP96*. *NUP98* splice variant includes only the *NUP98* portion of the *NUP98-96* mRNA and its expression is very low compared to the other variant. NUP98-NUP96 precursor polypeptide undergoes autoproteolysis and is cleaved into two peptides: NUP98 and NUP96 [8]. This coexpression from the same mRNA and its subsequent cleavage promotes and regulates the correct localization of the two mature peptides to the NPC. However, two pools of NUP98 have been found: an NPC-bound fraction (on both the nucleoplasmic and cytoplasmic domains) and an off-pore intranuclear fraction[31] (Figure 4). Phosphorylation of NUP98 during mitosis causes it to release from the pore and it is possible that this post-translational modification could promote NUP98 cycling to and from the NPC.[31,32]
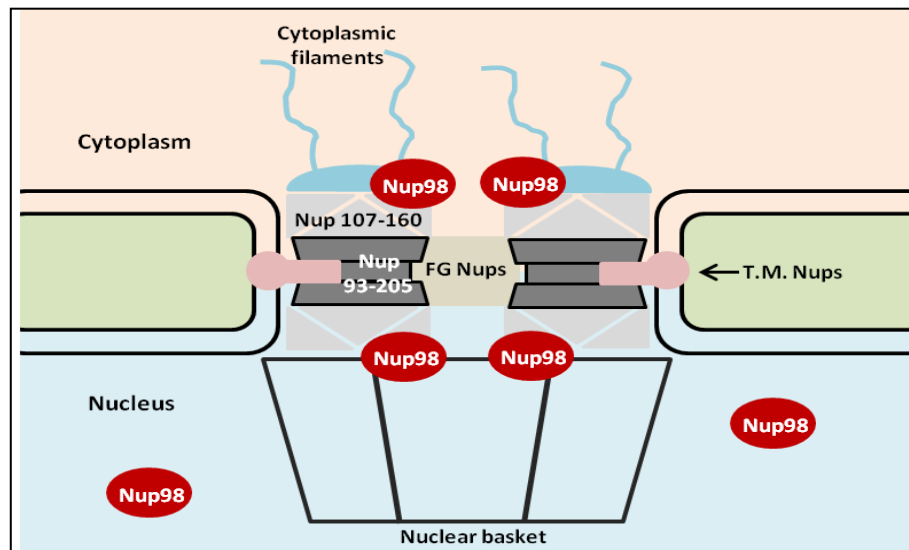
Figure 4: The nuclear pore complex (NPC). The core structure of the NPC is maintained by the NUP93-205 complex (black) and the supporting structures of the NUP107-160 complex (grey) on the nuclear and cytoplasmic faces of the pore. The cytoplasmic filaments (blue) and the nuclear basket (black) allow the pore to sense the cytoplasmic and nuclear compartments, respectively, and regulate transport of macromolecules through the NPC via binding sites on FG Nups in the barrel of the pore. The entire pore is anchored to the nuclear envelope by protein–protein contacts between transmembrane Nups, T.M. Nups (pink) and the NPC scaffold. NUP98 is found on the nuclear and cytoplasmic faces of the pore (red) as well as in the nucleoplasm. (Modified from Franks *et al., Trends in Cell Biology*, 2013)

Approximately a third of all nucleoporin proteins contain repeats of Phe-X-Phe-Gly amino acid residues, or Gly-Leu-Phe-Gly (GLFG) residues, collectively called FG repeats. However, NUP98 is distinct from other FG nucleoporins because it contains multiple nontandem GLFG repeats[8]. The GLFG repeats are thought to function as docking sites for different proteins. They have been shown to bind nuclear exportin 1 protein CRM1 (XPO1) and the mRNA export factor TAP, among others. The nontandem FG repeats of NUP98 are intersected by a coiled-coil domain, the Gle2-binding sequence (GLEBS) motif in the N-terminal portion of NUP98 that interacts with the RNA export factor RAE1 (Gle2). Finally, the C-terminal end of NUP98 contains an RNA-binding motif that is also involved in the export of mature forms of RNA[8] (Figure 5). There is evidence that NUP98 is involved in RNA export and protein import[33].
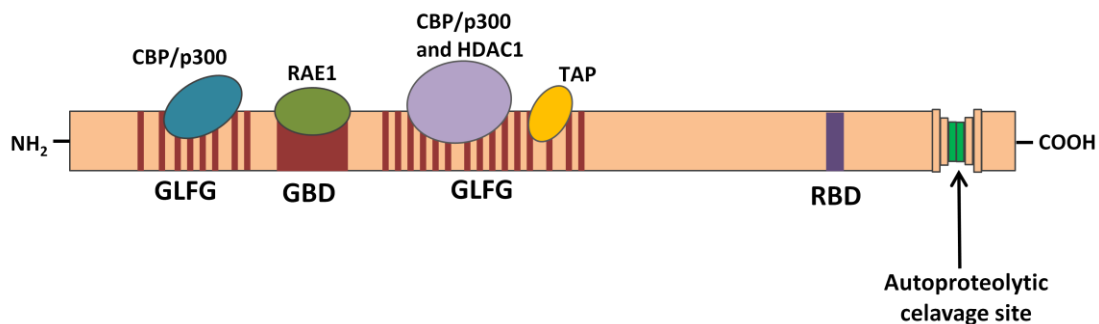
Figure 5: Schematic representation of the NUP98 protein. Red lines indicate GLFG repeats, red box indicates the Gle2-binding domain (GBD), purple box indicates the RNA-binding domain (RBD) and the green lines represents the autoproteolytic cleavage site. Known NUP98 interacting proteins are indicated: CBP/p300 (blue), TAP (yellow), RAE (green) and HDAC1 (purple) (Modified from Gough *et al*., *Blood,* 2011)

In addition to its well established role in mediating nucleo-cytoplasmic transport, recent studies have shown that the intranuclear fraction of NUP98 could also localize to promoters of multiple genes[28-30,34]. Moreover, the downregulation of different target genes after the knockdown of NUP98 suggests that it could have important functions as a transcription factor in *Drosophila*[30,34]. Supporting these findings, it was found that the FG repeat region of NUP98 is able to recruit p300 to activate transcription or HDAC1 to repress transcription[33] (Figure 5). Later, Liang *et al*[35] demonstrated that NUP98 dynamically associates with the human genome in a manner that is tightly linked to differentiation status and developmental gene expression. They found that in human progenitor cells NUP98 is highly associated to active genes and that it is functionally important for their expression. However, in differentiated cells, NUP98 mainly interacts with silent chromatin domains. Importantly, it seems that both the intranuclear and the NPC-bound fraction of NUP98 are involved in the developmental gene regulation. Although the

mechanisms by which this nucleoporin is targeted to chromatin remain unclear, it is known that at least part of its DNA binding activity resides in the FG repeat region[33].

Finally, FG repeats also allow NUP98 to interact with the dynein light chain, DYNLT1, which suggests a role of NUP98 in the process of chromosome segregation[36].

Frequent allelic loss at the 11p locus has been reported in AML and indeed, Loss of Heterozygosity (LOH) of the *NUP98* gene has been considered an adverse prognostic factor in Acute Myeloid Leukemia (AML)[37]. However, the molecular mechanisms involving *NUP98* gene locus and its function in leukemogenesis are still unknown.

2.1.2   *HOXA9*

Homeobox A9 (HOXA9) is a member of the highly conserved *HOX* protein family of transcription factors, which play key roles in both development and hematopiesis[38].

In mammals, 39 genes have been identified, separated into four clusters (A-D) located in four different chromosomes. *HOX* genes are numbered such that genes of different clusters with the same number have greatest similarity and functional redundancy (paralogs). (Figure 6)
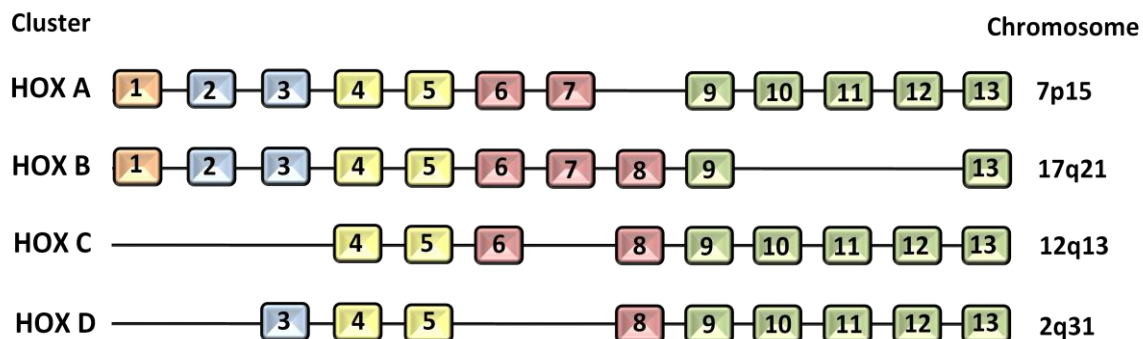
Figure 6: Clustered HOX genes. Schematic structure of HOX genes clusters. The 39 HOX genes are located on four different chromosomes. Human HOX genes are shown by colors and blank spaces shown missing genes. (Modified from Alharbi *et al*, *Leukemia*, 2013)

In particular, the *HOXA9* gene is located in chromosome 7p15, is 7.8 kb long and is composed by two exons.

HOX proteins bind to DNA through a homeodomain (HD), which is found in the C-terminus of the protein. The HD is highly conserved between HOX proteins and includes conserved tyrosine residues which may regulate HOX activity. However, this regulation may be different for each type of HOX protein. For example, the HD-tyrosine phosphorylation of HOXA9 increases the binding affinity to its target genes whereas the affinity of HOXA10 to the same genes is

43

decreased.[39,40]   HOX proteins need cofactor interactions in order to increase the selectivity, affinity and specificity for their DNA-binding sites. The most important cofactors are pre-B-cell leukemia (PBX) and myeloid ectopic insertion site (MEIS) families. Indeed, HOX proteins harbor specific binding domains for this cofactors: PBX motif (PM) and MEIS domain (MD).[41] Additionally, HOX-PBX binding requires a hexapeptide motif containing a critical tryptophan residue in the HOX proteins[42]. (Figure 7)
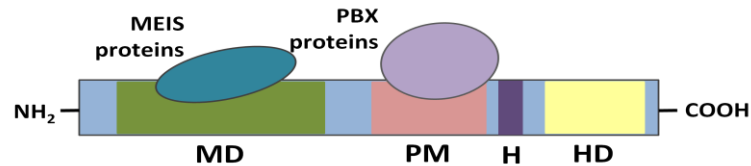
Figure 7: Schematic representation of the HOXA9 protein. Green box represents MEIS domain (MD), pink box corresponds to PBX motif (PM), purple box represents the hexapeptide (H) and yellow box would be the homeodomain (HD).

HOX gene transcription during hematopoiesis is tightly regulated in a temporal manner. Maximal expression of *HOX1–6* occurs in most primitive hematopoietic stem cells (CD34$^+$CD38$^-$) in humans, with downregulation of these genes during CD38+ differentiation. Subsequently, *HOX7–11* expression is maximal during commitment (CD34$^+$CD38$^+$) with downregulation as differentiation proceeds[43]. Indeed, the expression of *HOX* genes is almost absent in CD34$^-$ cells, which are considered differentiated bone marrow cells. Interestingly, the different *HOX* clusters also have specific patterns of lineage-restricted expression: *HOXA* genes are expressed in myeloid cells, *HOXB* genes in erythroid cells and *HOXC* genes in lymphoid cells. *HOXD* genes are not expressed in hematopoiesis despite having similar regulatory regions to the other clusters [41].

The function of HOX proteins in normal hematopoiesis and leukemia has been widely studied. In particular, HOXA9 is required for the maintenance of hematopoietic progenitor status and promotes proliferation. Its overexpression in murine bone marrow induces myeloid progenitor expansion and, after a very long latency, leads to leukemia. Interestingly, HOXA9 overexpression only results in a partial inhibition of pre-B-cell differentiation and does not affect T-cell development.[44] Thus, it has been hypothesized that HOXA9 is involved in the selection of myeloid versus lymphoid lineage commitment.[43] On the other hand, *HOXA9$^{-/-}$* mice showed deficiencies in myeloid and lymphoid cells with a significant defect in their repopulating ability, as well as the corresponding reduction in spleen cellularity and size[45].

HOXA9 expression is controlled by different transcriptional activators such as MLL and caudal-type HOX transcription factor family (CDX1, CDX2 and CDX4). Also, as *HOX* genes are distributed in clusters, they are especially sensitive to changes in the chromosomal organization[41].

HOXA9 regulates transcription by interacting with enhancers. Recent genome-wide studies have made it possible to identify some potential target genes that are directly related to cell proliferation and survival. Most of them are upregulated (*MEIS1*, different members of *HOX* family, *SOX4, CAMK2D, FOXP1, PIM1, EVI1, etc.*) but HOXA9 is also able to downregulate some targets like *RUNX* genes[46].

HOXA9 may also bind DNA as heterodimer with MEIS1 or PBX proteins or as a heterotrimer with both of them to regulate different genes. Among the targets of these transcriptional activator complexes we found the *MYB, MEF2C or FLT3* genes. Notably, the coexpression of HOXA9 and MEIS1 in mice significantly reduces the time of latency in the induction of AML and there is increasing evidence of the importance that these factors appear to have in the development of AML[47]. Different studies have identified the consensus binding site for HOXA9 complexes. PBX-HOXA9 would bind to ATGAT<u>TTACGAC</u> sequence, MEIS1-HOXA9 consensus target would be <u>TGACAGTTAT/C</u> and PBX-MEIS binds to TGAT<u>TGACAG</u>. PBX consensus binding sites are not underlined, HOXA9 consensus binding sites are underlined and MEIS consensus binding sites are broken underlined. Surprisingly, a DNA binding site specific for the MEIS1-HOXA9-PBX heterotrimer has not been detected but it seems to be interacting with PBX-HOXA9 consensus sites. Hence, MEIS1 could bind to a PBX-HOXA9 dimer bound to DNA without interacting to its specific recognition site. Indeed, MEIS1 enhances in vitro HOXA9-PBX protein complex formation[48].

It is therefore unsurprising that the overexpression of HOXA9 is a highly frequent event in AML. Its deregulation is associated in some cases with chromosomal rearrangements that involve HOXA9 upstream regulators, such as MLL, but also with specific mutations, like *NPM1* mutations[41]. *NPM1* mutations are the most common genetic aberration in AML, reported in about 35% of adult patients. Such mutations induce the recolocation of NPM1 into the cytoplasm, which leads to the overexpression of *HOXA9* and other *HOX* genes[49].

HOX genes expression has become an important prognostic factor in AML[41]. In particular, *HOXA9* has been the single gene correlated with the worst outcome and relapse of disease and short survival, among all the genes investigated in AML patients[50]. Subsequently, low *HOXA9*

expression, as well as low expression of *MEIS1*, were favorable predictors for AML patient outcome and good response to therapy[41].

2.2   Molecular description

NUP98-HOXA9, the fusion gene originated by the chromosomal translocation, encodes the FG repeat-rich portion of the nucleoporin *NUP98,* fused to the homeodomain region and PBX heterodimerization domain of *HOXA9* (Figure 8).
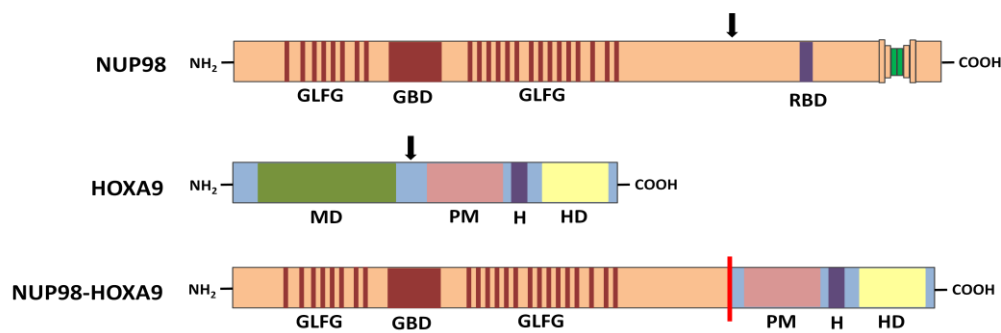
Figure 8: Schematic representation of NUP98, HOXA9 and NUP98-HOXA9 proteins. Arrows indicate the breakpoints that originate the chromosomal translocation

Different studies have described four types of fusion sequences in patients with t(7;11)(p15;p15), which appeared to be the result of the alternative splicing of *NUP98* and *HOXA9*[14]. All of them are in-frame and originate fusion proteins with a similar number of amino acids (Table 2 and figure 9). Although it has not been analysed experimentally, it is expected that the different fusions act similarly as they retain the same functional domains. Leukemic cells from these patients contain at least two different fusion products, but Type I (fusion between *NUP98* exon 12 and *HOXA9* exon 1b) is present in all the patients.

NUP98-HOXA9 fusion gene Type I has 13 exons and originates a chimeric protein with 579 aminoacids and 60kDa of molecular weight.

Table 2: Types of *NUP98-HOXA9* transcripts

FUSION TYPE	BREAKPOINT	Nº OF NUCLEOTIDES	Nº OF AMINOACID
Type I	NUP98 exon 12 – HOXA9 exon 1b	1737	579
Type II	NUP98 exon 12 – HOXA9 exon 2	1647	549
Type III	NUP98 exon 11 – HOXA9 exon 1b	1596	532
Type IV	NUP98 exon 11 – HOXA9 exon 2	1506	502
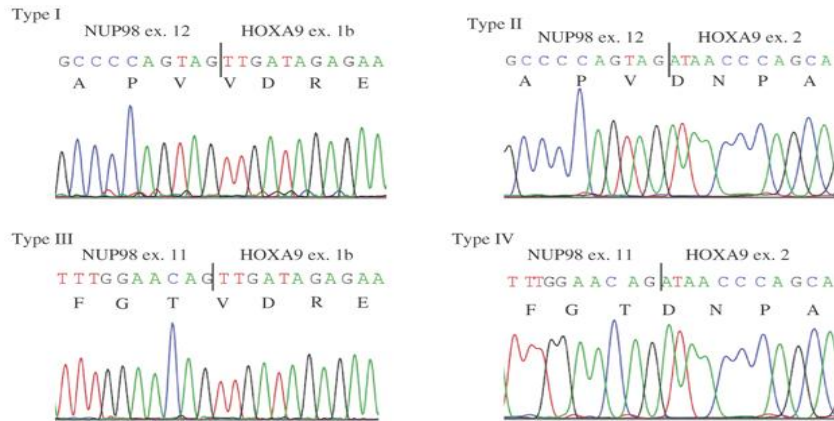
Figure 9: **Fusion point sequence in each type of transcript**. The corresponding amino acids are listed below the nucleotide sequence. Black lines represent the breakpoints. (Modified from Chou *et a.l*, *Leukemia*, 2009)

2.3 Subcellular localization

NUP98-HOXA9 is located almost exclusively in the nucleus and it is never found, neither as a part of nor associated with the NPC. Interestingly, during mitosis, the fusion protein is concentrated at kinetochores and along chromosome arms[51], unlike NUP98 wt, which is distributed throughout the cell. This association is detected in prophase and persists through metaphase and anaphase (Figure 10). Importantly, these novel intranuclear localizations of NH require both the GLFG repeats of NUP98 and the homeodomain of the HOX protein.
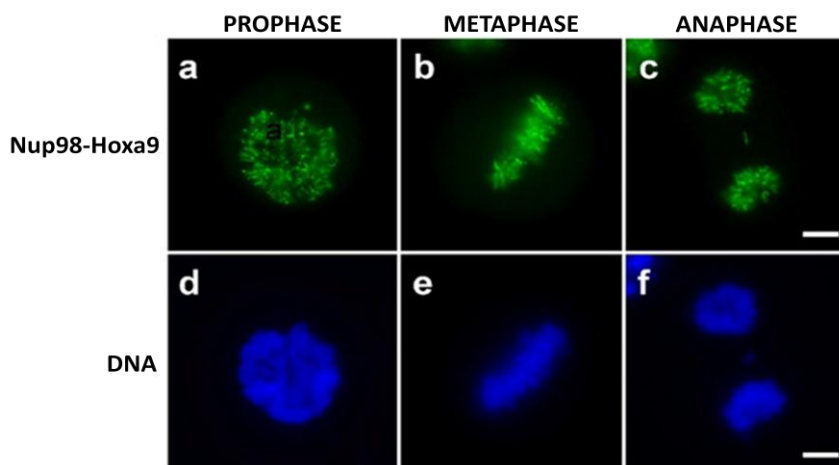
Figure 10: Subcellular localization of NUP98-HOXA9. GFP-Nup98 fusions were highly concentrated on chromosomes during mitosis and exhibit a punctate staining pattern along mitotic chromosomes. (Modified from S. Xu *et al.,* Molecular Biology of the Cell, 2010)

2.4   Mechanism of action

NUP98-HOXA9 is an oncogenic fusion protein that in murine bone marrow results in a rapid, polyclonal myeloproliferative disease progressing to AML by 7-8 months[26]. However, in human primary hematopoietic cells, NH induces long-term proliferation, blocks the differentiation and dramatically increases the number of primitive cells, but additional genetic changes are required for the progression to the overt leukemia [52,53].

Specifically, enforced expression of NH in hHP confers a proliferative advantage and increases self-renewal. It also alters differentiation by inhibiting erythropoiesis, delaying neutrophil maturation and blocking myeloid colony formation[54].

It has been commonly accepted that NH acts as an oncogenic transcription factor[53-55]. Indeed, some studies have analyzed changes in the expression profile in presence of NUP98-HOXA9 in hHP and extensive changes with a preponderance of upregulated genes have been detected[53-56]. Among the deregulated genes we can find oncogenes, transcription factors, growth factors, cell cycle regulators, tumor suppressors and factors involved in hematopoietic differentiation[53]. However, it has not been determined whether these genes are directly regulated by the chimera or their overexpression is a downstream effect. Based on the involvement of HOXA9 in leukemogenesis and its function as transcription factor, it was accepted that the homeodomain of HOXA9 conferred NUP98-HOXA9 the ability to bind to DNA to regulate gene expression[52]. However, there are evidences that suggest that NH also induces oncogenic mechanisms beyond those mediated by the homeodomain. On one hand, NH induces a leukemia that is preceded by a myeloproliferative phase, whereas the leukemia caused by overexpression of HOXA9 is not preceded by myeloproliferation[25]. On the other hand, mutations in the homeodomain do not abolish most of the effects of NH on hematopoietic differentiation[52]. Thus, it seems clear that the NUP98 region of the fusion protein is required for the induction of the leukemic transformation driven by NH. However, the specific role of NUP98 and how such a broad range of different partners can bring similar AML phenotypes is still not clear. NUP98 could be mediating the interaction with other proteins through the FG-repeat domain and it has been suggested that NUP98-HOXA9 could interact with p300[55,57] (transcriptional activator), HDACs[37](transcriptional inhibitor), CRM1[58,59] (protein involved in nuclear transport) or AES[60] (transcriptional Regulator of the TLE family), but limited information is available regarding the features, architecture or the effects in leukemogenesis of these interactions. Interestingly, during interphase, there are dynamic interactions between NH and endogenous NUP98 that lead to mislocalization of the

intranuclear fraction of NUP98, but do not alter the level of NUP98 at the nuclear pore complex. During mitosis, this interaction would be disrupted[51]. In addition, it has been suggested that the interaction of NH with DYNLT1 could cause defective chromosomal segregation, contributing to the oncogenic transformation[36].

Importantly, recent studies that propose a role for NUP98 in transcriptional regulation[30,35] uncover a very interesting aspect for the study of NUP98 leukemic fusions that must be explored. NUP98, together with its ability to interact with different proteins, could allow NH to regulate the expression of different target genes.

Therefore, there is still a long way to completely elucidate the functional contribution of each of the moieties of this chimeric protein.

2.5   Cooperating oncogenic events and mutations

It is accepted that *NUP98-HOXA9* fusion gene is a class II mutation, because most of the described collaborating events are class I proliferation and survival mutations. However, this would be another example of how simplistic the *Two hit hypothesis* is, since NH does not exclusively play a role in blocking the differentiation, but it also induces uncontrolled proliferation of progenitor cells. In contrast, based on the model proposed by Welch *et al*[10], we could consider NH as a driver mutation that plays a causative role in the development of AML but needs the cooperation of different oncogenes to enhance its leukemogenic potential.

Several studies have demonstrated an enforced strong transcription of MEIS1 in NH cellular models. A collaborator role of this factor with the fusion protein in leukemogenesis has been suggested [26,27,54,55]. Co-expression of NH with MEIS1 shortens the period of AML development in transplanted mice from 230 to 142 days[55].

On the other hand, several studies have detected the NH fusion gene in patients with Chronic Myeloid Leukemia (CML), a clonal stem cell disease caused by the BCR-ABL fusion oncoprotein. The *BCR-ABL* fusion gene is originated by the chromosomal translocation t(9;22)(q34;q22), whose derivative chromosome is known as Philadelphia chromosome (Ph). Interestingly, there have also been patients who developed NH-induced myeloproliferation and subsequently acquired a Ph chromosome in their leukemic cells when AML was diagnosed, suggesting that NH and BCR-ABL may interact genetically in human leukemia and influence disease progression[57,61]. In fact, coexpression of NH with BCR-ABL reduces the period to AML in transplanted mice even further than MEIS1, from 230 to 21 days[55].

NH is strongly associated with activating mutations in *KIT, NRAS, KRAS, WT-1* and also *FLT3-ITD* mutations, each of which is a class I mutation that confers a proliferation or survival advantage to hematopoietic cells [8,14,33].

2.6 Other oncogenic NUP fusions

Although *NUP98-HOXA9* is the most frequent, and it is considered the prototype, *NUP98* is known to fuse to other twenty-eight different partner genes to produce abnormal fusion proteins, listed in Table 3. They are caused primarily by balanced translocations and inversions in the malignant cells of patients with a wide array of distinct hematopoietic malignancies, mostly AML, but also including T-ALL and MDS[8,33]

All of these chimeric fusion proteins have the N-terminal half of the NUP98 protein in common, which includes the FG repeat region. The C-terminal partner in these chimeras can be one of a wide variety of proteins that can be divided into two general classes: 1) homeodomain transcription factors and 2) other, typically nuclear and often nucleic acid-binding proteins, such as topoisomerases or the putative RNA helicase,DDX10[51].

Just like NH, the other NUP98 fusion proteins also induce proliferation and impair the differentiation in hHP. Likewise, they cause MDS that progress to AML in mice with variable penetrance, suggesting that cooperating oncogenic events are needed for the complete leukemic transformation[33,62-66]. However, the molecular mechanisms by which these fusion proteins, involving such different partners, cause similar leukemic phenotypes, are completely unknown. Many of the NUP98 translocation partner are transcription factors with a DNA binding domain (like the homeodomain transcription factors) and it is postulated that, as NH, they deregulate the expression of genes involved in hematopoiesis[8]. In other cases, NUP98 is fused to epigenetic regulators, like histone methyltransferase (NSD1, NSD3 or MLL) or histone demethylase (JARID1A) and it seems that they modulate gene expression by modifying chromatin at numerous regions genome-wide.[8,31] However, there is no explanation for those cases in which NUP98 is fused to a gene that does not have a DNA-binding activity or that has an unknown function.

Importantly, NUP98 is not the only nucleoporin implicated in the pathogenesis of hematological malignancies. Four other fusion genes involving *NUP214* have been described (Table 3).

Table3: Nucleoporin gene rearrangements in hematologic malignancies. AML = acute myeloid leukemia, CML-BC = chronic myelogenous leukemia in blast crisis, MDS = myelodysplastic syndrome, t-AML = therapy-related acute myeloid leukemia, t-MDS = therapy-related myelodysplastic syndrome, APL = acute promyelocytic leukemia, T-ALL = T-cell Acute lymphoblastic leukemia, B-ALL = B-cell acute lym-phoblastic leukemia, and AUL = acute undifferentiated leukemia. (Modified from Takeda *et al.; Semin Cancer Bio,* 2014)

Rearrangement	Fusion trasncript	Disease
NUP98		
t(7;11)(p15;p15)	**NUP98-HOXA9**	AML/MDS, t-AML/MDS, CML
t(7;11)(p15;p15)	NUP98-HOXA11	CML-BC
t(7;11)(p15;p15)	NUP98-HOXA13	AML,MDS
t(11;12)(p15;q13)	NUP98-HOXC11	AML
t(11;12)(p15;q13)	NUP98-HOXC13	AML
t(2;11)(q35;p15)	NUP98-HOXD11	Pediatric AML
t(2;11)(q31;p15)	NUP98-HOXD13	AML, t-AML
t(1;11)(q23;p15)	NUP98-PMX1	AML, t-MDS/AML
t(9;11)(q34;p15)	NUP98-PRRX2	t-AML
t(10;11)(q23;p15)	NUP98-HHEX	AML
inv(11)(p15;q22)	NUP98-DDX10	AML, MDS, CML
t(11;20)(p15;q11)	NUP98-TOP1	AML, t-MDS
t(9;11)(p22;p15)	NUP98-PSIP1	AML, MDS
t(5;11)(q31;p15)	NUP98-NSD1	Pediatric AML
t(8;11)(p11.2;p15)	NUP98-NSD3	AML
t(3;11)(p24;p15)	NUP98-TOP2B	AML (Monoblastic)
Complex (12p13)	NUP98-JARID1A	AML (Megakaryoblastic)
t(11;17)(p15;p13)	NUP98-PHF23	AML
Complex (3p25)	NUP98-ANKRD28	MDS/AML
t(6;11)(q24.1;p15.5)	NUP98-CCDC28A	AML (Megakaryoblastic), T-ALL
Complex (3q29)	NUP98-IQCG	Biphenotypic T-ALL/AML
t(11;18)(p15;q12)	NUP98-SETBP1	Pediatric T-ALL
t(4;11)(q21;p15)	NUP98-RAP1GSD1	Adult T-ALL
t(10;11)(q25;p15)	NUP98-Adducin 3	T-ALL
t(X;11)(q28;p15)	NUP98-HMGB3	t-AML
t(3;11)(q12;p15)	NUP98-LOC348801	AML
inv(11)(p15;q23)	NUP98-MLL	AML
t(11;12)(p15;q13)	NUP98-RARG	APL
t(3;11)(p11;p15)	NUP98-PU1F1	t-AML
NUP214		
t(6;9)(p23;q34)	DEK-NUP214	AML
del(9)(q34)	SET-NUP214	T-ALL, AML, AUL
der(5)t(5;9)(q35;q34)	SQSTM1-NUP214	T-ALL
Amplified episomes	NUP214-ABL1	T-ALL, B-ALL

3   ENHANCERS AND TRANSCRIPTIONAL REGULATION

Enhancers are DNA *cis*-regulatory elements that enable the regulation of target gene expression and that are able to function close to the promoters or at large distances upstream or downstream from them. Importantly, multiple activators and repressors can bind and modulate an individual enhancer and multiple enhancers can determine a single gene's precise pattern of expression[67]. Currently, there is a growing number of studies that point to the importance of enhancers in orchestrating important transcriptional networks and highlight the fact that the misregulation of enhancer properties is central to the pathogenesis of cancer and other diseases known as "enhanceropathies".

The initial recognition of enhancers requires pioneer transcription factors that can bind to their consensus sequences and facilitate the binding of additional transcription factors to open chromatin. GATA1, FOX proteins or Pu.1 are some of the identified pioneer factors[68]. Additionally, the *communication* between enhancers and promoter regions seems to be necessary for the proper transcriptional regulation. Cohesin has been validated as a major regulator of this enhancer-promoter interaction. Its ability to hold together sister chromatids is used to form or stabilize loops of chromatin that bring enhancers and promoters in close proximity[69] (Figure 11A).
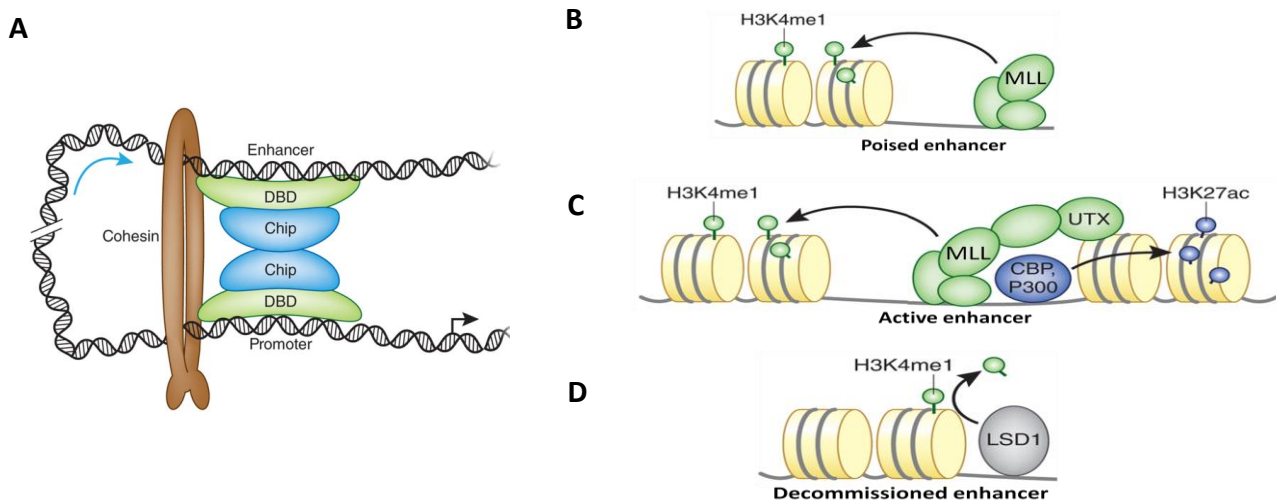
Figure 11: Structure of enhancers. (A) Schematic representation of the loop that enables the enhancer-promoter communication. Chip (in Drosophila) or LDB1 (the mammalian homolog) is thought to interact with DNA-binding transcription factors (DBDs) at enhancers and promoters and through a self-interaction domain to help bridge this communication. The ring-like cohesin structure is thought to hold the enhancer and promoter together. (B-C) Schematic representations of the chromatin signatures associated with different enhancer states and the mediators responsible for each of them: H3K4me1 implemented by MLL in poised and active enhancers (A), H3K27ac implemented by CBP/p300 that is characteristic of active enhacncers (B) and the removal of the H3K4me1 in enhancer dcommissioning. (Images modified from E. Smith & A. Shilatifard, Nat Struct Mol Biol. 2014)

Enhancers have been classified into poised and active states. Chromatin signatures in these different states have been identified. Both classes maintain the enrichment for H3K4me1 that is implemented by MLL3 and MLL4 (lysine (K)-specific methyltransferases, components of COMPASS family). During differentiation, however, many enhancers become silenced by losing the H3K4me1 that is removed by the demethylase LSD, in a process called enhancer decommissioning. H3K27ac, implemented by CBP and p300, is present in active enhancers. Thus, CBP and p300 are found at active promoters and their occupancy at intergenic or intragenic regions has indeed served as a useful and convenient marker for enhancers (Figure 11 B-D). UTX is the H3K27 deacetylase that would remove the active mark and enable the return to poised state. Active enhancers also have the ability to recruit Polimerase II (PolII)[70].Most of them are transcribed by PolII generating transcripts named enhancer RNAs (eRNA). The appearance of eRNAs precedes the activation of nearby genes, but their function is not completely understood yet [67].

A number of genome-wide studies on transcription factor binding in multiple different cell systems have shown that many transcription factors tend to co-localize with other factors on chromatin[71]. More recently, the existence of super-enhancers, which are large genomic regions (several kilobases) containing clusters of closely spaced transcription factor binding regions, has been demonstrated. These large super-enhancers are characterized by very high levels of Mediator subunit 1 (MED1) binding and seem to regulate cell identity[71-73]. Hence, the fact that enhancers harbor binding sites for a large number of transcription factors highlights the importance of these regulatory elements in normal cell functions and points to that the global network of enhancer–promoter interactions may be much more complex than it was previously believed[74].

Enhanceropaties are originated by the alteration of the enhancer sequences or mutations in proteins that either directly interact with enhancers or regulate enhancer function[67,75]. Importantly, many oncogenes or tumor suppressor genes are regulated by enhancers but also, many oncogenes are transcription factors that regulate the gene expression by interacting with enhancers. All of this helps to explain the potential role that these regulatory regions seem to play in cancer pathogenesis. Pharmacological approaches that target transcription factors that are recruited to the enhancers of tumor suppressor genes or oncogenes ("Enhancer Therapy"), are being developed with promising results in the fight against enhancer-mediated cancers[75].

OBJECTIVES

NUP98-HOXA9-associated AML is one of the most aggressive forms of leukemia, highly refractory to intensive treatment and with a dismal overall survival. Current knowledge about the molecular mechanisms underlying this fusion protein, something that would trigger the identification of new and more efficient therapeutic targets, is still limited.

This project was aimed at studying the oncogenic role of this leukemic fusion gene in a human cellular modeled context, with the purpose of providing some proofs of concept that allow to explore a better designed targeted therapy for these patients. For this aim our objectives were:

1. Generate *in vitro* human cellular models that constitutively express NUP98-HOXA9.

2. Based in these models:

   a. Identify and analyze the specific DNA binding regions of NH, their target genes and the global expression profile induced by the fusion protein.

   b. Identify new NH-interacting proteins and the role of each moiety of the chimera in the NH driven leukemogenesis

   c. Describe some of the molecular mechanism of action of this fusion protein

   d. Propose new therapeutic targets and a possible effective treatment for these patients.

METHODS

1 CELLULAR MODELS

1.1 Cell culture

hHP (CD34+) were purified using immunomagnetic beads (Miltenyi Biotech, Germany) from Human umbilical cord blood (CB). hHP were cultured in Iscove's modified Dulbecco's media (IMDM) containing 20% BIT 9500 Serum substitute (STEMCELL Technologies), 1% penicillin-streptomycin (Invitrogen), 0,2% β-mercaptoethanol and 10 ng/ml SCF, megakaryocyte growth and development factor (TPO), FLT3 ligand (FLT3L), IL-3, and IL-6.

HEK293T and HEK293FT cells were cultured in Dulbecco modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (Invitrogen).

1.2 Retroviral constructs

NUP98-HOXA9 coding region sequence was PCR amplified from cDNA from a patient with primers, forward 5'-ATGTTTAACAAATCATTTGGAACACCCT-3', and reverse 5'-TCACTCGTCTTTTGCTCGGTCTTTGTTG-3' to be cloned into pMSCV-IRES-GFP retroviral vector, and named pMSCV-NH. The coding region of *HOXA9* was amplified from a cDNA control with primers, forward 5'-ATGGCCACCACTGGGGCCCTGGGCAACT-3', and reverse 5'-TTACTCGTCTTTTGCTCGGTCTTTGTTG-3'. It was cloned into the same retroviral vector and named pMSCV-HOXA9. The *NUP98* coding region sequence was kindly provided by Dr Maureen A. Powers and PCR amplified with primers, forward 5'-ATGTTTAACAAATCATTTGGAACACCCT-3', and reverse 5'-TCACTGTCCTTTTTTCTCTACCTGAGGT-3'. It was cloned into the same retroviral vector and named pMSCV-NUP98. We included FLAG and HA as tag in the three coding regions.

The *PBX3* coding region was amplified from the vector PBX3-pSP65 (Plasmid #21033 from Addgene, http://www.addgene.org) and cloned into the pMSCV-IRES-GFP retroviral vector and named pMSCV-PBX3, using the forward primer 5'-ATGAAACCAGCGCTCTTCAGCGTCCTGT-3' and the reverse 5'-GTTAGAGGTATCCGAGTGCACACTTC-3' with FLAG as tag.

1.3 Retroviral production and transduction

Retrovirus for each construct was produced in 293FT cells by cotransfecting viral plasmids along with M57 and RD114 packaging plasmid in a 1:1.2:0.3 ratio using Calcium-phosphate protocol. 12 hours later, the precipitate was washed away, and viral supernatants were collected

at 24 and 36 hours. The supernatant was centrifuged at 1800 rpm and filtered through a 0.45μm filter[76,77].

Transduction of hHP (CD34+ cells) was carried out on plates coated with Retronectin (Fisher Scientific, Pittsburgh, PA) preloaded with virus. We transduced 2×10^6 cells with 3 rounds of retrovirus using retronectin-coated dishes in the same media used for hHP culture with the addition of 8 μg/ml Polybrene. Transduction of 293FT cells was performed in the same way but without Retronectin.

hHP and 293FT cells were stably transduced with the different retroviral vectors and GFP positive cells were sorted on BD Influx™ cell sorter (BD Biosciences).

1.4   Patient samples

The leftover material of 8 AML t(7;11)(p15;p15) samples were used for RNA extraction. Three of them came from the MLL Münchner Leukämielabor GmbH (München, Germany) and the other five from the cytogenetic laboratories of the University of Navarra (Pamplona, Spain). Informed consent was obtained according to the Declaration of Helsinki (October 2008).

2   IDENTIFICATION OF GENOMIC BINDING SITES

2.1   Chromatin immunoprecipitation sequencing (ChIP-seq)

Chromatin immunoprecipitation (ChIP) was performed as described before[78]. 15 min of formaldehyde fixation in ≈ 10 millions of cells from 293FT models were performed to yield the best combination of *in vivo* fixed chromatin, high DNA recovery and small average size of chromatin fragments. In our analyses, chromatin was sonicated and sheared to an average length of 200 bp. Immunoprecipitation was carried out using Protein A/G Plus-Agarose beads (Santa Cruz) and the antibodies anti-FLAG M2 (Sigma-Aldrich) and anti-HA tag antibody-ChIP Grade (ab9110, abcam). DNA was purified with phenol/chlorophorm extraction followed by ethanol precipitation. Eluted DNA fragments were analysed by ultrasequencing. Libraries was constructed using the ChIP-seq sample preparation kit (Illumina, San Diego, California, USA) and were sequenced in a Genome Analyzer IIx (GA2, Illumina) single 36-base read run. Raw sequences were defined as reads passing purity filter before genome alignment. Alignment was performed with BWA[79] versus the human sequence assembly (GRCh37/hg19, Feb 2009) under

default settings permitting alignments with 1 mismatch in 40 base reads. Peak detection was performed with MACS 2.0[80] following developer's technical recommendations (pvalue cutoff for peak detection = 1×10^5). MACS pipeline was used to find differential peaks in pair-wise comparisons. Significant ChIP-seq peaks were stablished at FDR ≤ 5%. The generated bed files were uploaded to the USCS genome browser (http://genome.ucsc.edu/) to visualize the enrichment peaks obtained with Findpeaks analysis. Venn diagrams used for gene set comparisons were performed at http://www.cmbi.ru.nl/cdd/biovenn/.

2.2   *In silico* data analysis

Gene lists were functionally annotated using Ingenuity Pathways Analysis software (Ingenuity® Systems, http://www.ingenuity.com/products/ipa). Enrichment of genes associated with specific biological functions, canonical pathways and diseases was determined relative to the Ingenuity knowledge database. The significance level cut off used pas p > 0.05 after Benjamin-Hoechberg multiple testing correction (B-H p-value).

MEME-ChIP[81] was employed to find enriched motifs in significant ChIP-seq peaks. They were also scanned for detect individual HOXA9, MEIS1 and HOXA9/MEIS1/PBX motifs using *FIMO* tool (from MEME Suite web server) [82]. *oOPOSSUM*[83] (http://opossum.cisreg.ca/oPOSSUM3/) was used to identify the over-represented transcription factor binding sites from the top 100 NH target genes. (Figure 12)

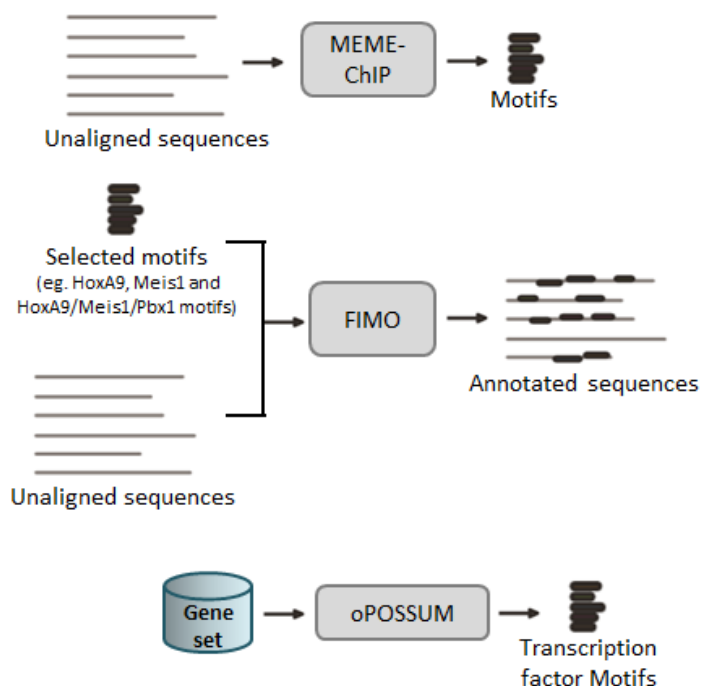**Figure   12:   Overview   of   the bioinformatics tools used for the motif analysis.** MEME-ChIP scans unaligned sequences from ChIP-seq analysis to identify described motifs. FIMO identifies a set of selected motifs in unaligned sequences from ChIP-seq   analysis.   oOPOSSUM   identifies known motifs of transcription factors in the regulatory regions of a gene set (Image modified from Bailey et al[84])

3   VALIDATION OF TARGET GENES

3.1   qChIP

ChIP was performed as described above (ChIP-seq paragraph) with $2X10^6$ cells per sample (hHP or 293FT cellular models). Quantitative PCR (qPCR) of eluted DNA was performed in 12ul with SYBR Green dye (Applied Biosystems) for validating the selected target genomic regions of NH obtained from the ChIP-seq data on an ABI 7900HT sequence detection system. Three biological replicates were performed and the input DNA, 'unbound' (wash) fraction of the no antibody control and the 'bound' fractions for each antibody was included for each experiment. In a house designed primers used for validation are listed in Table 4. The amount of immunoprecitpiated DNA relative to input was calculated for each experiment.

Table 4: qChIP primers

ENHANCERS	HOXA9	5'- GCTTACAATACCTCCTCCATCAA-3'
		5'- CTCTAAACCTCAGGCCACATC -3'
	MEIS1	5'- TGCCATTCATACCCTGCTATAC -3'
		5'- TTGGAATCCTAGTGGATGTTTCT -3'
	PBX3	5'- GCTACTCCAGGAACACTAAACC -3'
		5'- CGGTCACAGTCAGAGAGAAATG -3'
	MET	5'- GGGCATTCTGCTCCTGTTAT -3'
		5'- ATCCCAAGCGGAATATACTAACC -3'
	BRAF	5'- CACAAGGGCACCTGTGAATA -3'
		5'- TGGATTGCATCATAACCCTGAA -3'
	AF9	5'- CCTTAGGCACCCACATGTATATT -3'
		5'- GCCTCTGATTAGCCTGTGTATG -3'
	NF1	5'- GGTACAGGTCTATATGTGTGTCTAAA -3'
		5'- AGAGTCTGCCATTAACCTTGTAA -3'
	PTEN	5'- GTACATGACCTTGGACGAGTTAT -3'
		5'- CTCCATTCGACCCTCACAAA -3'
PROMOTERS	HOXA9	5'- CGGCACGATCCCTTTACAT -3'
		5'- CCCGTCCAGCAGAACAATAA -3'
	MEIS1	5'- TGTAAGACGCGACCTGTTATG -3'
		5'- GCGTGTGTAAAGTGTGTGTTG -3'
	PBX3	5'- TCTTTCTTCTCCTCCCTCGT -3'
		5'- CAGCATCCTGGATTGATCGT -3'
	AFF3	5'- CTCTTTAGGAGCCACGATGATAC -3'
		5'- CTGGGTGTCGACTTCAAACT -3'

PTEN	5'- ATGTGGCGGGACTCTTTATG -3'	
	5'- GCGGCTCAACTCTCAAACT -3'	
BIRC3	5'- GCAGAGCTTTCCACCATGAA -3'	
	5'- GGTCAACTTCTCCAGGCTACTA -3'	
FILIP1L	5'- TGAAGTCCCTGAAGCAACAA -3'	
	5'- CTGTCGGGCTAACAAAGTCA -3'	
SMAD1	5'- TGTCTCGTTCCTTTCCCTTTAC -3'	
	5'- CTAACAAAGACAGGGAGCAGAG -3'	

3.2 RNA extraction and quantitative Real-time RT-PCR

RNA from the different hHP cellular models was extracted using the RNeasy Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instruction and TaqMan® Gold RT-PCR Kit (Applied Biosystems, Foster City, California, USA) was used for the reverse-transcription. The qRT-PCR was performed on 384-well plates, with each independent cDNA included in triplicate, using quantitative PCR with SYBR Green dyes (Applied Biosystems) on a 7900HT Fast Real-Time PCR System (Applied Biosystems), following the manufacturer's protocol. The expression of the endogenous human housekeeping gene *GAPDH* was used to normalize the data and they are expressed as the mean of $2^{-\Delta Ct}$ values obtained for each sample after normalization with normal human hematopoietic progenitors (SDS Program; Applied Biosystems). Statistical significances on level of expression of the selected target genes were determined with significance levels of *P*<0.05 and *P*<0.01 calculated using T-test. Dissociation curves were performed for each set of primers. In a house designed primers are listed in Table 5.

To analyse the expression of *hsa-miR-181b* in hHP-NH cellular model, total RNA, including miRNAs, was extracted using the miRNeasy Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions for *Purification of total RNA, including Small RNAs from Animal Cells*. 10 ng of RNA per sample were reverse transcribed using the MicroRNA Reverse Transcription Kit (Applied Biosystems). Three replicates of cDNA were prepared for each sample. qRT-PCR was performed using TaqMan Fast Universal PCR Master Mix on a 7500 Fast Real-Time PCR System following the manufacturer's protocol. Each independent cDNA was run in triplicate in a 384-well plate. Data were normalized on the expression of the artificial Homo sapiens miRNA control *RNU19* and relative expression was calculated using the comparative Ct method as described above. A pre-developed TaqMan Assays (Applied Biosystems) primer for *has-miR-181b* (462578_mat) was used.

Table 5: Sybr Green qPCR primers

HOXA9	5'-GTGGTTCTCCTCCAGTTGATAG-3'
	5'-AGTTGGCTGCTGGGTTATT-3'
MEIS1	5'-CTAACACACCCTTACCCTTCTG-3'
	5'-TCTATCATGGGCTGCACTATTC-3'
PBX3	5'-CATCACAGTGTCACAGGTATCC-3'
	5'-CAGCATAGAGGTTGGCTTCTT-3'
MYB	5'-TGTAGGGAGAGGTGGCATAA-3'
	5'-GTCTCTTGTGTGCCTGGTAAA-3'
MEF2C	5'-GCAACAGCAACACCTACATAA -3'
	5'-GTAGAAGGCAGGGAGAGATTTG-3'
HOXA7	5'-AGTTCCACTTCAACCGCTAC-3'
	5'-CCTTCGTCCTTATGCTCTTTCT-3'
BIRC3	5'-CAAGCCAGTTACCCTCATCTAC-3'
	5'-CTGAATGGTCTTCTCCAGGTTC-3'
BIRC2	5'-CTAGTCTGGGATCCACCTCTAA-3'
	5'-TGTTCCAAGGTGGGAGATAATG-3'
LMO2	5'-GATGACAATGCGGGTGAAAG-3'
	5'-GTCAGAGTTGATGAGGAGGTATC-3'
RARA	5'-CGAACAACAGCTCAGAACAAC-3'
	5'-GGCGAACTCCACAGTCTTAAT-3'
PML	5'-TGTACGCCTTCTCCATCAAAG-3'
	5'-GACTCCATCTTGATGACCTTCC-3'

3.3   Luciferase assay

Luciferase constructs were made following amplification of the three selected regulatory genomic regions of *HOXA9, PBX3* and *MEIS1* by subcloning in the pGL3-Promoter Vector (Promega Biotech Ibérica S.L) with the firefly luciferase gene, previous to the SV40 promoter. PCR products were obtained using the following primers: *HOXA9*, sense 5'-CACGG TACCGGGTTTCCCAGCTTTTTCTC-3', antisense 5'-TTCCTCGAGGAGAAGGGGGACAAGAGGAC-3'; *PBX3* sense 5'-CACGGTACCTGTTCTGAGACTGCCCATTG-3', antisense 5'-TTCCTCGAGA AATTGCTGCACAGGTGAGA-3'; *MEIS1* sense 5'-CACGGTACCCCCCTTTAACTTGAGAATCAGC-3', antisense 5'-TTCCTCGAGTTCAATGCAGCAGCTGAACTAT-3'. The insert identities were verified by sequencing and the following specific bacterial artificial chromosomes (BACs) were used as a PCR template: RP11–197I05 (*HOXA9*), RP11-638B02 (*PBX3*) and RP11-678O18 (*MEIS1*). A total of 10^5 HEK293FT cells were transfected with 500 ng of the firefly luciferase reporter vector containing the *HOXA9, PBX3* and *MEIS1* regulatory genomic regions and 200 ng of the control vector

containing Renilla luciferase pRL-CMV (Promega) using Calcium-phosphate mediated transfection protocol. Five hundred nanograms of pMSCV-NH or the empty vector were used to analyze the effect of NUP98-HOXA9 expression on luciferase signal. Luciferase assays were performed 48 h after transfection using the Dual Luciferase Reporter Assay System (Promega). The results were expressed as relative luciferase activity (%), calculated by normalizing the ratio of firefly luciferase to Renilla luciferase luminiscence. The reporter assay was independently performed three times for each of the experiments, including all samples in triplicate.

3.4   Gene expression microararys

RNAs from three different clones of hHP-NH cellular model and 5 primary samples from patients with t(7;11)(p15;p15) were isolated and hybridized in Array SurePrint G3 Human Gene Expression 8x60K v2 (Agilent Technologies) in NIMgenetics (www.nimgenetics.com, Madrid, Spain). Arrays were examined using the DNA Microarray Scanner C (Agilent Instruments) and microarray background subtraction was carried out using *normexp* method. To normalize the dataset, we performed loess within arrays normalization and quantiles between arrays normalization. Differentially expressed genes were obtained by applying linear models with R limma package[85] (Bioconductor project, http://www.bioconductor.org). To account for multiple hypotheses testing, the estimated significance level (p value) was adjusted using Benjamini & Hochberg False Discovery Rate (FDR) correction. Those genes with FDR <0.05 were selected as differentially expressed between hHP-NH or patient samples and controls (hHP-empty vector or hHP wt, respectively).

3.5   Gene Set Enrichment Analysis (GSEA)

We applied gene set enrichment analysis (GSEA) to integrate global gene expression data and the enrichment in NH targets. Genes were ranked based on limma moderated t statistic. After Kolmogorov-Smirnoff testing, those gene sets showing FDR <0.05, a well-established cut-off for the identification of biologically relevant gene sets [86], were considered enriched between classes under comparison. The enrichment score from the gene set after it has been normalized across analyzed gene set (NES) was also shown.

4   EVALUATION OF TREATMENT EFFICACY

4.1   Viability assay

HXR9 and CXR9 peptides were synthesized by Biosynthesis® (http://www.biosyn.com/) as previously described[47]. $5x10^3$ hHP-NH/hHP-emtpy vector sorted cells were plated per well in 96-well plates in triplicate and allowed to recover for 20 hours. Then they were treated with different doses of the active peptide HXR9 or the control peptide CXR9 in a range of 0 µM to 250 µM for 48 hours to determine the LC50. Cell viability was assessed by adding WST1 cell proliferation reagent (Roche, Basel, Switzerland) and reading the plates at 450 nm to measure optical density, according to the manufacturer's instructions. Then, we treated the hHP-NH with 13uM HXR9 (LC50) and measured cell viability at 0h, 24h, 48h, 72h and 144h.

To assess the effect of Panabinostat (LB589, purchased from Selleck) in hHP-NH, sorted cells were plated in triplicate with titrating doses of the drug (purchased from Selleck) (range, 0 µM to 1 µM) for 72 hours. Cell viability was analysed in a similar manner, using WST-1.

For each assay, three independent clones of hHP-NH/hHP-empty vector were used in three different experiments.

4.2   CFU assay

Aliquots of $8x10^3$ or $4x10^3$ of the hHP-NH/hHP-empty vector sorted cells were plated into 35 mm culture dishes in 1,5 ml of MethoCult M4230 methylcellulose medium (StemCell Technologies) containing 20% of BIT 9500 Serum substitute (STEMCELL Technologies) and 20ng/ul each of human recombinant SCF, IL-3, IL-6 and GM-CSF, and 3U/ml of EPO. Two aliquots of each cellular model were treated with 13uM of both HXR9 and CXR9 peptides (LC50) or with 4nM of LBH589. Cultures were incubated at 37°C in a humidified atmosphere for 10 days and then the number of colonies was counted for each case. This experiment was performed with two different clones of hHP-NH/hHP-empty vector for each treatment.

CFU assay with hHP-NUP98 cellular model was performed in the same manner but without treatment.

4.3 Apoptosis assay

3.75×10^5 hHP-NH/hHP-emtpy vector sorted cells were plated per well in 24-well plates in triplicate and allowed to recover for 20 hours. Then, the cells were treated with HXR9/CXR9 (13uM and 30uM) or LB589 (10nM and 30nM) for 24 hours. Apoptosis was assessed after the treatment using FITC annexin V and annexin V binding buffer (BD Biosciences) and DAPI staining (to exclude the nonviable cells) on FACSCanto II (BD Biosciences, San Jose, California, USA) following the manufacturer's instructions.

5 PROTEIN-PROTEIN INTERACTION STUDY

5.1 Protein co-immunoprecipitation and Immunoblotting

For p300, HDAC1, MEIS1 and PBX3 co-immunoprecipitation with NUP98-HOXA9, 4.10^6 HEK293FT cells were transfected with pMSCV-NH and/or pMSCV-HOXA9 vectors. 48h post-transfection, cells were washed with cold PBS + protease inhibitors (Roche), and lysed for 1 hour at 4°C with 0.5 ml of lysis buffer (50mM Tris-HCl pH 7.5, 150mM NaCl and 1% NP-40 and protease inhibitors (Roche)). The lysate was centrifugated at 13,000×rpm for 15 min at 4°C. The supernatants were transferred to new tubes and were precleared using 20ul Protein A/G Plus-Agarose beads (Santa Cruz) for 30 min at 4°C. Then, in accordance with the manufacturer's instructions, the different antibodies were added to the supernatants (anti-p300 (Santa Cruz), anti-HDAC1 (Thermo Scientific), anti-MEIS1 ChIP Grade (Abcam) and anti-PBX3 (Santa Cruz). After overnight incubation at 4°C, 20ul Protein A/G Plus-Agarose beads were added and incubated for another 2 hours. Beads were washed with 900 µl of lysis buffer 5 times, followed by centrifugation at 2400×rpm for 2 min at 4°C. Beads were washed one time with cold PBS and bound proteins were eluted by boiling with 40 µl of 2X Laemmli Buffer (100mM Tris-HCl pH 6.8, 20% Glycerol, 4% SDS, 300 mM β-mercaptoethanol, 0.2% Bromophenol Blue) for 3 min at 95°C. Input lysates and immunoprecipitated proteins were loaded onto SDS-PAGE gels (9% gel) followed by immunoblotting. After transfer, PVDF membranes (Millipore, Billerica, Massachusetts, USA) were incubated with anti-FLAG M2 (Sigma-Aldrich) primary antibody to detect Nup98-Hoxa9 and/or Hoxa9 and with anti-GAPDH or α-tubulin (monoclonal antibodies, CNIO) as endogenous control.

5.2   Mass Spectrometry

We used the 293FT-NH, 293FT-HOXA9 and 293FT-NUP98 models to conduct the corresponding immunopreticiptations using anti-FALG antibody (Sigma Aldrich), followed by mass spectrometry in three replicates for each model. Cellular model transducted with the retroviral empty vector with GFP gene flanked by FLAG-tag was submitted to the same process and used as a control in the assay.

Sample Preparation

Proteins in the pull-downs were subjected to label free proteome analysis. Samples were digested by means of the standard FASP protocol[1]. Briefly, samples were resuspended in UT buffer (8M urea in 100 Mm Tris-HCl, pH=8.01). Proteins were then reduced with 10 mM DTT, alkylated using 50 mM iodoacetamide for 20 min in the dark and the excess of reagents was washed out with UA twice. Proteins were digested with endoproteinase Lys-C (Wako) during 6 hours in a wet chamber (1:50 enzyme to substrate ratio). Finally, samples were diluted in 50 mM ammonium bicarbonate to reduce the urea concentration to 1M and subsequently digested with Trypsin Gold (Promega) overnight at 37 °C. Resulting peptides were further desalted and concentrated using homemade reversed phase micro-columns filled with Poros Oligo R3 beads (Life Technologies). The samples were dried using the Speed-Vac and dissolved in 30 µL of 0.1% formic acid (FA).

LC-MS/MS analysis

Desalted peptides were separated by reversed-phase chromatography using a nanoLC Ultra system (Eksigent), directly coupled with a LTQ-Orbitrap Velos instrument (Thermo Fisher Scientific) via nanoelectrospray source (ProxeonBiosystem). Peptides were loaded onto the column (Dr. Maisch, Reprosil-Pur C18 GmbH 3 ⬚m, 200x0.075 mm), with a previous trapping column step (Prot Trap Column 0.3 x 10 mm, ReproSil C18-AQ, 5 µm, 120Å, SGE), during 10 min with a flow rate of 2.5 ⬚l/min of loading buffer (0.1% FA). Elution from the column was made with a 120 min linear gradient (buffer A: 2% ACN, 0.1%FA; buffer B: 100% ACN, 0.1%FA) at 300 nL/min. The peptides were directly electrosprayed into the mass spectrometer using a PicoTip emitter (360/20 OD/ID µm tip ID 10 µm, New Objective) a 1.4 kV spray voltage with a heated capillary temperature of 325°C and S-Lens of 60%. Mass spectra were acquired in a data-dependent manner, with an automatic switch between MS and MS/MS scans using a top 20 method with a threshold signal of 800 counts. MS spectra were acquired with a resolution of

60000 (FWHM) at 400 m/z in the Orbitrap, scanning a mass range between 350 and 1500 *m/z*. Peptide fragmentation was performed using collision induced dissociation (CID) and fragment ions were detected in the linear ion trap. The normalized collision energy was set to 35%, the Q value to 0.25 and the activation time to 10 ms. The maximum ion injection times for the survey scan and the MS/MS scans were 500 ms and 100 ms respectively and the ion target values were set to 1E6 and 5000, respectively for each scan mode.

RESULTS

1   NUP98-HOXA9 human cellular models construction

First of all, we cloned the cDNA of *NUP98-HOXA9* fused to FLAG-tag at the 3'-end and HA-tag at the 5'-end into the retroviral vector pMSCV-IRES-GFP (Figure 13A). We generated retroviruses in HEK293T cells and efficiently transduced both Human Hematopoietic Progenitors (hHP) and the HEK293FT human cell line. hHP (CD34+) were isolated from human cord blood samples using CD34 immunomagnetic microbeads and maintained in cytokine-stimulated suspension culture. Green Fluorescence protein (GFP) allowed us to select the transducted cells using BD Influx™ cell sorter. Thus, we established two human cellular models that constitutively express *NUP98-HOXA9*: hHP-NH and 293FT-NH (Figure 13B and C). According to the published data[27,53], long-term cultures of the hHP-NH cellular model clearly showed an increase in proliferation due to the presence of the fusion protein, supporting so far the fitness of the model for further studies (Figure 13D). However, hHP-NH cells grow 3.5 weeks more, on average, than scramble cells, but only up to 8 weeks. This fact makes this cellular model very hard to work with, especially if we need to obtain big amounts of cellular material. Therefore, to solve this problem, we have also created the 293FT-NH model from a human immortalized cell line.

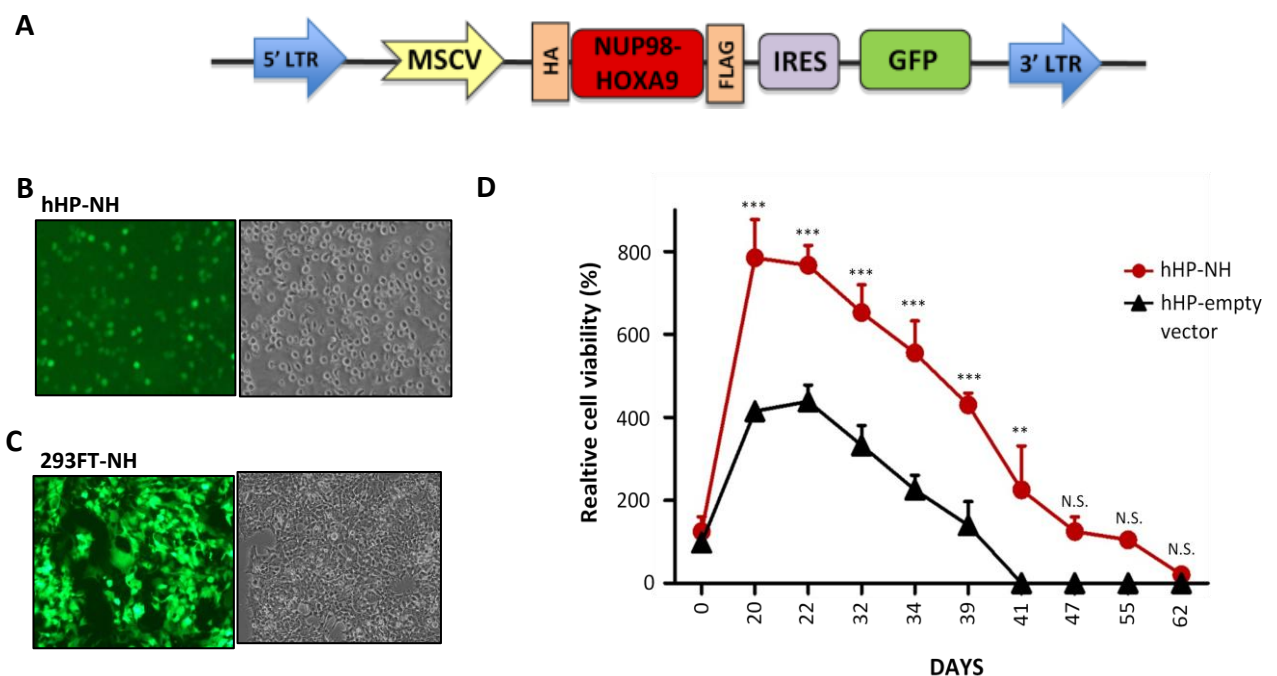 **Figure 13: NUP98-HOXA9 human cellular models construction**. (A) Representation of the pMSCV-NH retroviral vector. ((B) hHP-NH cellular model: Microscopy image of hHP transduced with pMSCV-NH retroviruses and sorted (GFP+) (C) 293FT-NH cellular model: Microscopy image of HEK293FT cells transduced with pMSCV-NH retroviruses and sorted (GFP+) D) Long-term culture of hHP-NH and hHP-empty vector cellular models for 62 days.

2   Identification of NUP98-HOXA9 genomic binding regions

To explore the transcription factor role of NH, we decided to perform a ChIP-seq analysis in order to identify its DNA binding sites, trying to extract from this data an idea about the possible oncogenic mechanism triggered by the fusion protein.

ChIP-seq approach consists in a chromatin immunoprecipitation (ChIP) followed by a massive sequencing of the eluted DNA fragments. We used the 293FT-NH cellular model to immunoprecipitate NH attached to the DNA, using an anti-FLAG antibody. The sequencing of these NH-binding sites allowed us to identify 4471 significant genomic regions (FDR < 0,05), located within a 5 kb distance from annotated Transcrption Start Sites (TSS), that corresponded to regulatory regions of 1363 different genes and 17 miRNAs (Figure 14A, Table 6 and Table S1).

A detailed sequence analysis of all DNA bounded *NH* regions using MEME-ChIP[84] identified a significant enrichment in consensus binding motifs of important leukemic transcription factors, as detailed in Table 7. We detected the presence of binding sites for several *HOX* genes, including *HOXA9*, which confirms the importance of DNA binding of the homeobox domain of this fusion protein. We also identified a highly significant enrichment of the motif CA/gTTT, present in one-third of all binding sites. This motif has not previously been associated with any known transcription factor; it appears to be a novel specific binding site for *NH*.

Table 6: Target miRNAs of NUP98-HOXA9 identified in the ChIP-seq

miRNA	FDR
hsa-miR-181A2	0.00
hsa-miR-181b	0.00
hsa-miR-586	0.00
hsa-miR-548F1	0.00
hsa-miR-581	0.00
hsa-miR-550-1	0.00
hsa-miR-194-1	0.00
hsa-miR-128-1	0.00
hsa-miR-615	0.00
hsa-miR-297	0.00
hsa-miR-125B2	0.00
hsa-miR-576	0.00
hsa-LET7I	0.00
hsa-miR-181A1	0.01
hsa-miR-1245	0.02
hsa-miR-30A	0.02
hsa-miR-626	0.04
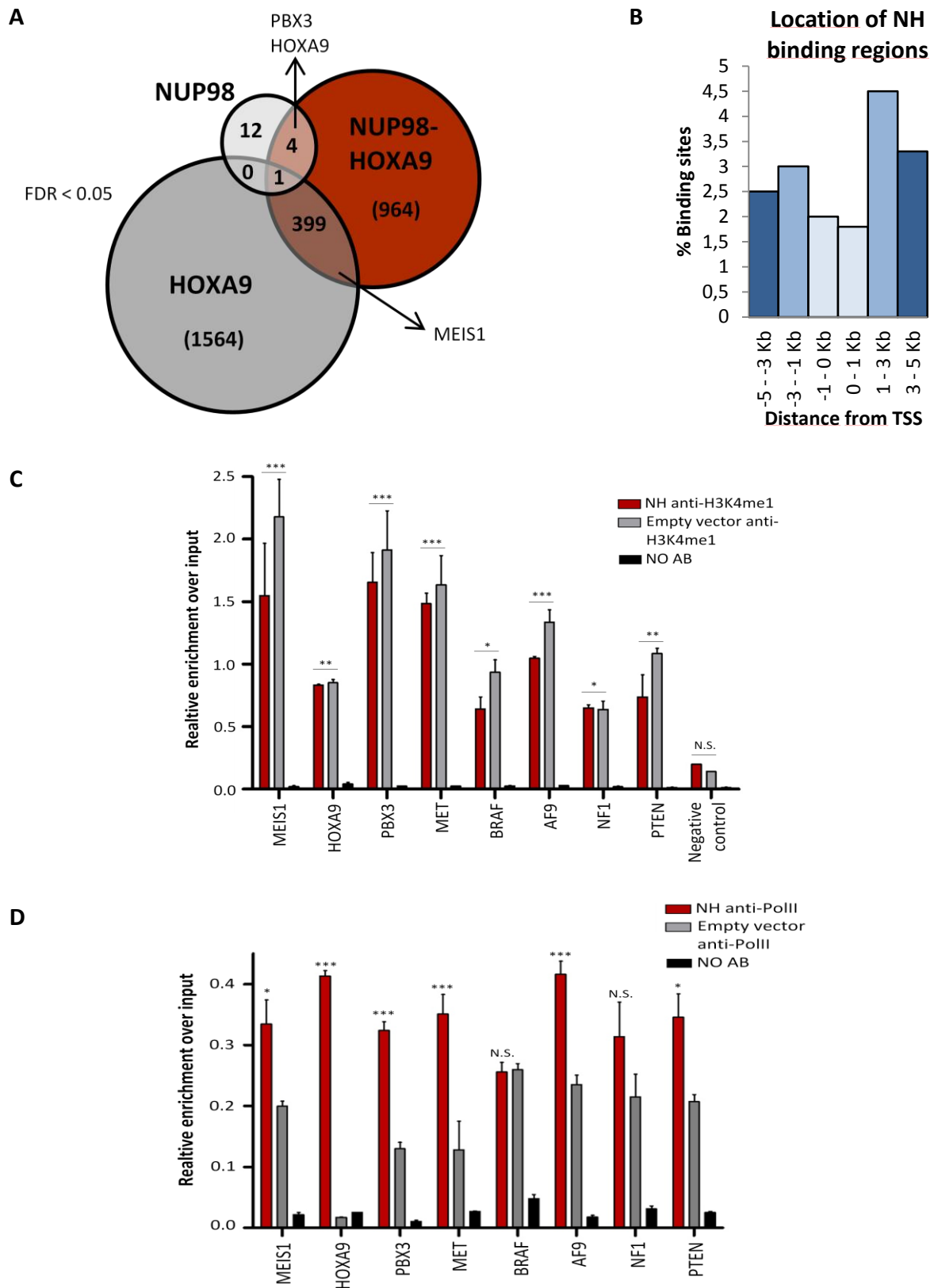
Figure 14: Identification of NUP98-HOXA9 genomic binding regions (A) Venn diagrams of the overlapping NH, HOXA9 and NUP98 target genes indentified with the ChIP-seq approach on HEK293F human models (FDR<0.05). (B) Plot that represents the localization of NH binding sites in distance (Kb) from the TSS. (C) H3K4me1 qChIP fold enrichment on the selected NH target regions. MEIS1 promoter region was used as a negative control. Average of 3 experiments is shown. Error bars represent SEM.Plot that represents the main functions of the NUP98 wt target genes (IPA output) (D) PolII qChIP fold enrichment on the eight selected NH target regions. Average of 3 experiments is shown.

Table 7: Identification of DNA binding motifs in NH target regions using MEME-ChIP.
(MEME-ChIP p-value < 1.10^{-6})

Motif Found (DREME program)	Known transcription factor	Positive sequences matching the motif	P-value	E-value (score MEME-ChIP)
TTTTAT	Hoxd13 Hoxd10	1013/4471	1.8e-35	1.2e-30
TTTATG	Hoxa9 Hoxc9 CDX2	330/4471	5.9e-7	0.037
ATTAC	Hoxa10 Irx3 CDX1	696/4471	3e-12	2e-7
CATTT	Unknown	1421/4471	8.5e-24	5.9e-19
CAGTG	TAL1::GATA1 complex	1450/4471	3.7e-21	2.5e-16
GGAAA	NFATC2 RELA STAT1	809/4471	9e-8	0.0058
CATAAA	ARID3A	525/4471	9.4e-13	6.1e-8

HOX genes

Possible NH binding motive

A further analysis of these binding sequences indicated that they were preferentially located more than 1 kb upstream from the TSS (only less than 20% of them located at promoter regions) (Figure 14B). This distribution strongly suggested that the enhancer regions likely were the preferred binding targets of the fusion gene. The chromatin mark that best predicts both poised and active enhancers is the monomethylated form of H3 lysine 4 (H3K4me1)[67]. We performed a ChIP by immunoprecipitating this epigenetic mark both in presence and absence of the fusion protein using the 293FT cellular models. Then, from the eluted DNA, we performed a quantitative PCR to analyzed the binding sites of NH corresponding to eight selected target genes with an undoubted role in myeloid leukemogenesis (*MEIS1, HOXA9, PBX3, MET, BRAF, AF9, PTEN* and *NF1).* We were able to demonstrate an enrichment of H3K4me1 at the eight genomic regions (Figure 14C). We also detected an increase in the presence of PolII at these NH-binding sites when the fusion protein is expressed, which is consistent with an active form of enhancers[67,70,87] (Figure 14D). In addition, we applied the oPOSSUM system[83], an *in silico* analysis to identify over-represented transcription factor binding sites in the regulatory regions from a set

of genes. We analysed the top 100 NH target genes obtained in the ChIP-seq (FDR < 0,0001). The transcription factors that may be binding to the same regulatory regions than NH are listed in Table 8 (only shown those with a Fisher score <0,01). Notably, there is an over-representation of FOX proteins, considered pioneer factors that are responsible for the initial recognition of enhancers and that facilitate the binding of additional transcription factors to those regulatory sites [67]. Taken together these approaches support a working model for NH where its role in transcriptional regulation is mostly mediated by its binding to enhancers.

Table 8: Identification of known transcription factor binding motifs in NH target regions using oOPOSSUM
(Fisher score < 0.005. Fox proteins in orange)

TRASCRIPTION FACTOR	Fisher score
HLF	$3,70 \times 10^7$
NFIL3	$4,68 \times 10^6$
CF2-II	$1,63 \times 10^5$
Broad-complex_1	$3,65 \times 10^5$
ATHB5	$5,68 \times 10^5$
HMG-IY	$6,03 \times 10^5$
Ddit3-Cebpa	$7,43 \times 10^5$
hb	$8,19 \times 10^5$
HNF1A	$8,45 \times 10^5$
Broad-complex_4	$9,96 \times 10^5$
Lhx3	$1,30 \times 10^4$
Athb-1	$1,64 \times 10^4$
SQUA	$2,79 \times 10^4$
Foxd3	$2,83 \times 10^4$
Lhx3	$4,67 \times 10^4$
Foxq1	$6,74 \times 10^4$
Broad-complex_3	$8,66 \times 10^4$
NKX3-1	$1,65 \times 10^3$
Foxa2	$2,17 \times 10^3$
FOXI1	$2,30 \times 10^3$
TBP	$2,30 \times 10^3$
Ovo	$2,66 \times 10^3$
MYB.ph3	$2,81 \times 10^3$
FOXD1	$3,64 \times 10^3$
SOX9	$4,83 \times 10^3$
PEND	$6,88 \times 10^3$
SRY	$8,40 \times 10^3$

3 Identification of *HOXA9* and *NUP98 wt* genomic binding regions

We further investigated the functional contribution to the DNA binding profile of the two moieties that compose the fusion protein. We cloned the coding region of *HOXA9 wt* and *NUP98 wt* fused to FLAG-tag and HA-tag in the retroviral vector pMSCV-IRES-GFP (Figure 15A). We then established two new cellular models, 293FT-HOXA9 and 293FT-NUP98, which allowed us to perform separate ChIP-seq analyses. By combining the obtained ChIP-seq data (Tables S3 and S4), we observed that one third of the NH target genes were common to HOXA9 wt target genes (Figure 14A), indicating that the homeobox contribution to the fusion protein is clearly involved in its DNA binding. Ingenuity Pathways Analysis (IPA, http://www.ingenuity.com) allowed us to demonstrate that HOXA9 target genes are mostly involved in cell cycle, leukemia and hematopoiesis (Figure 15B) and that the location of the sites has a very similar distribution to those of NH, which would also correspond mainly to enhancer regions, consistently with previous results[46] (Figure 15C).

Importantly, we assessed the ChIP-seq of NUP98 wt, only 17 target genes were identified as direct genomic targets of the protein (Anexo table 1). Strikingly, all of them appeared to have a directly role in either leukemogenesis or hematopoiesis (Figure 15D). In addition, analyzing the location of the binding regions, we observed that the possible modulatory effect of NUP98 wt on gene transcription would be leading primarily by binding to promoters, since half of the regions are located between 0 and 1 kb upstream of the TSS of the possible target genes and the other half to more than 5 kb (Figure 15E).

To go a step further into this possible role of NUP98 in hematopoietic differentiation, we generated a hHP cellular model that overexpress NUP98 wt (hHP-NUP98) and demonstrated that most of the selected candidate targets were upregulated. Interestingly, NUP98 wt was also able to induce the downregulation of *BIRC3*, a well known tumor suppressor gene (Figure 16A). We then performed a common Colony Forming Unit (CFU) Assay to compare the ability of differentiation and proliferation of the hHP-NUP98 cells with the hHP-empty vector model. CFU assay is an *in vitro* assay based on the ability of hematopoietic progenitors to proliferate and differentiate into colonies in a semi-solid media in response to cytokine stimulation. The colonies formed can be counted and characterized according to their unique morphology. We importantly showed a highly significant increase in colony formation of all types (CFU-G, CFU-GM, BFU-E, CFU-M and CFU-GEMM) in hHP-NUP98 (Figure 16B and 16C). These results lead us to validate

the function of NUP98 wt in hematopoiesis and think about a possible oncogenic role in hematological malignancies, which would need to be further studied.
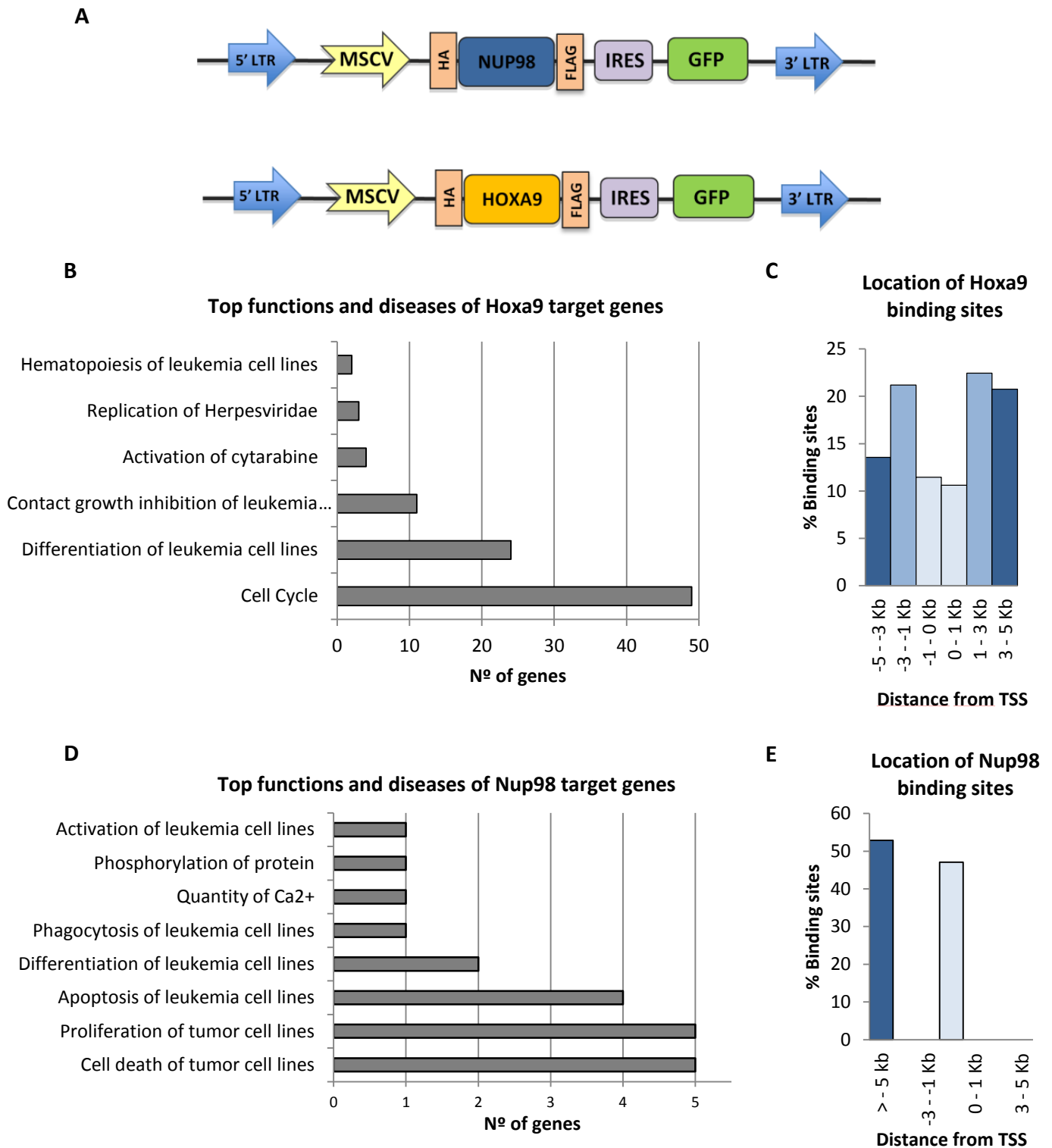
Figure 15: Identification of *HOXA9* and *NUP98 wt* genomic binding regions. (A) Representation of the pMSCV-NUP98 and pMSCV-HOXA9 retroviral vectors. (B) Plot that represents the main functions of the *HOXA9 wt* target genes (IPA output) (C) Plot that represents the localization of NUP98 wt binding sites in distance (Kb) from the TSS. (D) Plot that represents the main functions of the NUP98 wt target genes (IPA output) (D) Plot that represents the localization of NUP98 wt binding sites in distance (Kb) from the TSS.
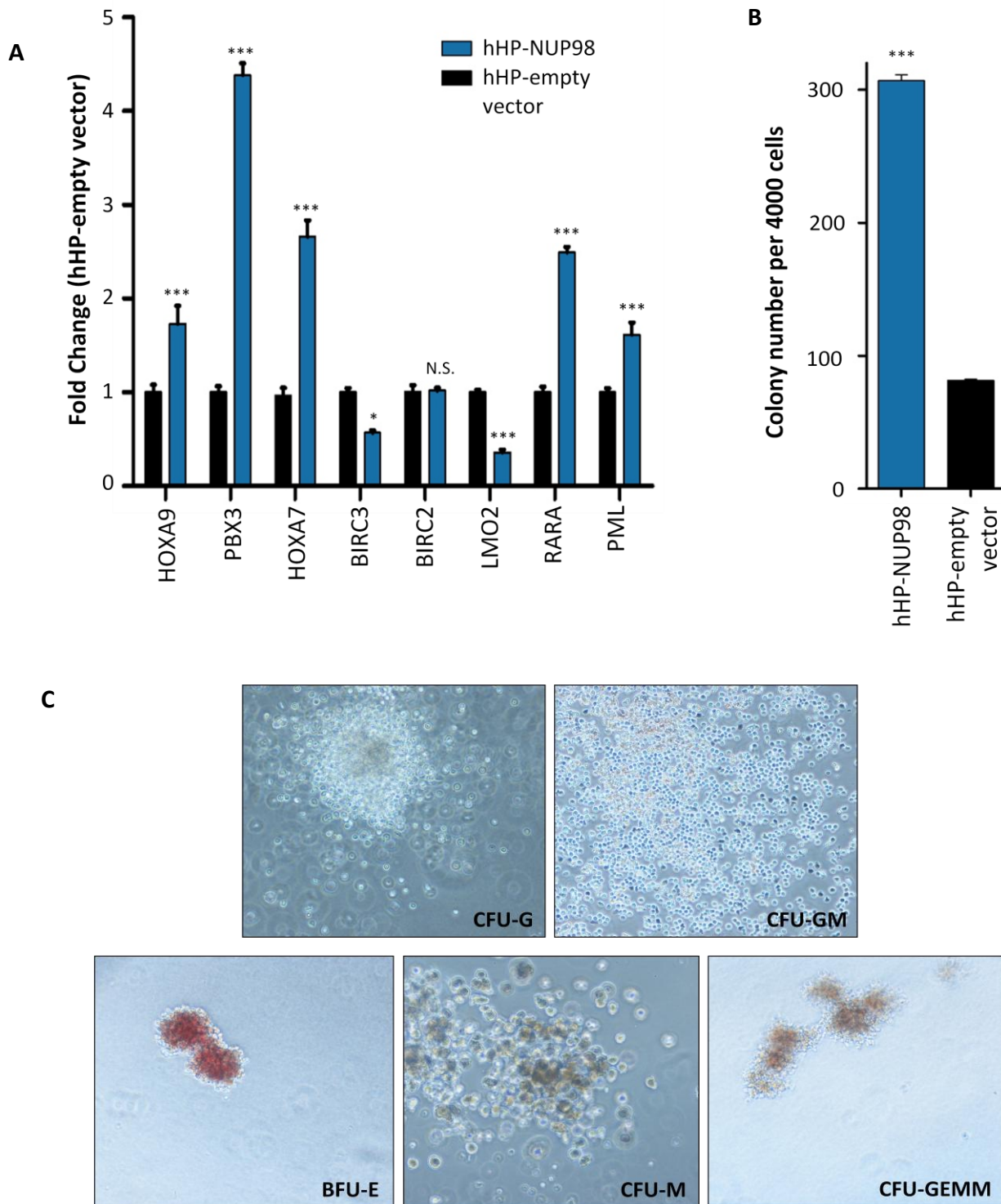
Figure 16: Nup98 plays a role in hematopoiesis. (A) Expression analysis by qRT-PCR for the eight selected NUP98 wt target genes in the *hHP-NUP98* cellular model. The expression of the endogenous human housekeeping gene GAPDH was used to normalize the data and they are expressed as the mean of $2^{-\Delta Ct}$ values obtained for each sample after normalization with hHP-empty vector model. Error bars represent SEM. (B) Colony-forming assay with hHP-NUP98 cells and hHP-emtpy vector cells. Average number of colonies per dish from 2 independent experiments after 10 days are shown (only colonies with > 50 cells/colony were counted). Error bars represent SEM. (C) Microscopy images of the different types of colonies identified in the hHP-Nup98 cellular model CFU assay: CFU-G (Colony forming unit-granulocyte), CFU-GM (Colony forming unit-granulocyte, macrophage), BFU-E (Burst forming unit-erythorid), CFU-M (Colony forming unit-macrophage) and CFU-GEMM (Colony forming unit-granulocyte, erythrocyte, macrophage, megakaryocyte)

4 NUP98-HOXA9 target genes analysis

Our next step was to focus on the study of the identify NH target genes and try to find some molecular mechanism that may explain, at least partially, its effect on leukemogenesis. When we analyzed this set of target genes using Ingenuity Pathways Analysis, we found enrichment in functions such as survival, cell cycle, leukemia and development and function of the hematologic system (Figure 17A). Among these genes, we found oncogenes implicated in the development of leukemia, but also interesting tumor suppressor genes such as *NF1*, *PTEN or BIRC3*.

Furthermore, regarding the canonical pathways associated to NH, only Protein kinase A signaling were found significantly enriched, whereas we found that most of the genes were involved in a wide range of different molecular pathways. In light of these data, we conclude that the target genes of NH are involved on different molecular pathways but targeting the same biological functions.

Importantly, we validated, in the human hematopoietic progenitor model (hHP-NH), the DNA direct binding of NH to the regulatory regions of the eight leukemic target genes that we selected before (Figure 17B), meaning that the ChIP-seq results can be perfectly extrapolated from the cellular model with 293FT cell line to the human hematopoietic progenitor model.

A

Top Functions and diseases of *Nup98-Hoxa9* target genes

Nº of genes

B
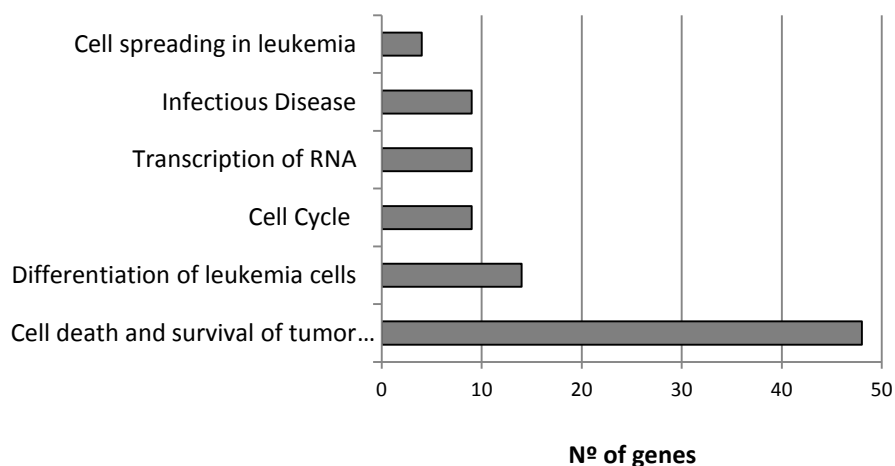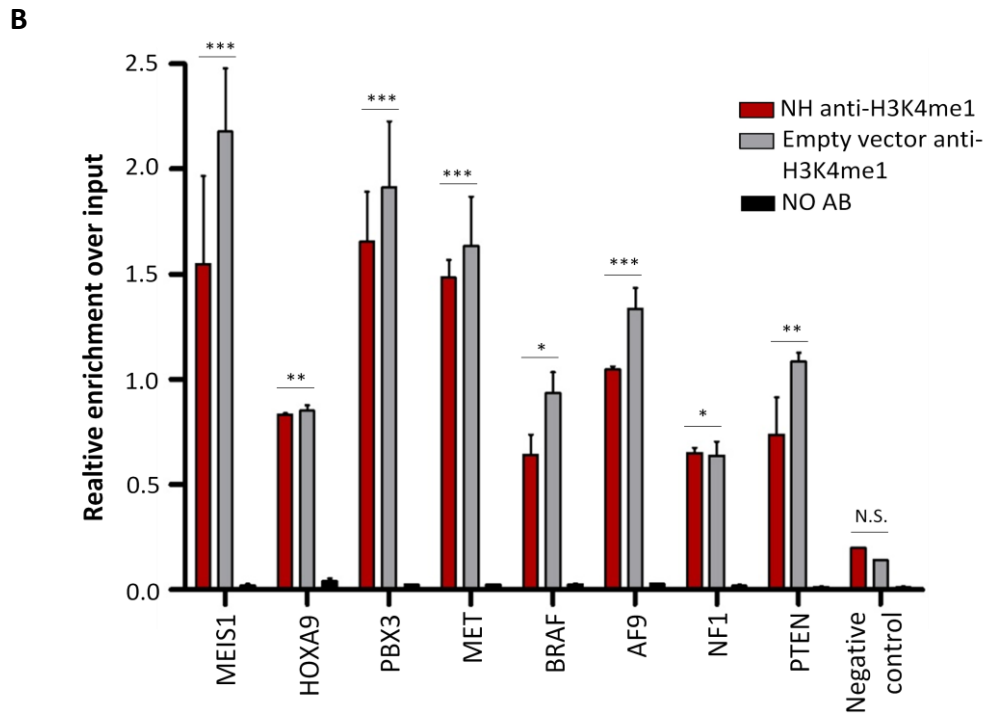
Figure 17: NUP98-HOXA9 target genes. (A) Plot that represents the main functions of the NH target genes (IPA output) (B) qChIP fold enrichment of NH in hHP-NH cellular model and hHP in the eight selected regions. The average of 3 experiments is shown. Error bars represent standard error of the mean (SEM). (C) PolII qChIP fold enrichment on the eight selected NH target regions. Average of 3 experiments is shown. qChIP fold enrichment of NH in hHP-NH cellular model and hHP in the eight selected regions. The average of 3 experiments is shown. Error bars represent standard error of the mean (SEM).

4.1 HOXA9-PBX3-MEIS axis

The ChIP-seq data demonstrated that *PBX3* and *HOXA9* are both common targets of the fusion protein as well as of the wild type NUP98 (Figure 14A). In addition, we also found that both the fusion protein and HOXA9 bind to the regulatory regions of *MEIS1*. These three transcription factors, PBX3, HOXA9 and MEIS1*,* form an activator complex that regulates the expression of genes involved in the AML induction[26,47,48,88,89]. Accordingly, we showed the significantly overexpression of these three genes in the hHP-NH cellular model, when comparing the hHP transducted with the empty vector (Figure 18A).

To determine whether NH has a direct involvement in the deregulation of *PBX3, HOXA9* and *MEIS1*, we subcloned their identified enhancer regions upsetream to the luciferase gene. For the three regions, we observed a significant increase in the luciferase activity when NH was also expressed (Figure 18B) demonstrating that NH directly induces the overexpression of the three target genes by binding to their enhancer regions. Remarkably, the findings were observed also in primary samples of three patients bearing the fusion gene, where we also showed the overexpression of these target genes (Figure 18C)

From the ChIP-seq analysis of NH, we also found that the fusion gene binds to the regulatory regions of three *hsa-miR-181* family members (*hsa-miR-181a1, hsa-miR-181a2* and *hsa-miR-181b*). *PBX3 and HOXA9* are potential targets of the *miR-181* family (particularly, *miR-181b*). In our hHP-NH cellular model, we observed a significant downregulation (Figure 18D) of *miR-181b,* most likely contributing to greater overexpression of the HOX/PBX3 signature.

All together, these data highlighted the important role of these three target genes in the oncogenic process induced by *NH*.

Thus, given the potential relevance of the HOXA9-PBX3-MEIS1 axis, we assessed the molecular interactions between PBX3, HOXA9 and MEIS1, and their relationship with the chimeric NH protein. We generated a new cellular model from 293FT cells by co-expressing NH and *HOXA9* and performed different co-immunoprecipitation studies (Co-IPs). We demonstrated that *HOXA9* forms a stable heterocomplex with the *PBX3* and *MEIS1* proteins (Figure 19A). Furthermore, among the known target genes of MEIS1-HOXA9-PBX3, we selected *MYB, MEF2C* and *FLT3* and demonstrated that the complex was sufficiently functional to overexpress them in the hHP-NH cellular model (Figure 18E).
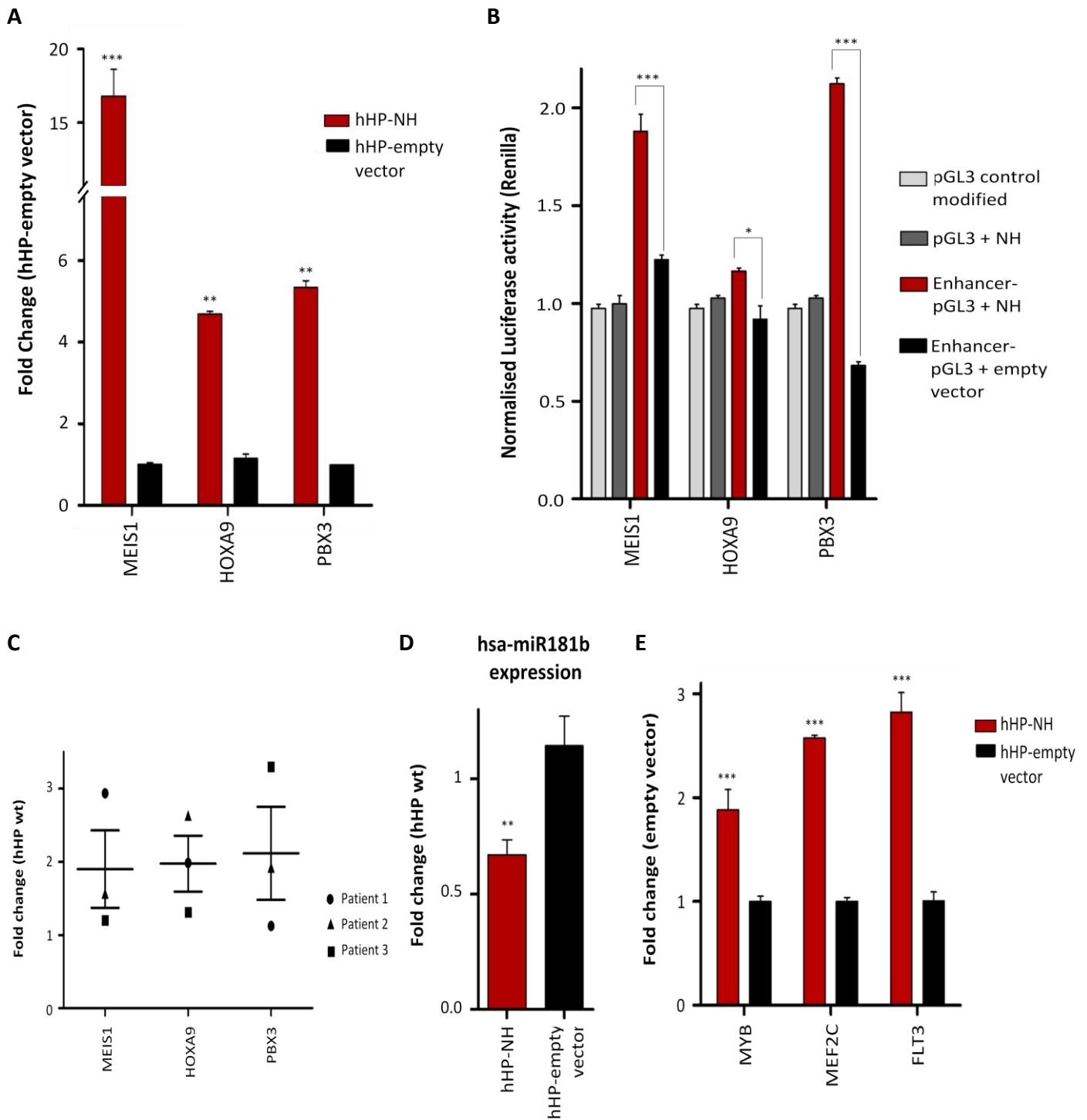
Figure 18: HOXA9-PBX3-MEIS axis. (A) Expression analysis by qRT-PCR of *MEIS1*, *HOXA9* and *PBX3* in the *hHP-NH* cellular model. The expression of the endogenous human housekeeping gene GAPDH was used to normalize the data and they are expressed as the mean of $2^{-\Delta Ct}$ values obtained for each sample after normalization with hHP-empty vector model. Error bars represent SEM. (B) Luciferase assay to evaluate the activity of NH on the enhancers of MEIS1, HOXA9 and PBX3. Data are presented as mean from two separate experiments with n=3 for each experiment. Error bars represent SEM. (C) Expression analysis by qRT-PCR that showed the overexpression of *MEIS1*, *HOXA9* and *PBX3* in three primary samples of patients with t(7;11)(p15;p15). Error bars represent SEM. (D) Expression analysis by qRT-PCR that showed the downregulation of *has-miR181b* in *hHP-NH* model compared to the *hHP-empty vector*. Data were normalized on the expression of the artificial Homo sapiens miRNA control RNU19. (E) Expression analysis by qRT-PCR of *MYB*, *MEF2C* and *FLT3* in the *hHP-NH* cellular model.

4.1.1 **Interaction between NUP98-HOXA9 and PBX3**

The chimeric NUP98-HOXA9 retains the short ANWL motif of HOXA9, which binds the PBX family, but the MEIS1 *i*nteraction domain is lost [26]. Consistent with this, we demonstrated that NH was indeed able to interact with the complex through PBX3 but not through MEIS1 by performing co-immunoprecipitations in the 293FT cellular model (Figure 19A). We scanned the DNA target sequences of NH for the HOXA9-MEIS1-PBX binding motif (ATGATTTATGGC[48]) and found 1981 co-occurrences with p ≤ 0.0001 using FIMO (Table S2) suggesting a strong cooperation between NH and the complex itself in inducing the final transcriptional profile.

NH could recruit the HOXA9-PBX3-MEIS1 complex through PBX3 and direct it to its own regulatory target regions. However, it has been shown that the complex co-occupies cellular promoters that drive leukemogenesis[90]; therefore, the loop that communicates enhancers with promoters could enable the binding of the complex to the promoters of the same genes thanks to the interaction with NH.

 To investigate this aspect, we cloned the coding region of PBX3 fused to FLAG-tag (Figure 19B) and the coding region of NH without any tag in the retroviral vector. We generated a new cellular model with these vectors by co-transfecting 293FT cells and performed two distinct qChIP approaches. On one hand, we analyzed the PBX3 enrichment in the same regulatory NH target regions and we observed that PBX3 is only binding to the enhancers of *MEIS1, MET, BRAF* and *NF1* in the presence of NH and not interact to any of the enhancers whether NH is not expressed (Figure 19C). Likewise, we observed that PBX3 is not binding to any promoter neither in the absence nor presence of the fusion protein (Figure 19D).

In the light of these results, we could conclude that the interaction of NH with the complex seems not to be crucial to its oncogenic functions.
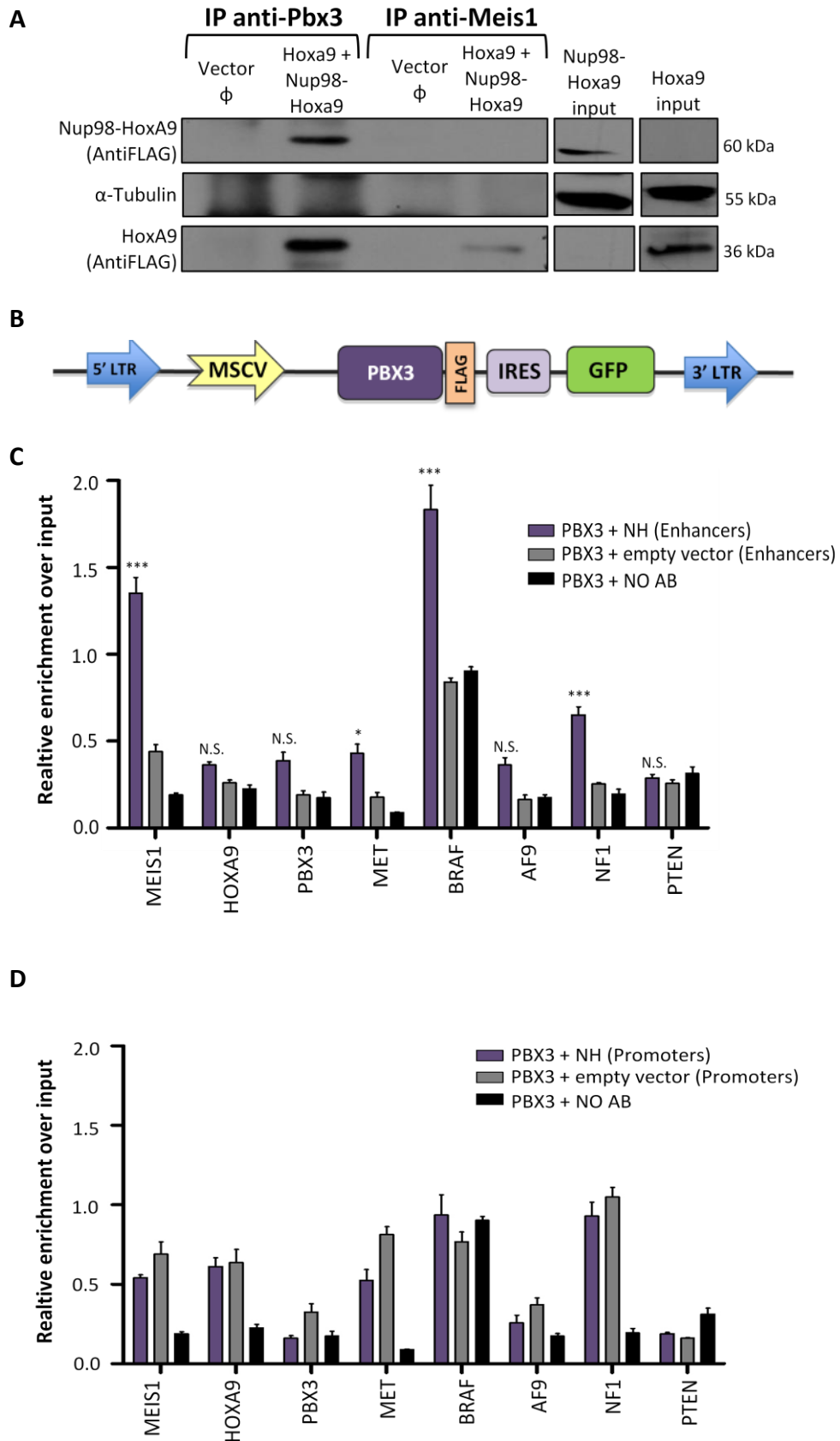
Figure 19: Interaction of NUP98-HOXA9 with PBX3. (A) Analysis of NH and MEIS1-HOXA9-PBX3 complex interactions by co-immunoprecipitation. Input lysates and immunoprecipitated proteins were analyze by immunoblotting using anti-FLAG antibody. α-tubulin protein levels were used as load control (B) Representation of the pMSCV-PBX3 retroviral vector. (C-D) PBX3 (anti-FLAG) qChIP fold enrichment on the eight selected target genes: enahcers (C) and promoters (D). Error bars represent standard error of the mean (SEM).

4.1.2   **Chemical disruption of the HOXA9-PBX3-MEIS complex**

We next decided to evaluate the biological relevance of this oncogenic axis in the NH-AML context using the HXR9 peptide that works as a specific inhibitor of the HOXA9-PBX interactions. HXR9 is an 18-amino-acid peptide that harbors the hexapeptide sequence of HOXA9 that can bind to PBX3 and 9 C-terminal arginine residues (R9) that facilitate cell entry[47]. This peptide seems to have high therapeutic potential in several tumor types including cytogenetically abnormal AML (CA-AML) that overexpresses HOX-PBX signature[47,91,92]. We evaluated the sensitivity to HXR9 of our hHP-NH model. Sorted cells were plated in triplicate with titrating doses of the drug to obtain the median Lethal Dose (LC_{50}, is the dose required to kill half of the cells after 24 hours of treatment) that was 13µM. Supporting our hypothesis, we observed an inhibitory effect on cell viability (30%) after 13µM HXR9 treatment at different time points. (Figure 20A). As control, we used the treatment with the peptide CXR9 that lacks a functional hexapeptide sequence but includes de R9 sequence. Furthermore, we also treated hHP-NH colony forming cells with 13µM HXR9 and observed a significant decrease in the number and the size of the colonies compared with cells treated with the control peptide. This response seemed to be conditioned by the presence of the chimeric protein since it was not observed when we treated wild type human hematopoietic cells (hHP) (Figure 20B and 20C). Apoptotic cells were identified using flow cytometry and APC-labeled anti-annexin V antibody. Consistently, an increase in apoptosis was observed when cells were treated with 13µM and 30µM HXR9 for 24 hours compared with controls (Figure 20D).

These results showed the role of NH in the activation of the HOXA9-PBX3-MEIS1 axis and provided evidence for its actionability as a new therapeutic approach for this leukemia.
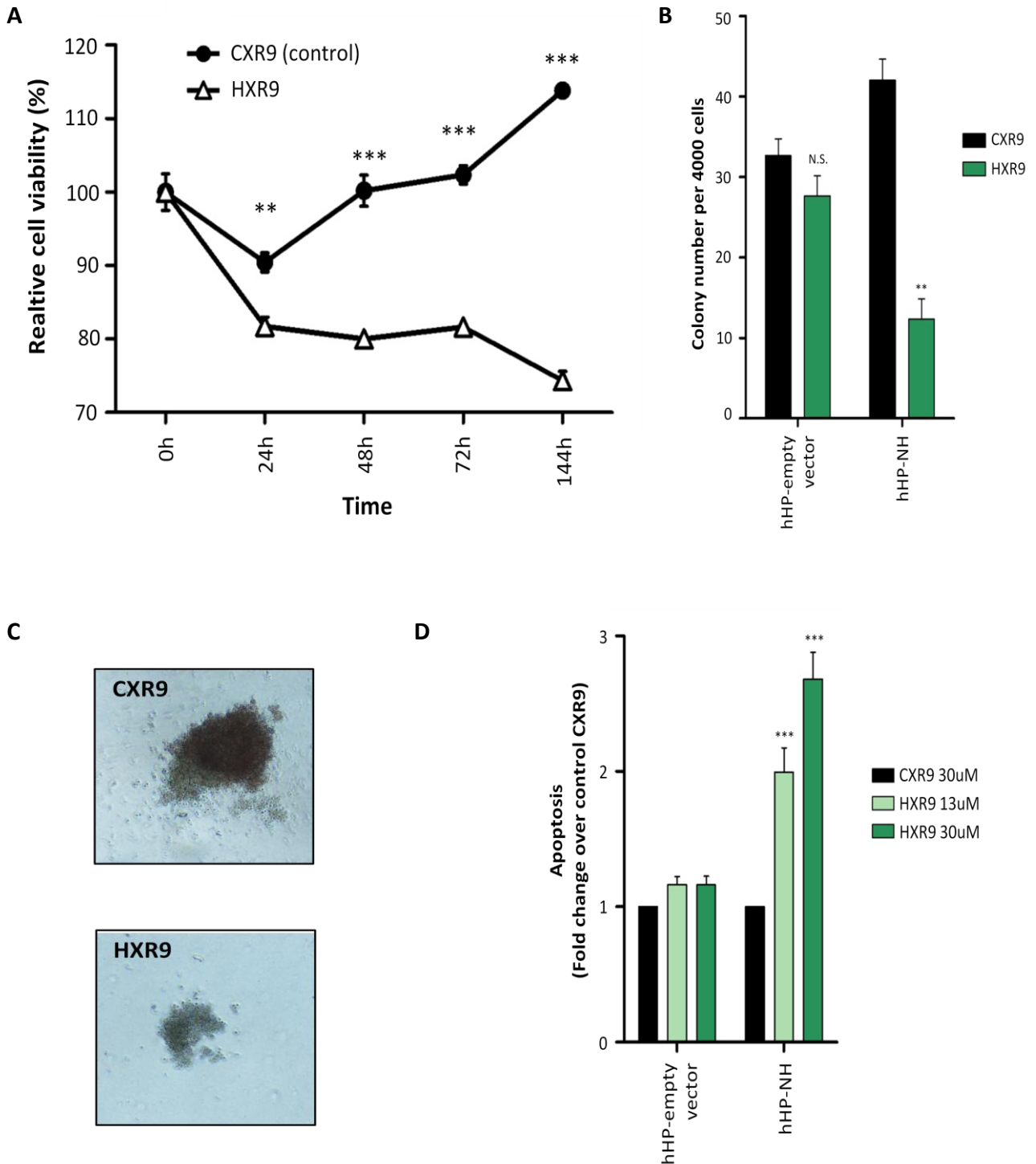
Figure 20: Chemical disruption of the HOXA9-PBX3-MEIS complex. (A) Analysis of the response to HXR9 treatment. hHP-NH cells were exposed to 13uM of HXR9/CXR9 and cell viability was assessed at different time points using WST-1. Averaged normalized optical density (OD) values of three independent experiments are shown. (B) Colony-forming assay with hHP-NH cells and hHP-emtpy vector cells treated with 13uM HXR9/CXR9. Average number of colonies per dish from 2 independent experiments after 10 days are shown (only colonies with > 50 cells/colony were counted). Error bars represent SEM. (C) Morphology of colonies of hHP-NH treated cells. (D) hHP-NH and hHP-empty vector cell apoptosis assay after the treatment with 13 uM and 30uM of HXR9 peptide for 24 hours. Cells were stained with annexin V and analyzed using flow cytometry. The plot shows the averaged fold changes of 3 independent experiments compared to CXR9 (30uM). Error bars represent SEM.

4.2 Effect of NUP98-HOXA9 on its other target genes

To gain insight into the molecular mechanisms that could explain the global expression profiling of this oncogenic model, we performed gene expression analysis of three independent clones of hHP-NH (Table S5) and three independent clones of hHP-empty vecto were used as control. From previous studies, only a subset of these data is publically available [27,53-55] and the most salient feature of these clones was a preponderance of upregulated genes. However, we did not want to focus our study on analyzing the genes up or down-regulated in the hHP-NH model since it has already been discussed before, but understand the effect of NH in the regulation of gene expression. Thus, we wanted to discover what was happening with the target genes of the fusion protein and applying GSEA analysis to our expression profile data, we found that NH seemed to induce overexpression and down-regulation of its target genes, previously identified in the ChIP-seq experiments (Figure 21A).

Since the chimeric protein retains the FG repeat domain of NUP98, NH may interact with different transcriptional regulatory cofactors, including the known activator p300 and the inhibitor HDAC1. To demonstrate that the activator-repressor role of NH may rely on these interactions, we assessed their potential influence on global gene regulation in our human cellular model. As expected, the Co-IPs performed in 293FT cellular models demonstrated the direct interactions of NH with both p300 and HDAC1 (Figure 21B).

We then used expression data to assess the role of these interactions in the transcriptional regulation driven by NH. We selected eight target genes, four (*MEIS1, HOXA9, PBX3* and *AFF3*) that were upregulated in the hHP-NH model and four (*BIRC3, SMAD1, FILIP1L* and *PTEN*) that were downregulated. We then designed primers to analyze the occupancy of their promoters by performing different qChIP assays. We found that p300 bound to the promoters of the overexpressed genes only when NH was present. On the other hand, we also observed an enrichment of HDAC1 within the promoter regions of the downregulated genes when the fusion protein was present, thereby demonstrating its hypothesised dual role (Figures 21C and 21D).
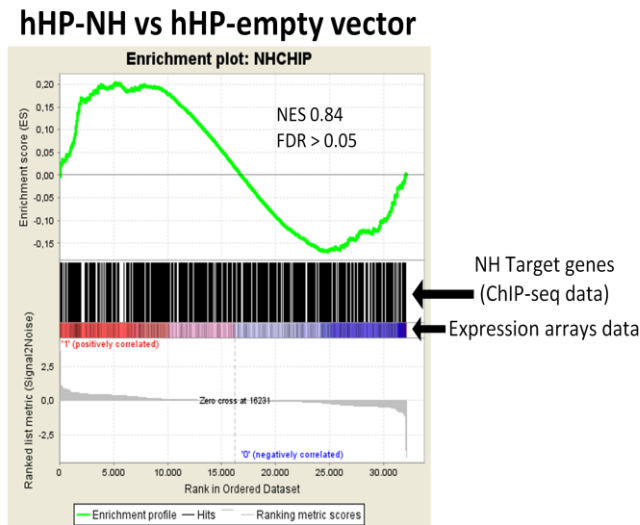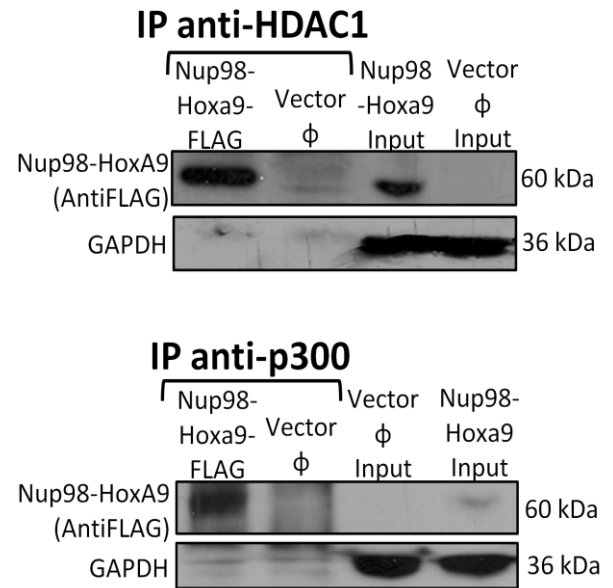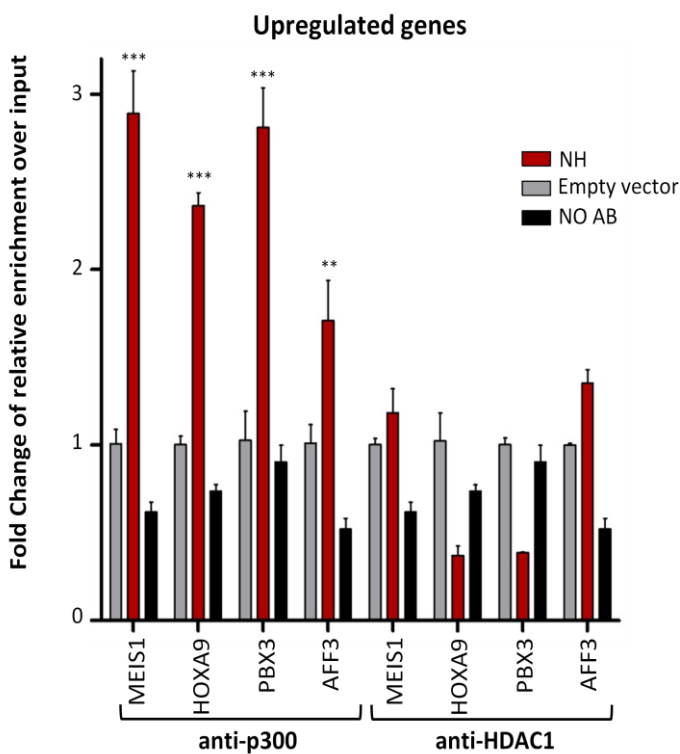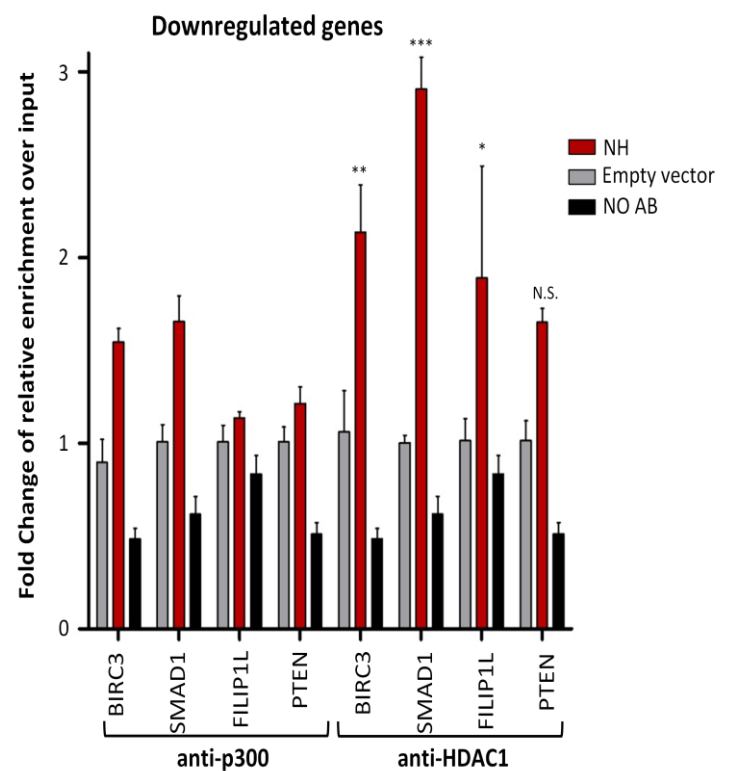
Figure 21: Effect of NUP98-HOXA9 on its other target genes. (A) GSEA plots using the NHChIP-seq gene set and the expression array data from hHP-NH cellular. (B) Analysis of NH and p300/HDAC1 interactions by co-immunoprecipitation. HEK293FT cells were transfected with pMSCV-Nup98-Hoxa9 or pMSCV-empty vectors. 48h post-transfection, the immunoprecitpitation was performed by using anti-p300 and anti-HDAC1 antibodies. Proteins were analyze by immunoblotting using anti-FLAG antibody. GAPDH protein levels were used as load control. (C) qChIP fold enrichment of p300 and HDAC1 in the regulatory regions of four upregulated and four downregulated target genes of NH in 293FT cellular models. The average of 3 experiments is shown. Error bars represent SEM.

5 In vitro analysis of the efficacy of the HDAC inhibitors on hHP- NUP98-HOXA9 cellular model

The interaction of NH with HDAC1 and its presence in the promoter of down-regulated genes prompted us to evaluate the effectiveness of a treatment with HDACs inhibitors (*HDACi*). *HDACi* are giving very promising results in the treatment of different hematologic malignancies, but to date they have not been studied for patients with the t(7;11)(p15,p15) translocation. Thus, we analyzed the sensitivity of our cellular model to the *pan-HDACi* Panobinostat (LBH589) compared with the sensitivity of other AML fusion genes for which the efficacy of this drug has already been demonstrated (*MLL-AF9* and *AML1-ETO*). We observed a dramatic inhibitory effect that, unexpectedly, was much higher (LC50$_{hHP-NH}$≈4nM) than that the one observed in the other leukemic models of hHP, hHP-MLL-AF9 (LC50$_{hHP-MLL\_AF9}$≈30nM) and hHP-AML1-ETO (LC50$_{hHP-AML1\_ETO}$≈200nM) fusion genes (Figure 22A). Indeed, the treatment with low doses (4nM) of LBH589 completely abrogated the hHP-NH cells capability to form colonies in the CFU assay (Figure 22B). Further, hHP-NH cells treated with 4nM and 30nM doses of LBH589 underwent apoptosis within 24h, whereas the same concentration of drug had no effect in the hHP-empty vector model (Figure 22C).

Summarizing, these experiments demonstrated that (1) the chimeric protein is capable of inducing a deep change in the expression profile (both activating and repressing target genes and (2) the repressing activity carried with collaboration of HDAC proteins clearly make those cells that harbor the fusion gene extremely sensitive to the treatment with epigenetic inhibitors.

A
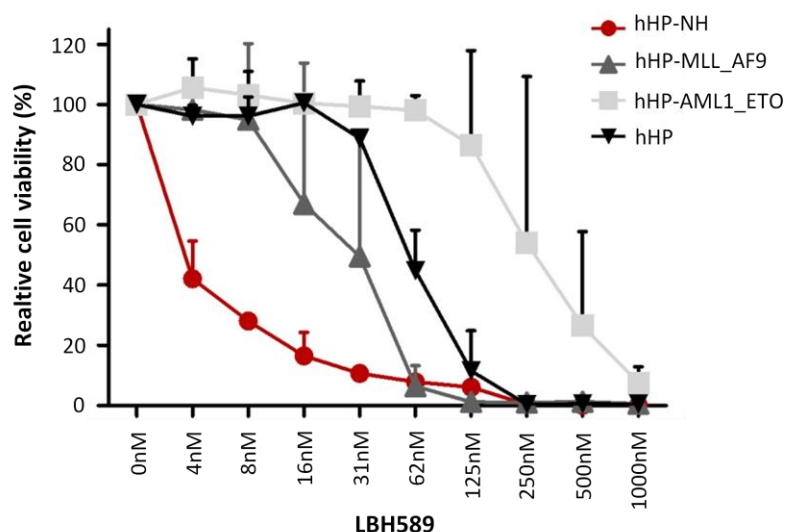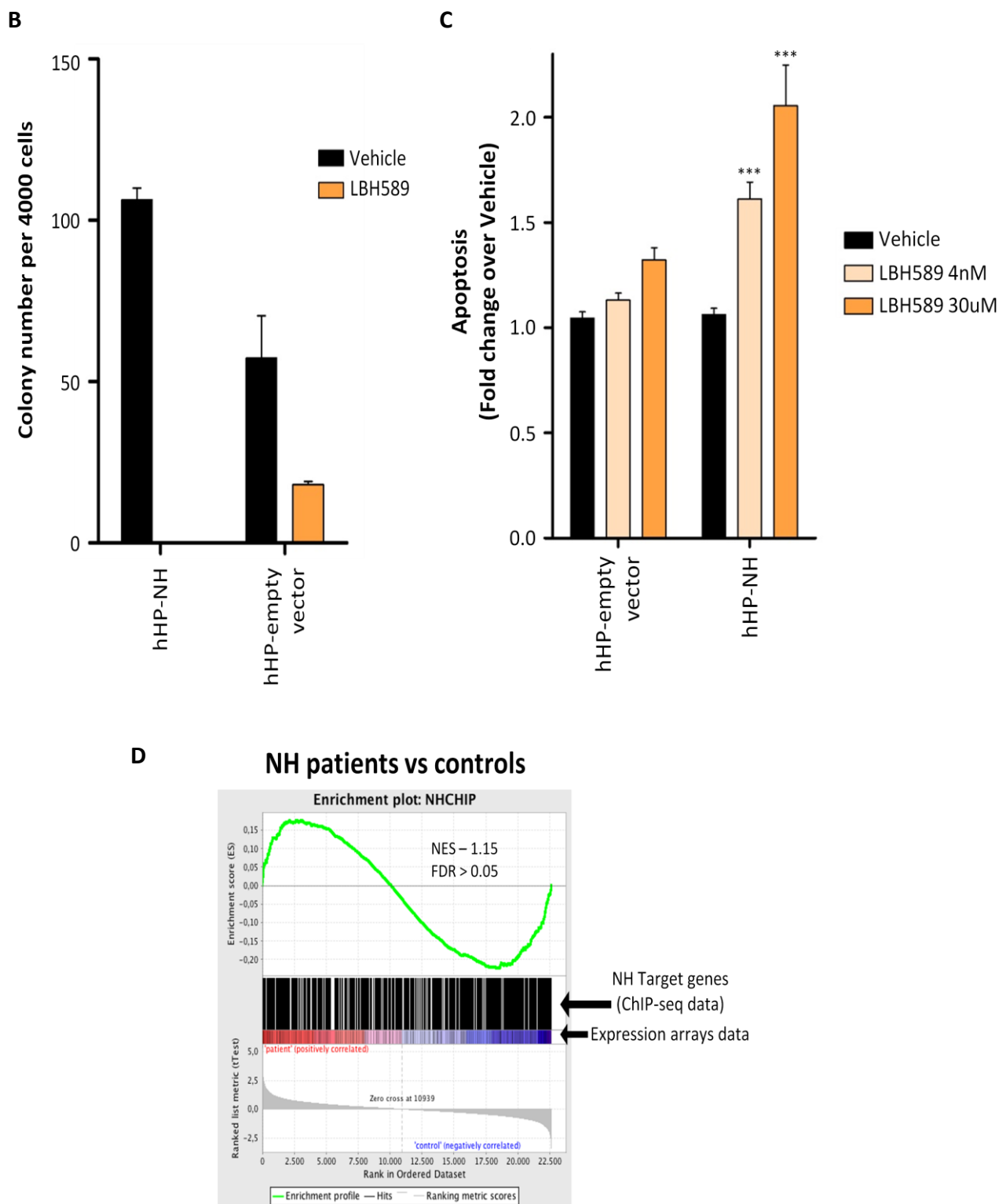
Figure 22: Treatment with HDAC inhibitors (A) hHP-NH cells were exposed for 72 hours to serial dilutions of panobinostat (LBH589) followed by addition of WST-1 to analyze cell viability. Averaged normalized optical density (OD) values are shown compared to vehicle. (B) Colony-forming assay with hHP-NH cells and hHP-emtpy vector cells treated with 4nM LBH589. Average number of colonies per dish from 2 independent experiments after 10 days is shown (only colonies with > 50 cells/colony were counted). Error bars represent SEM. (C) hHP-NH and hHP-empty vector cell apoptosis assay after the treatment with 4 nM and 30nM of HXR9 peptide for 24 hours. Cells were stained with annexin V and analyzed using flow cytometry. The plot shows the averaged fold changes of 3 independent experiments compared to vehicle. Error bars represent SEM. (D) GSEA plots using the NH ChIP-seq gene set and the expression array data from NH primary samples.

6   Patients with t(7;11)(p15;p15)

We obtained RNA samples from a further five patients with t(7;11)(p15,p15). To explore the activator-repressor role of NH in primary samples, we performed expression arrays with these samples and analyzed the data using GSEA (Table S6). We used three independent clones of hHP wt as control. Importantly, we found that the identify target genes behave similarly in primary samples and in the hHP-NH cellular model (Figure 22D). We showed both overexpressed and down-regulated target genes, thereby validating our finding that the fusion gene plays a dual role in transcriptional regulation.

Moreover, we used single-sample gene set enrichment analysis (ssGSEA) to assess gene set activation scores for the five tumor samples. ssGSEA is a rank-based comparison of the expression levels of genes in the gene set with all other genes in an expression profile[93]. In this analysis we included different gene sets and analysed the five expression profiles separately. The obtained pattern reflects a high level of homogeneity between samples and controls, but we observed two distinct clusters, meaning that the expression profiles of NH primary samples cluster together. Interestingly, one of the differentially expressed gene sets between patients and controls was the NOTCH pathway genes.

7   NUP98-HOXA9 binding-proteins identification by Mass Spectrometry

NH-interacting proteins have been tried to be identified by traditional yeast two-hybrid assays and it has allowed to find some proteins, like the Amino-terminal Enhancer of Split (AES) which collaborates with NH in the in deregulating transcription and proliferation[60]. However, Mass spectrometry is rapidly maturing as a powerful tool for proteomic analysis and specifically for the identification of protein-protein interaction. Therefore, in this study, we sought to identify and compare proteins that interact with NH, HOXA9 wt and NUP98 wt, using a Mass spectrometry assay.

We used the 293FT-NH, 293FT-HOXA9 and 293FT-NUP98 models to conduct the corresponding immunopreticiptations, followed by mass spectrometry. For each model, all peptides that appeared at least once in the controls were excluded and we selected only those

that appeared in all the three biological replicates analyzed. Table S7 (A-F) contain the lists of the proteins identified during these analyses, that interact with NH, HOXA9 or NUP98, respectively.

As it is observed in Figure 23, we found a total of 88 NH interacting-proteins, 198 for HOXA9 and 52 for NUP98. Interestingly, there are 38 proteins that only bind to NH, suggesting that the fusion protein acquires a new spatial conformation that allows it to interact with some proteins that are not able to bind to neither HOXA9 wt nor NUP98 wt (Figure 24A , Table S7A). Among them, we found proteins involved in chromatin remodeling (SMARCD2 or ARID1A), cytoskeletal reorganization (ABI1, ABI2, NCKAP1 or CYFIP1) or leukemogenesis (PBX2). However, the most striking finding was that NH binds to almost all of the components of the Ccr4-Not complex: CNOT1, CNOT3, CNOT6L, CNOT7, CNOT8, CNOT10 and RQCD1.
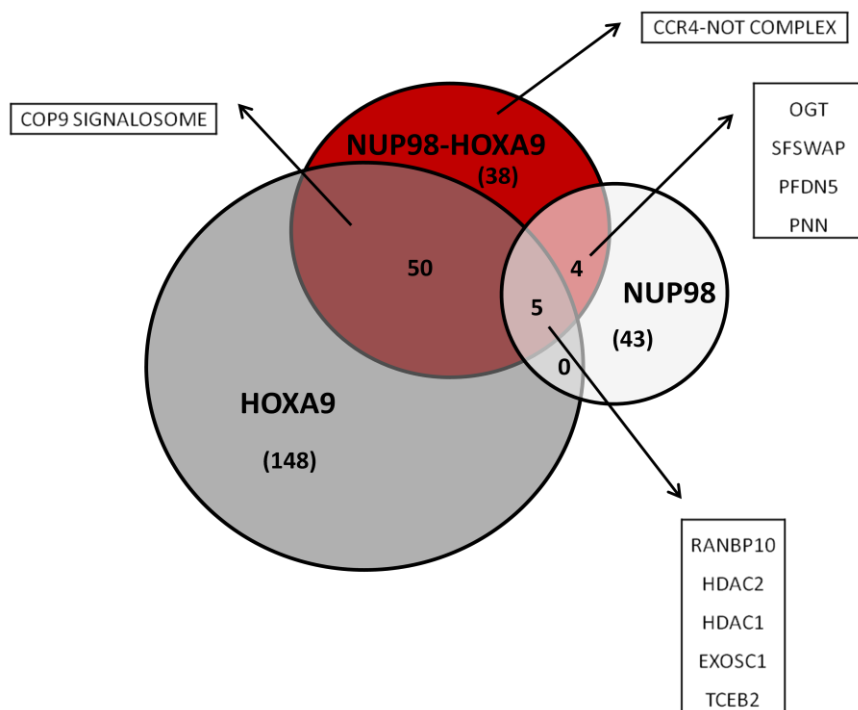
Figure 23: Mass spectrometry results. Venn diagrams of the overlapping NH, HOXA9 and NUP98 interacting -proteins indentified with the Mass Spectrometry approach on HEK293F human models

Among the binding proteins that are common to NH and HOXA9 (Table S7D), we found proteins responsible of the DNA repair (RPA3, RPA2, RAD18 or SMC3) and cell cycle regulators (CDKN2A or SP3). However, again, we made a surprising discovery: both HOXA9 and NH were able to interact to all of the members of the *COP9 signalosome* (GPS1, COPS2, COPS4, COPS5, COPS6, COPS7A, COPS7B and COPS8), that is involved in the ubiquitin–proteasome pathway (Figure 24B, Table S7D). We found four proteins that interact with both NH and NUP98 wt (Figure 24C, Table S7E). Notably, two of these proteins are RNA splicing regulators (SFRS8 and

PNN[94]), which supports the importance that aberrant splicing is acquiring in the pathogenesis of AML[95]. On the other hand, OGT is a glycosyltransferase that forms a core complex with HCF-1 and BAP1 that can differentially recruit histone-modifying enzymes to regulate gene expression and thereby preserve normal hematopoiesis[96]. The latter protein of this group, but no less interesting, would be PFDN5 or MM1. It is a subunit of prefoldin, a molecular chaperone complex that binds and stabilizes newly synthesized polypeptides.

Finally, among the five interacting proteins that are common to the three entities, the presence of HDAC1 and HDAC2 is remarkable since it validates the use of HDAC inhibitors for the treatment of patients harboring the NUP98-HOXA9 fusion protein (Figure 24D, Table S7F).

A

Only NUP98-HOXA9 interacting proteins
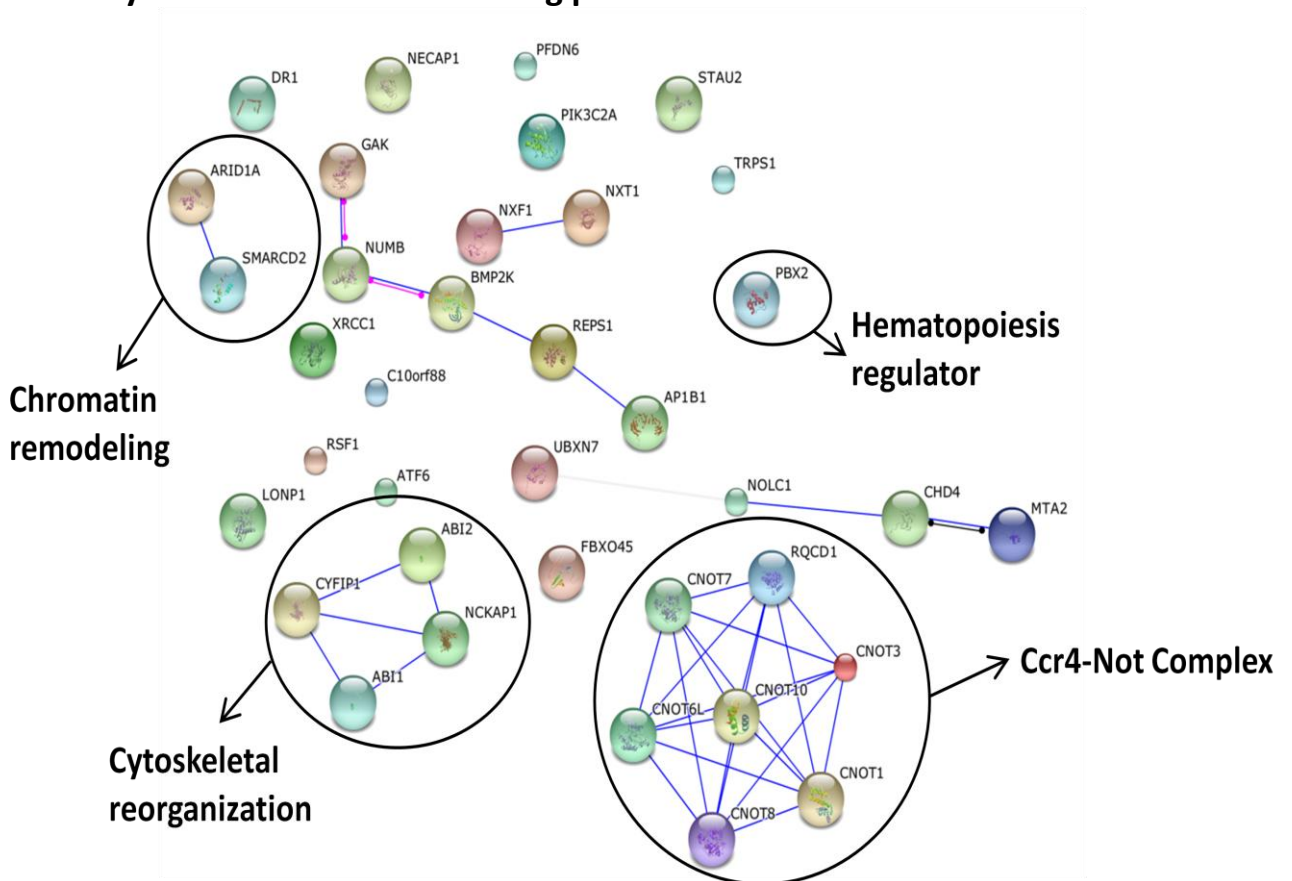
B   **Common NUP98-HOXA9 and HOXA9 wt interacting proteins**
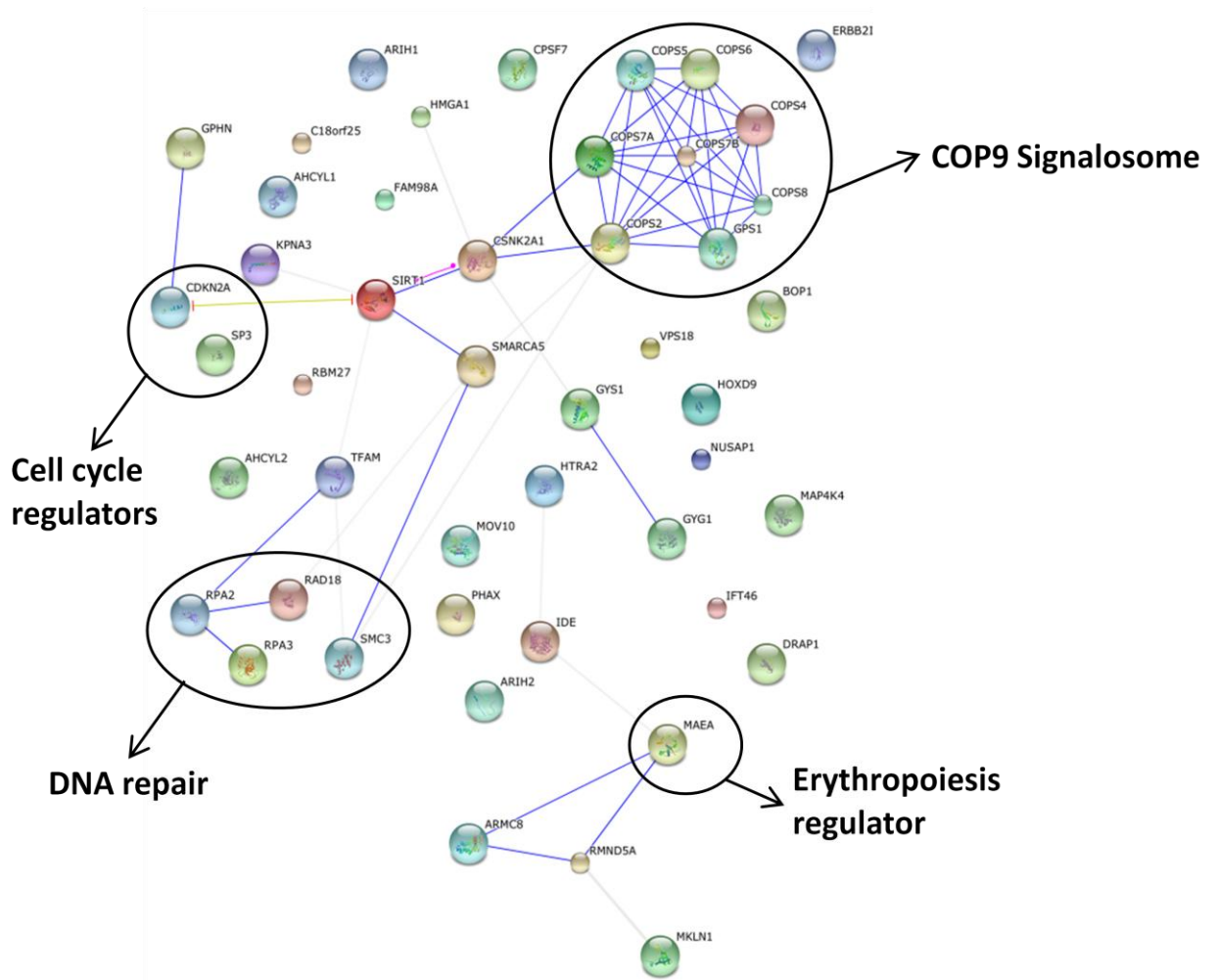
C   Common NUP98-HOXA9 and NUP98 wt interacting proteins
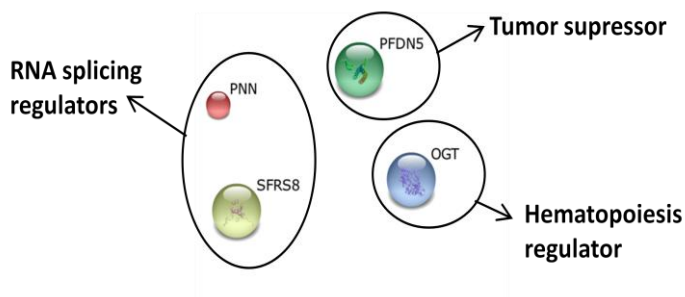
D   Common NUP98-HOXA9, HOXA9 wt and NUP98 wt interacting proteins
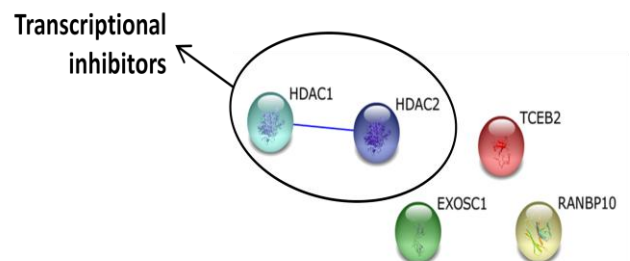
Figure 24: Identification of interacting-proteins. (A-D) STRING 9.1 diagrams of the interacting-proteins indentified with the Mass Spectrometry.

DISCUSSION

NUP98-HOXA9-associated AML is one of the most aggressive forms of leukemia. It is highly refractory to intensive treatment and has a dismal prognosis, only 2-5% of patients surviving in five-years. The molecular mechanisms underlying the occurrence of this fusion protein are poorly understood. Advancing knowledge in this area could trigger the identification of new and more efficient therapeutic targets. The present work provides proof of concept to explore a better designed targeted therapy for these patients.

NUP98-HOXA9 genomic binding sites

By identifying its DNA binding sites, we have shown that the transcription factor activity of NH is carried out mainly through enhancers. Each of the eight selected region was characterized as enhancer by detecting the enrichment of the H3K4me1 epigenetic mark. We also observed an increase in the presence of PolII in seven of the eight target sites when NH is present, indicating that the fusion protein could induce the activation of its target enhancers somehow. Moreover, through two different *in silico* analysis of the binding sequences, *MEME-ChIP* and *oOPOSSUM*, we were able to identify DNA binding motifs of several enhancer pioneer transcription factors (FOX proteins and GATA1) in the same target regions of NH. Transcriptional regulation carried out by HOXA9 wt is known to be mostly mediated by its binding to enhancers. Our results, therefore, prompted us to suggest that the NH binding to DNA is most likely driven by the HOXA9 moiety. Indeed, we found that one third of the NH target genes were common to HOXA9 wt.

Interestingly, a detailed study of the DNA binding regions allowed us to identify a consensus sequence (CA/gTTT) that is present in one-third of genomic regions where NH is bound, but is not associated with any known transcription factor. This specificity represents a unique and promising finding to be further explored.

Unexpected role of NUP98 in hematopoiesis

Notably, this effect depends not only on the presence of the HOXA9 homeodomain in the fusion, but also on the contribution of the NUP98 moiety to the chimeric protein. We have shown that this moiety also participates in the interaction of the fusion protein with DNA in a human cellular context. An unexpected finding was that not only the nucleoporin NUP98 showed direct DNA binding activity, but also that many of the genes involved in the interaction play a role in the regulation and differentiation of hematopoiesis. Recent studies have suggested that NUP98 acts as a transcription factor in a tissue and status dependent way[35]. We have demonstrated that its overexpression in hHP induces the upregulation of important genes such

as *HOXA9, PBX3, HOXA7, RARA* and *PML*, but also the downregulation of the tumor suppressor gene *BIRC3*. This role of NUP98 as a hematopoietic transcription factor needs to be further explored as it could open the door to important insights into hematopoiesis and leukemia. This finding indirectly provides an explanation to why a nucleoporin, mainly though as a structural protein, is so commonly rearranged in AML; it has more than 20 different partners that result in chimeric proteins with leukemogenic features. This proposed role of NUP98 led us to speculate that the loss of one wt allele of the gene, as a consequence of the chromosomal translocation, could have some effect on the blockage of hematopoietic differentiation, a common feature of myeloid leukemias. Supporting this hypothesis, it has been reported that endogenous NUP98 interacts with NH in the leukemic cells[51], an alternative mechanism that would block the transcription factor activity of the nucleoporin. Further, that haploinsuficiency of the NUP98 wt was required to induce full leukemic transformation could explain why a complete leukemic transformation has not been demonstrated in a retroviral model overexpressing NUP98-HOXA9 in hHP, where the two wt alleles are intact [54]. As we have recently published for other fusion genes[97], genetically engineered cell models that generate the chromosomal translocation inside the cell should provide a more appropriate tool to evaluate this hypothesis.

MESI1-HOXA9-PBX3 axis deregulation

It is interesting to note some clinical similarities found between *NUP98-HOXA9* and *MLL* gene rearrangements that are also events associated with aggressive myeloid leukemias. In both cases, it has been observed that there is an induction of erythroid hyperplasia, an abnormal erythroid maturation and a similar *in vitro* mode of transformation[27]. These findings suggested that both chromosomal rearrangements might be sharing some oncogenic mechanisms. Accordingly, we found that NH directly induces the overexpression of the three best known target genes of MLL-AF9 and MLL-AF4[47,98,99]fusion proteins: *MEIS1, HOXA9* and *PBX3.* These three transcription factors can form a transcriptional activator heterocomplex that regulates genes involved in the development of AML[47,48]. Additionally, it has been recently shown that the overexpression of a 4-homeobox-gene signature (composed of *PBX3, HOXA9, HOXA7* and *HOXA11*) is an independent predictor of poor overall survival in patients with cytogenetically abnormal AML (CA-AML)[89].

We demonstrated that the interaction of NH with the specific enhancers of *MEIS1, HOXA9* and *PBX3* directly induces their overexpression. In addition, we validated the formation of the actively functional heterocomplex MEIS1-HOXA9-PBX3 in cells expressing the fusion protein,

since it is able to induce its specific expression signature. Interestingly, we showed that NH could interact with this complex through PBX3. However, it seems that PBX3 could be collaborating with NH in transcriptional regulation of only a few genes. Therefore, these results suggest that recruitment of the complex by NH could be a mechanistic artifact that is not relevant for the NH function as transcriptional regulator. This matches previous studies that suggested that Pbx-interaction motif is not required for the oncogenic mechanism induced by NH[26].

Since Morgan *et al.*[100] designed the peptide HXR9, which is able to disrupt HOXA9-PBX interactions and shows effective anti-tumor effects in melanoma models and other solid tumors[91,92], recent results suggest that HXR9 could improve the survival of the leukemia patients with increased expression of the HOXA/PBX3 signature, as is the case for MLL-rearranged leukemia[47]. The role of PBX proteins, HOXA9 and MEIS1 in the context of NUP98-HOXA9 leukemia has been previously discussed[26], but with these new findings, a revision of their function in this context was required. Thus, we evaluated the efficacy of HXR9 treatment in our hHP-NH model and observed a clear inhibitory effect on cell viability. This effect is not as dramatic as in other leukemic fusions studied previously[47], indicating that the deregulation of the PBX3-HOXA9-MEIS1 axis is probably not the only driver determining the transformative power of NH.

Target miRNAs of NUP98-HOXA9

Among the 17 target miRNAs of NH that we identified in the ChIP-seq analysis (Table 3), we found three *miR181* family members:*hsa-miR-181a1, hsa-miR-181a2* and *hsa-miR-181b*. It has been shown that the miR181 family acts as a regulator of myeloid differentiation[101] and that its downregulation contributes to the poor prognosis of cytogenetically abnormal AML patients[89]. Interestingly, *PBX3, HOXA9, HOXA7* and *HOXA11* are potential targets of the *miR-181* family (particularly, *miR-181b*). In our hHP-NH cellular model we also observed the significant downregulation (Figure 9) of *miR-181b* which clearly entails an additional mechanism of the fusion protein to obtain an even greater overexpression HOX/PBX3 signature. Another miRNA that could be regulated by NH is *hsa-miR-128,* which targets the expression of genes NF1[102], MLL and AF4[103], all of them with an important role in human leukemia. *NF1* is a tumor suppressor gene whose downregulation is a common event in AML[104]. On the other hand, MLL is frequently involved in chromosomal translocations in aggressive human lymphoid and myeloid leukemias. One of these MLL-associated leukemias results from a balanced translocation between *MLL* and *AF4*. *hsa-miR-128* is able to downregulate the expression of *MLL, AF4*, and both *MLL-AF4* and

AF4-MLL fusion genes[103]. The data suggests that the deregulation of this miRNA by NH could serve an important function in the development of the leukemia. Finally, hsa-miR-30A has been proposed as a tumor suppressor gene in CML that is also deregulated by BCR-ABL[105].

Other NUP98-HOXA9 target genes

Although this work has focused specifically on three of the target genes of NH for their key role in AML (*MEIS1*, *HOXA9* and *PBX3*), the rest must not be ignored. Here, we summarize the most relevant features of the selected target genes and their relation to AML and other hematological cancers:

- *MET* is an oncogene that acts as receptor tyrosine kinase. A recurrent activating mutation in *MET that is* also described in small cell lung cancer and breast cancer, has been found in AML[106]. *MET* signaling is a requisite pathway in the growth and survival of AML cells in nearly half of the primary clinical samples[107].

- *BRAF* is a serine/threonine protein kinase involved in cell division, differentiation, and secretion. Mutations in this gene have also been associated with various types of cancer. In AML, *BRAF* mutations, as well as other mutations of genes more downstream in the *RAS-BRAF* pathway, are very frequent[108].

- *AF9* is a member of the YEATS domain superfamily with functions in transcriptional elongation. *AF9* is one of the most common translocation partners of MLL fusion proteins that are responsible for a subset of acute leukemias. It has been proposed as an attractive pharmacological target in such diseases[109].

- *NF1* is a tumor suppressor gene that negatively regulates the RAS signal transduction pathway and, as discussed above, it appears mutated or deleted very frequently in AML[110].

- *PTEN* is another tumor suppressor gene that is inactivated in many human cancers, including AML and other hematological malignancies[111]. *PTEN* is a negative regulator of AKT/PKB signaling pathway.

- *AFF3 is a* nuclear transcription factor that may function in lymphoid development and oncogenesis. It has been found in fusion with MLL in acute lymphoblastic leukemia patients[112].

- *BIRC3* negatively regulates MAP3K14, an activator of the non-canonical pathway of NF-κB signalling. This gene is frequently mutated and deleted in Chronic Lymphocytic Leukemia (CLL). CLLs harbouring *BIRC3* disruption display constitutive NF-κB activation that is associates with unfavorable clinical and genetic features and predicts poor outcomes independent of other risk factors[113].

- *SMAD1* belongs to the SMAD family and it is a signal transducer and a transcriptional modulator that mediate multiple signalling pathways involved in a range of biological activities including cell growth, apoptosis, morphogenesis, development and immune responses. Smad-signalling circuitry is intimately linked to normal and leukemic hematopoiesis[114].

- *FILIP1L* is a novel tumor suppressor that inhibits the canonical WNT signalling and has its expression downregulated in various cancers. Its great potential as a cancer therapeutic target has been proposed.[115]

Given the important functions of these genes in tumorigenesis and leukemia development, the oncogenic mechanism induced by NH is likely the result of the collaboration of all of them. In this regard, we decided to assess the effect of NH on the rest of its target genes, both in our model and, more relevantly, in primary samples. We observed that the fusion protein seemed to have the capacity to induce both upregulation and downregulation. This result confirms that the preponderance of overexpressed genes described the hHP-NH signatures[27,54,55] is most probably due to a downstream effect (for example, gene activation induced by PBX3-HOXA9-MEIS1), rather than the direct action of NH.

In order to better understand this activator-repressor role that NH seems to have on its target genes, we explored the capacity of the Nup98 moiety of the fusion protein to recruit transcriptional regulators, such as the coactivator p300 or the corepressor HDAC1[51]. We demonstrated, for the first time, this dual role, showing the presence, along with the fusion protein, of p300 and HDAC1 in the regulatory regions of the overexpressed and downregulated NH-target genes, respectively.

Treatment with HDAC inhibitor

HDAC inhibitors are a novel class of anticancer agents that block the activity of HDACs with promising therapeutic expectations in different types of myeloid malignancies. Since in our model we identified the involvement a molecular mechanism that justifies the use of these drugs in patients harboring the NH fusion protein, we assayed in vitro the potential use of LBH589 (Panobinostat). We observed a high inhibitory effect in hHP-NH after treatment with that was more efficient in our model than in others myleoid cell lines bearing fusion genes such as *MLL-AF9* or *AML1-ETO*.

Likewise, we suggested that a p300 inhibitor like C646, that induces cell cycle arrest and apoptosis selectively in *AML-ETO* positive AML cells[116], could also be tested in the cellular models and in patients that harbor NH.

However, in patients with t(7;11)(p15;p15), we found a stronger presence of downregulated target genes, indicating that the treatment with HDAC1 could be even more effective. Patients with the translocation should be included in Panobinostat clinical trials.

Probably the best therapeutic option for these patients would be the use of combination therapies that include the proposals made in this study.

New NUP98-HOXA9 interacting-proteins

The Mass spectrometry results confirm that NH is not only an aberrant transcription factor but that its ability to interact with many other proteins allows it to expand its leukemogenic potential to achieve the high degree of malignancy that characterizes it. We found that NH binds to almost all the components of the CCR4-NOT complex and that this capacity could be acquired from the specific spatial conformation of the fusion protein, since none of the proteins of the complex seem to interact neither with NUP98 wt norHOXA9 wt. The CCR4-NOT complex is a unique, essential and conserved multi-subunit complex[117]. It has been studied mostly in *Saccharomyces cerevisiae* but the orthologous genes have been found in humans. The complex is composed by two functional modules with different subunits that comprise two different enzymatic activities: deadenylation and ubiquitination. The combination of these two activities allows coordinating multiple functions such as mRNA degradation, protein synthesis and degradation, RNA nuclear export and chromatin modification. In this broad range of processes, it has a direct involvement in transcriptional regulation or DNA damage response, making it essential for cell viability [118]. Given these important functions, the newly identified ability of NH to interact with the CCR4-NOT complex and its role in leukemogenesis open up an exciting research field.

Another striking finding came up when we identified, among the binding proteins that are common to NH and HOXA9, all of the COP9 *signalosome* components. It is a conserved protein complex that operates in the ubiquitin–proteasome pathway. Although its mechanism of action is not yet fully understood, it is rapidly emerging as a key player in the DNA-damage response, cell-cycle control and gene expression[119]. Further studies are necessary to elucidate the consequences of the interaction between the fusion protein and the intriguing and multifunctional COP9 *signalosome* in leukemic transformation. Importantly, NH and Hoxa9 also bind to MAEA, a key protein in normal differentiation of erythroid cells and macrophages[120]. As NH induces erythroid hyperplasia and abnormal erythroid maturation[27], it is very likely that the interaction with MAEA plays a role in this process.

In addition, we found that NH and NUP98 wt interact with OGT, a protein whose usual role in hematopoiesis could be altered because of this interaction, helping in the induction of the impaired blood differentiation that is characteristic of NH-driven AML. Regarding the interaction between NH and MM1, it is noteworthy that mutations in MM1 are often observed in patients with leukemia or lymphoma and it has been considered as a tumor suppressor that acts by repressing the transcriptional activity of the proto-oncogene *c-Myc*[87]. The interaction with the fusion protein could free c-myc, allowing it to carry out its oncogenic activity.

Finally, mass spectrometry results let us find out that both HDAC1 and HDAC2 are able to interact with NH, HOXA9 and NUP98. It has already been proposed that the retained FG-repeat domain of Nup98 in the fusion protein could be mediating the interaction with HDAC1[37]. However, the possibility that HOXA9 wt also binds to HDACs had not ever been considered. Furthermore, the fact that the fusion protein is also able to recruit HDAC2 supports the effectiveness of the treatment with Panobinostat, the tested pan-HDACs inhibitor, in this kind of leukemia.

We have been highly restrictive in the analysis of the MS results, which explains why we have not been able to detect some NH interacting-proteins already identified in previous works. However, this approach made it possible to find a large number of novel collaborators of the fusion protein that would contribute to the discovery of new molecular mechanisms underlying this oncogene. Features and architecture of these interactions need to be further explored.

Final remarks

Our work describes most of the molecular mechanisms that underlay the leukemogenic activity of the NH fusion protein. A surprising new function for the nucleoporin NUP98 in hematopoiesis has been demonstrated. We have described how NH has the ability to interact with HDAC1 and p300 to inhibit and induce gene expression respectively.

With the evidences on the activation of the PBX3-MEIS1-HOXA9 axis, and the supporting data coming from primary samples, we have provided new biological rationales for widening the therapeutic reservoir for the patients.

It is important to note that since it is now clear that the NUP98 moiety plays a crucial role in leukemogenesis, other NUP98 fusions could share the same mechanism of action. The proposed therapeutic approaches could represent a breakthrough in the treatment of a larger number of cases of acute myeloid leukemia.

Finally, our description of a huge amount of novel NH interacting-proteins supports the idea that NH is not only an aberrant transcription factor, but that its ability to interact with many other proteins allows it to expand its leukemogenic potential to achieve the high degree of malignancy that characterizes it.

CONCLUSIONES/CONCLUSIONS

CONCLUSIONES

1. Hemos generado con éxito dos modelos celulares humanos que expresan NUP98-HOXA9 de forma constitutiva. Uno de ellos se ha generado a partir de células embrionarias epiteliales y otro a partir de progenitores hematopoyéticos humanos. Éste último fue capaz de reproducir la proliferación aberrante inducida por la proteína de fusión.

2. Nuestro trabajo supone la primera descripción de los sitios de unión al DNA específicos de NUP98-HOXA, la mayoría de los cuales son regiones reguladoras "potenciadoras". Hemos identificado un nuevo motivo de unión específico para NUP98-HOXA9 (CA/gTTT) que está presente en un tercio de sus sitios diana. NUP98-HOXA9 se une a las regiones reguladoras de 1363 genes y 17 microARNs que están principalmente implicados en procesos de muerte celular, supervivencia, ciclo celular y desarrollo de la leucemia.

3. NUP98-HOXA9 activa directamente la expresión de los genes que forman el complejo MEIS1-HOXA9-PBX3 mediante la interacción con sus regiones reguladoras "potenciadoras", constituyendo, esta activación, uno de los mecanismos patogénicos inducidos por la proteína de fusión. La desestabilización de este complejo usando el péptido HXR9 en progenitores hematopoyéticos humanos que expresan NUP98-HOXA9 tiene un efecto directo y específico en la viabilidad celular. Este resultado nos lleva a proponer al eje MEIS1-HOXA9-PBX3 como una adecuada y prometedora diana terapéutica que debe ser evaluada en pacientes con la transclocación t(7;11)(p15;p15).

4. Al analizar la función cada una de las dos partes que forman la proteína de fusión, encontramos que un tercio de los genes diana de NUP98-HOXA9 son comunes a HOXA9 wt, lo que confirma la importancia del homeodominio en la capacidad de interacción con el DNA. Por otro lado, proponemos una papel para NUP98 wt en la regulación hematopoyética, ya que identificamos 17 regiones de unión al DNA y todas ellas parecen tener una función en hematopoyesis o leucemogénesis. Además, la sobreexpresión de NUP98 wt en progenitores hematopoyéticos humanos activa la expresión de genes diana implicados directamente en el desarrollo de la LMA e induce proliferación celular. Estos resultados sugieren, por tanto, un papel oncogénico para NUP98 en neoplasias hematológicas.

5. El análisis de la expresion génica en modelos celulares humanos y muestras primarias de pacientes sugieren que NUP98-HOXA9 podría tener un papel activador-represor en la regulación transcripcional que estaría determinado por su interacción con p300 (activador transcripcional) y con HDAC1 (inhibidor transcripcional).

6. La función represora de la transcripción génica que parece tener NUP98-HOXA9 en colaboración con HDACs, hace que las células que expresan el gen de fusión sean extremadamente sensibles al tratamiento con el inhibidor de HDAC LBH589 (Panobinostat). Además, observamos que el efecto de LBH589 es mucho mayor en el modelo celular que expresa NUP98-HOXA9 que en otros modelos leucémicos para los que la sensibilidad a esta droga ya había sido demostrada. Por tanto, proponemos firmemente que los inhibidores de HDAC, solos o en combinación con otros agentes quimioterapéuticos, sean probados en pacientes con este tipo de leucemia.

7. El análisis mediante espectrometría de masas nos ha permitido identificar y comparar un gran número de nuevas proteínas que interaccionan con NUP98-HOXA9, HOXA9 and NUP98. Estos resultados confirman que la proteína de fusión no es solamente un factor de transcripción oncogénico, si no que su capacidad de interacción con otras proteínas le permite ampliar su potencial leucemogénico para conseguir el alto grado de malignidad que le caracteriza. Sin embargo, la arquitectura y características de estas interacciones necesitan ser estudiadas en mayor profundidad.

8. En conjunto, estos datos indican que el proceso oncogénico inducido por el gen de fusión NUP98-HOXA9 implica una enorme variedad de mecanismos moleculares: la regulación directa e indirecta de la expresión génica a través de regiones *potenciadoras*, modificaciones epigenéticas y una amplia red de interacciones proteicas. Este escenario inducido por la translocación cromosómica puede ser responsable de la compleja biología que presentan los progenitores leucémicos y probablemente explica la falta respuesta a los tratamientos actuales disponibles para esta enfermedad.

CONCLUSIONS

1. We successfully generated two human cellular models that expressed NUP98-HOXA9 in a constitutive manner. One was based on embryonic epithelial cells and the other on human hematopoietic progenitors. The latter was able to reproduce the aberrant proliferation induced by the fusion protein.

2. Our work provides the first description of the DNA binding sites of NUP98-HOXA9, most of which are enhancer regulatory regions. We identify a novel specific motif (CA/gTTT) for NH that is present in one-third of its target genomic regions. NH is binding to the regulatory region of 1363 genes and 17 miRNAs that are mainly involved in cell death, survival, cell cycle and leukemia development.

3. NUP98-HOXA9 directly activates the expression of the complex MEIS1-HOXA9-PBX3 by interacting with its respective enhancers and it constitutes one of the pathogenic mechanisms induced by the fusion protein. The disruption of this complex with the HXR9 peptide in human hematopoietic progenitors that express NUP98-HOXA9 has a direct effect on cell viability. This results led us to propose the MEIS1-HOXA9-PBX3 axis as a good therapeutic target to evaluate in patients with the t(7;11)(p15,p15).

4. Analyzing the role of each moiety of the fusion protein, we found that one-third of the NH target genes are common to HOXA9 wt, which confirms the importance of the homeobox domain in the ability to interact with the DNA. We also propose a function in hematological differentiation for the nucleoporin NUP98, since we identified 17 DNA binding sites and all of them appeared to play a direct role in hematopoiesis or leukemogenesis. In addition, its overexpression in hHP activates the expression of genes directly involved in AML development and induces cell proliferation. Therefore, these results also suggest an oncogenic role for NUP98 in hematological malignancies

5. Gene expression analysis in human cellular models and primary samples suggests an activator-repressor role to NUP98-HOXA9 in the transcriptional regulation that is determined by its interaction with both p300 (transcriptional activator) and HDAC1 (transcriptional inhibitor).

6. The repressing transcriptional activity carried by NH with collaboration of HDAC proteins clearly make those cells that harbor the fusion gene extremely sensitive to the treatment with the pan-HDAC inhibitor LBH589 (Panobinostat). We found that the inhibitory effect of LBS589 was much higher in the cellular model expressing NH than in other leukemic models for which the sensitivity to this drug has already been demonstrated. Thus, we strongly proposed that HDAC inhibitors, alone, or in combination with other chemotherapeutic agents, must be tested in NH driven AML patients.

7. Mass Spectrometry allowed us to identify and compare a huge number of novel binding-proteins of NH, HOXA9 and NUP98. These results confirmed that NH is not only an aberrant transcription factor, but that their ability to interact with many other proteins enable to expand its leukemogenic potential to achieve the high degree of malignancy that characterizes it. However, the features and architecture of these identified interactions need to be further explored.

8. Taking together, all data indicate that the leukemogenic process induced by fusion gene NUP98-HOXA9 involves several molecular mechanisms: direct and indirect gene regulation via enhancers, epigenetic modifications and a novel network of protein interaction. This profoundly aberrant landscape induced by the translocation may be responsible of the complex biology of the leukemic hematopoietic progenitor and probably explain the lack of response to the current treatments available for this disease.

BIBLIOGRAPHY

1.   Kumar CC. Genetic abnormalities and challenges in the treatment of acute myeloid leukemia. Genes Cancer. 2011;2:95-107.

2.   Estey E, Dohner H. Acute myeloid leukaemia. Lancet. 2006;368:1894-1907.

3.   Shimada H, Ichikawa H, Ohki M. Potential involvement of the AML1-MTG8 fusion protein in the granulocytic maturation characteristic of the t(8;21) acute myelogenous leukemia revealed by microarray analysis. Leukemia. 2002;16:874-885.

4.   Schlenk RF, Dohner H. Genomic applications in the clinic: use in treatment paradigm of acute myeloid leukemia. Hematology Am Soc Hematol Educ Program. 2013;2013:324-330.

5.   Kvinlaug BT, Chan WI, Bullinger L, et al. Common and overlapping oncogenic pathways contribute to the evolution of acute myeloid leukemias. Cancer Res. 2011;71:4117-4129.

6.   Higuchi M, O'Brien D, Kumaravelu P, Lenny N, Yeoh EJ, Downing JR. Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. Cancer Cell. 2002;1:63-74.

7.   Conway O'Brien E, Prideaux S, Chevassut T. The epigenetic landscape of acute myeloid leukemia. Adv Hematol. 2014;2014:103175.

8.   Gough SM, Slape CI, Aplan PD. NUP98 gene fusions and hematopoietic malignancies: common themes and new biologic insights. Blood. 2011;118:6247-6257.

9.   Murati A, Brecqueville M, Devillier R, Mozziconacci MJ, Gelsi-Boyer V, Birnbaum D. Myeloid malignancies: mutations, models and management. BMC Cancer. 2012;12:304.

10.   Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. Cell. 2012;150:264-278.

11.   Mrozek K, Marcucci G, Paschka P, Whitman SP, Bloomfield CD. Clinical relevance of mutations and gene-expression changes in adult acute myeloid leukemia with normal cytogenetics: are we ready for a prognostically prioritized molecular classification? Blood. 2007;109:431-448.

12.   Dohner H, Estey EH, Amadori S, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. Blood. 2010;115:453-474.

13.   Nambiar M, Kari V, Raghavan SC. Chromosomal translocations in cancer. Biochim Biophys Acta. 2008;1786:139-152.

14.   Chou WC, Chen CY, Hou HA, et al. Acute myeloid leukemia bearing t(7;11)(p15;p15) is a distinct cytogenetic entity with poor outcome and a distinct mutation profile: comparative analysis of 493 adult patients. Leukemia. 2009;23:1303-1310.

15.   Pommier Y, Leo E, Zhang H, Marchand C. DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. Chem Biol. 2010;17:421-433.

16.   Chhikara BS, Parang K. Development of cytarabine prodrugs and delivery systems for leukemia treatment. Expert Opin Drug Deliv. 2010;7:1399-1414.

17.   Felipe Rico J, Hassane DC, Guzman ML. Acute myelogenous leukemia stem cells: from Bench to Bedside. Cancer Lett. 2013;338:4-9.

18.   DeAngelo DJ, Spencer A, Bhalla KN, et al. Phase Ia/II, two-arm, open-label, dose-escalation study of oral panobinostat administered via two dosing schedules in patients with advanced hematologic malignancies. Leukemia. 2013;27:1628-1636.

19.   Bolden JE, Peart MJ, Johnstone RW. Anticancer activities of histone deacetylase inhibitors. Nat Rev Drug Discov. 2006;5:769-784.

20.   Fraga MF, Ballestar E, Villar-Garea A, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet. 2005;37:391-400.

21.   StatBite: FDA oncology drug product approvals in 2009. J Natl Cancer Inst. 2010;102:219.

22.   Mann BS, Johnson JR, Cohen MH, Justice R, Pazdur R. FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. Oncologist. 2007;12:1247-1252.

23. Sato Y, Abe S, Mise K, et al. Reciprocal translocation involving the short arms of chromosomes 7 and 11, t(7p-;11p+), associated with myeloid leukemia with maturation. Blood. 1987;70:1654-1658.

24. Nakamura T, Largaespada DA, Lee MP, et al. Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. Nat Genet. 1996;12:154-158.

25. Kroon E, Thorsteinsdottir U, Mayotte N, Nakamura T, Sauvageau G. NUP98-HOXA9 expression in hemopoietic stem cells induces chronic and acute myeloid leukemias in mice. EMBO J. 2001;20:350-361.

26. Calvo KR, Sykes DB, Pasillas MP, Kamps MP. Nup98-HoxA9 immortalizes myeloid progenitors, enforces expression of Hoxa9, Hoxa7 and Meis1, and alters cytokine-specific responses in a manner similar to that induced by retroviral co-expression of Hoxa9 and Meis1. Oncogene. 2002;21:4247-4256.

27. Abdul-Nabi AM, Yassin ER, Varghese N, Deshmukh H, Yaseen NR. In vitro transformation of primary human CD34+ cells by AML fusion oncogenes: early gene expression profiling reveals possible drug target in AML. PLoS One. 2010;5:e12464.

28. Ptak C, Aitchison JD, Wozniak RW. The multifunctional nuclear pore complex: a platform for controlling gene expression. Curr Opin Cell Biol. 2014;28C:46-53.

29. Capelson M, Hetzer MW. The role of nuclear pores in gene regulation, development and disease. EMBO Rep. 2009;10:697-705.

30. Kalverda B, Pickersgill H, Shloma VV, Fornerod M. Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm. Cell. 2010;140:360-371.

31. Franks TM, Hetzer MW. The role of Nup98 in transcription regulation in healthy and diseased cells. Trends Cell Biol. 2013;23:112-117.

32. Laurell E, Beck K, Krupina K, et al. Phosphorylation of Nup98 by multiple kinases is crucial for NPC disassembly during mitotic entry. Cell. 2011;144:539-550.

33. Takeda A, Yaseen NR. Nucleoporins and nucleocytoplasmic transport in hematologic malignancies. Semin Cancer Biol. 2014.

34. Capelson M, Liang Y, Schulte R, Mair W, Wagner U, Hetzer MW. Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes. Cell;140:372-383.

35. Liang Y, Franks TM, Marchetto MC, Gage FH, Hetzer MW. Dynamic association of NUP98 with the human genome. PLoS Genet. 2013;9:e1003308.

36. Sarma NJ, Yaseen NR. Dynein Light Chain 1 (DYNLT1) Interacts with Normal and Oncogenic Nucleoporins. PLoS One. 2013;8:e67032.

37. Moore MA, Chung KY, Plasilova M, et al. NUP98 dysregulation in myeloid leukemogenesis. Ann N Y Acad Sci. 2007;1106:114-142.

38. Collins C, Wang J, Miao H, et al. C/EBPalpha is an essential collaborator in Hoxa9/Meis1-mediated leukemogenesis. Proc Natl Acad Sci U S A. 2014.

39. Lindsey S, Huang W, Wang H, Horvath E, Zhu C, Eklund EA. Activation of SHP2 protein-tyrosine phosphatase increases HoxA10-induced repression of the genes encoding gp91(PHOX) and p67(PHOX). J Biol Chem. 2007;282:2237-2249.

40. Bei L, Lu Y, Eklund EA. HOXA9 activates transcription of the gene encoding gp91Phox during myeloid differentiation. J Biol Chem. 2005;280:12359-12370.

41. Alharbi RA, Pettengell R, Pandha HS, Morgan R. The role of HOX genes in normal hematopoiesis and acute leukemia. Leukemia. 2013;27:1000-1008.

42. Brumatti G, Salmanidis M, Kok CH, et al. HoxA9 regulated Bcl-2 expression mediates survival of myeloid progenitors and the severity of HoxA9-dependent leukemia. Oncotarget. 2013;4:1933-1947.

43. Eklund E. The role of Hox proteins in leukemogenesis: insights into key regulatory events in hematopoiesis. Crit Rev Oncog. 2011;16:65-76.

44.   Thorsteinsdottir U, Mamo A, Kroon E, et al. Overexpression of the myeloid leukemia-associated Hoxa9 gene in bone marrow cells induces stem cell expansion. Blood. 2002;99:121-129.

45.   Magnusson M, Brun AC, Lawrence HJ, Karlsson S. Hoxa9/hoxb3/hoxb4 compound null mice display severe hematopoietic defects. Exp Hematol. 2007;35:1421-1428.

46.   Huang Y, Sitwala K, Bronstein J, et al. Identification and characterization of Hoxa9 binding sites in hematopoietic cells. Blood. 2012;119:388-398.

47.   Li Z, Zhang Z, Li Y, et al. PBX3 is an important cofactor of HOXA9 in leukemogenesis. Blood. 2013;121:1422-1431.

48.   Shen WF, Rozenfeld S, Kwong A, Kom ves LG, Lawrence HJ, Largman C. HOXA9 forms triple complexes with PBX2 and MEIS1 in myeloid cells. Mol Cell Biol. 1999;19:3051-3061.

49.   Mullighan CG, Kennedy A, Zhou X, et al. Pediatric acute myeloid leukemia with NPM1 mutations is characterized by a gene expression profile with dysregulated HOX gene expression distinct from MLL-rearranged leukemias. Leukemia. 2007;21:2000-2009.

50.   Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science. 1999;286:531-537.

51.   Xu S, Powers MA. Nup98-homeodomain fusions interact with endogenous Nup98 during interphase and localize to kinetochores and chromosome arms during mitosis. Mol Biol Cell. 2010;21:1585-1596.

52.   Yassin ER, Sarma NJ, Abdul-Nabi AM, et al. Dissection of the transformation of primary human hematopoietic cells by the oncogene NUP98-HOXA9. PLoS One. 2009;4:e6719.

53.   Takeda A, Goolsby C, Yaseen NR. NUP98-HOXA9 induces long-term proliferation and blocks differentiation of primary human CD34+ hematopoietic cells. Cancer Res. 2006;66:6628-6637.

54.   Chung KY, Morrone G, Schuringa JJ, et al. Enforced expression of NUP98-HOXA9 in human CD34(+) cells enhances stem cell proliferation. Cancer Res. 2006;66:11781-11791.

55.   Ghannam G, Takeda A, Camarata T, Moore MA, Viale A, Yaseen NR. The oncogene Nup98-HOXA9 induces gene transcription in myeloid cells. J Biol Chem. 2004;279:866-875.

56.   Abdul-Nabi AM, Yassin ER, Varghese N, Deshmukh H, Yaseen NR. In vitro transformation of primary human CD34+ cells by AML fusion oncogenes: early gene expression profiling reveals possible drug target in AML. PLoS One;5:e12464.

57.   Mayotte N, Roy DC, Yao J, Kroon E, Sauvageau G. Oncogenic interaction between BCR-ABL and NUP98-HOXA9 demonstrated by the use of an in vitro purging culture system. Blood. 2002;100:4177-4184.

58.   Oka M, Asally M, Yasuda Y, Ogawa Y, Tachibana T, Yoneda Y. The mobile FG nucleoporin Nup98 is a cofactor for Crm1-dependent protein export. Mol Biol Cell. 2010;21:1885-1896.

59.   Takeda A, Sarma NJ, Abdul-Nabi AM, Yaseen NR. Inhibition of CRM1-mediated nuclear export of transcription factors by leukemogenic NUP98 fusion proteins. J Biol Chem. 2010;285:16248-16257.

60.   Sarma NJ, Yaseen NR. Amino-terminal enhancer of split (AES) interacts with the oncoprotein NUP98-HOXA9 and enhances its transforming ability. J Biol Chem. 2011;286:38989-39001.

61.   Dash AB, Williams IR, Kutok JL, et al. A murine model of CML blast crisis induced by cooperation between BCR/ABL and NUP98/HOXA9. Proc Natl Acad Sci U S A. 2002;99:7622-7627.

62.   Jankovic D, Gorello P, Liu T, et al. Leukemogenic mechanisms and targets of a NUP98/HHEX fusion in acute myeloid leukemia. Blood. 2008;111:5672-5682.

63.   Pineault N, Buske C, Feuring-Buske M, et al. Induction of acute myeloid leukemia in mice by the human leukemia-specific fusion gene NUP98-HOXD13 in concert with Meis1. Blood. 2003;101:4529-4538.

64.   Hirose K, Abramovich C, Argiropoulos B, Humphries RK. Leukemogenic properties of NUP98-PMX1 are linked to NUP98 and homeodomain sequence functions but not to binding properties of PMX1 to serum response factor. Oncogene. 2008;27:6056-6067.

65.   Gurevich RM, Aplan PD, Humphries RK. NUP98-topoisomerase I acute myeloid leukemia-associated fusion gene has potent leukemogenic activities independent of an engineered catalytic site mutation. Blood. 2004;104:1127-1136.

66.   Wang GG, Cai L, Pasillas MP, Kamps MP. NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. Nat Cell Biol. 2007;9:804-812.

67.   Smith E, Shilatifard A. Enhancer biology and enhanceropathies. Nat Struct Mol Biol. 2014;21:210-219.

68.   Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. Genes Dev. 2011;25:2227-2241.

69.   Kagey MH, Newman JJ, Bilodeau S, et al. Mediator and cohesin connect gene expression and chromatin architecture. Nature. 2010;467:430-435.

70.   De Santa F, Barozzi I, Mietton F, et al. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol. 2010;8:e1000384.

71.   Siersbaek R, Rabiee A, Nielsen R, et al. Transcription factor cooperativity in early adipogenic hotspots and super-enhancers. Cell Rep. 2014;7:1443-1455.

72.   Whyte WA, Orlando DA, Hnisz D, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013;153:307-319.

73.   Loven J, Hoke HA, Lin CY, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell. 2013;153:320-334.

74.   van Arensbergen J, van Steensel B, Bussemaker HJ. In search of the determinants of enhancer-promoter interaction specificity. Trends Cell Biol. 2014.

75.   Herz HM, Hu D, Shilatifard A. Enhancer malfunction in cancer. Mol Cell. 2014;53:859-866.

76.   Mulloy JC, Cammenga J, MacKenzie KL, Berguido FJ, Moore MA, Nimer SD. The AML1-ETO fusion protein promotes the expansion of human hematopoietic stem cells. Blood. 2002;99:15-23.

77.   Goyama S, Schibler J, Cunningham L, et al. Transcription factor RUNX1 promotes survival of acute myeloid leukemia cells. J Clin Invest. 2013;123:3876-3888.

78.   Ballestar E, Paz MF, Valle L, et al. Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. EMBO J. 2003;22:6335-6345.

79.   Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754-1760.

80.   Feng J, Liu T, Zhang Y. Using MACS to identify peaks from ChIP-Seq data. Curr Protoc Bioinformatics. 2011;Chapter 2:Unit 2 14.

81.   Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics. 2011;27:1696-1697.

82.   Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. Bioinformatics. 2011;27:1017-1018.

83.   Ho Sui SJ, Fulton DL, Arenillas DJ, Kwon AT, Wasserman WW. oPOSSUM: integrated tools for analysis of regulatory motif over-representation. Nucleic Acids Res. 2007;35:W245-252.

84.   Bailey TL, Boden M, Buske FA, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37:W202-208.

85.   Smyth GK, Michaud J, Scott HS. Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics. 2005;21:2067-2075.

86.   Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545-15550.

87.   Attema JL, Bert AG, Lim YY, et al. Identification of an enhancer that increases miR-200b~200a~429 gene expression in breast cancer cells. PLoS One. 2013;8:e75517.

88. Thorsteinsdottir U, Kroon E, Jerome L, Blasi F, Sauvageau G. Defining roles for HOX and MEIS1 genes in induction of acute myeloid leukemia. Mol Cell Biol. 2001;21:224-234.

89. Li Z, Huang H, Li Y, et al. Up-regulation of a HOXA-PBX3 homeobox-gene signature following down-regulation of miR-181 is associated with adverse prognosis in patients with cytogenetically abnormal AML. Blood. 2012;119:2314-2324.

90. Wang GG, Pasillas MP, Kamps MP. Persistent transactivation by meis1 replaces hox function in myeloid leukemogenesis models: evidence for co-occupancy of meis1-pbx and hox-pbx complexes on promoters of leukemia-associated genes. Mol Cell Biol. 2006;26:3902-3916.

91. Plowright L, Harrington KJ, Pandha HS, Morgan R. HOX transcription factors are potential therapeutic targets in non-small-cell lung cancer (targeting HOX genes in lung cancer). Br J Cancer. 2009;100:470-475.

92. Morgan R, Plowright L, Harrington KJ, Michael A, Pandha HS. Targeting HOX and PBX transcription factors in ovarian cancer. BMC Cancer. 2010;10:89.

93. Verhaak RG, Tamayo P, Yang JY, et al. Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. J Clin Invest. 2013;123:517-525.

94. Wu HP, Hsu SY, Wu WA, Hu JW, Ouyang P. Transgenic mice expressing mutant Pinin exhibit muscular dystrophy, nebulin deficiency and elevated expression of slow-type muscle fiber genes. Biochem Biophys Res Commun. 2014;443:313-320.

95. Li XY, Yao X, Li SN, et al. RNA-Seq profiling reveals aberrant RNA splicing in patient with adult acute myeloid leukemia during treatment. Eur Rev Med Pharmacol Sci. 2014;18:1426-1433.

96. Dey A, Seshasayee D, Noubade R, et al. Loss of the tumor suppressor BAP1 causes myeloid transformation. Science. 2012;337:1541-1546.

97. Torres R, Martin MC, Garcia A, Cigudosa JC, Ramirez JC, Rodriguez-Perales S. Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR-Cas9 system. Nat Commun. 2014;5:3964.

98. Wong P, Iwasaki M, Somervaille TC, So CW, Cleary ML. Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential. Genes Dev. 2007;21:2762-2774.

99. Faber J, Krivtsov AV, Stubbs MC, et al. HOXA9 is required for survival in human MLL-rearranged acute leukemias. Blood. 2009;113:2375-2385.

100. Morgan R, Pirard PM, Shears L, Sohal J, Pettengell R, Pandha HS. Antagonism of HOX/PBX dimer formation blocks the in vivo proliferation of melanoma. Cancer Res. 2007;67:5806-5813.

101. Su R, Lin HS, Zhang XH, et al. MiR-181 family: regulators of myeloid differentiation and acute myeloid leukemia as well as potential therapeutic targets. Oncogene. 2014;0.

102. Paschou M, Doxakis E. Neurofibromin 1 is a miRNA target in neurons. PLoS One. 2012;7:e46773.

103. Kotani A, Ha D, Hsieh J, et al. miR-128b is a potent glucocorticoid sensitizer in MLL-AF4 acute lymphocytic leukemia cells and exerts cooperative effects with miR-221. Blood. 2009;114:4169-4178.

104. Garcia-Orti L, Cristobal I, Cirauqui C, et al. Integration of SNP and mRNA arrays with microRNA profiling reveals that MiR-370 is upregulated and targets NF1 in acute myeloid leukemia. PLoS One. 2012;7:e47717.

105. Suresh S, McCallum L, Lu W, Lazar N, Perbal B, Irvine AE. MicroRNAs 130a/b are regulated by BCR-ABL and downregulate expression of CCN3 in CML. J Cell Commun Signal. 2011;5:183-191.

106. Loriaux MM, Levine RL, Tyner JW, et al. High-throughput sequence analysis of the tyrosine kinome in acute myeloid leukemia. Blood. 2008;111:4788-4796.

107. Kentsis A, Reed C, Rice KL, et al. Autocrine activation of the MET receptor tyrosine kinase in acute myeloid leukemia. Nat Med. 2012;18:1118-1122.

108.   Pedersen-Bjergaard J, Christiansen DH, Desta F, Andersen MK. Alternative genetic pathways and cooperating genetic abnormalities in the pathogenesis of therapy-related myelodysplasia and acute myeloid leukemia. Leukemia. 2006;20:1943-1949.

109.   Leach BI, Kuntimaddi A, Schmidt CR, Cierpicki T, Johnson SA, Bushweller JH. Leukemia fusion target AF9 is an intrinsically disordered transcriptional regulator that recruits multiple partners via coupled folding and binding. Structure. 2013;21:176-183.

110.   Parkin B, Ouillette P, Wang Y, et al. NF1 inactivation in adult acute myelogenous leukemia. Clin Cancer Res. 2010;16:4135-4147.

111.   Liu TC, Lin PM, Chang JG, Lee JP, Chen TP, Lin SF. Mutation analysis of PTEN/MMAC1 in acute myeloid leukemia. Am J Hematol. 2000;63:170-175.

112.   Luo Z, Lin C, Guest E, et al. The super elongation complex family of RNA polymerase II elongation factors: gene target specificity and transcriptional output. Mol Cell Biol. 2012;32:2608-2617.

113.   Gaidano G, Foa R, Dalla-Favera R. Molecular pathogenesis of chronic lymphocytic leukemia. J Clin Invest. 2012;122:3432-3438.

114.   Blank U, Karlsson S. The role of Smad signaling in hematopoiesis and translational hematology. Leukemia. 2011;25:1379-1388.

115.   Kwon M, Libutti SK. Filamin A interacting protein 1-like as a therapeutic target in cancer. Expert Opin Ther Targets. 2014:1-13.

116.   Gao XN, Lin J, Ning QY, et al. A histone acetyltransferase p300 inhibitor C646 induces cell cycle arrest and apoptosis selectively in AML1-ETO-positive AML cells. PLoS One. 2013;8:e55481.

117.   Collart MA, Panasenko OO, Nikolaev SI. The Not3/5 subunit of the Ccr4-Not complex: a central regulator of gene expression that integrates signals between the cytoplasm and the nucleus in eukaryotic cells. Cell Signal. 2013;25:743-751.

118.   Collart MA, Panasenko OO. The Ccr4--not complex. Gene. 2012;492:42-53.

119.   Wei N, Serino G, Deng XW. The COP9 signalosome: more than a protease. Trends Biochem Sci. 2008;33:592-600.

120.   Mao X, Shi X, Liu F, Li G, Hu L. Evaluation of erythroblast macrophage protein related to erythroblastic islands in patients with hematopoietic stem cell transplantation. Eur J Med Res. 2013;18:9.

APPENDIX I

- Supplementary tables

Supplementary Tables

Table S1: NUP98-HOXA9 target genes identified by ChIP-seq (see data on CD-room)

Table S2: Co-occurrences of HOXA9-MEIS1-PBX binding motif in the DNA target sequences of NH (FIMO results) (see data on CD-room)

Table S3: HOXA9 target genes identified by ChIP-seq (see data on CD-room)

Table S4: NUP98 target genes identified by ChIP-seq (see data on CD-room)

Table S5: Expression arrays in cellular models (see data on CD-room)

Table S6: Expression arrays in patients (see data on CD-room)

Table S7: Interacting-proteins indentify by Mass spectrometry assay

A. ONLY Nup98-Hoxa9 INTERACTING-PROTEINS	
CCR4-NOT transcription complex subunit 1	**CNOT1**
Isoform 2 of CCR4-NOT transcription complex subunit 1	**CNOT1**
CCR4-NOT transcription complex subunit 3	**CNOT3**
Nuclear RNA export factor 1	**NXF1**
Isoform 2 of CCR4-NOT transcription complex subunit 10	**CNOT10**
Cell differentiation protein RCD1 homolog	**RQCD1**
Isoform 2 of CCR4-NOT transcription complex subunit 7	**CNOT7**
CCR4-NOT transcription complex subunit 11	**CNOT11**
Isoform 3 of RalBP1-associated Eps domain-containing protein 1	**REPS1**
DNA repair protein XRCC1	**XRCC1**
AP-1 complex subunit beta-1 (Fragment)	**AP1B1**
AT-rich interactive domain-containing protein 1A	**ARID1A**
Phosphatidylinositol 4-phosphate 3-kinase C2 domain-containing subunit alpha	**PIK3C2A**
Isoform 2 of Remodeling and spacing factor 1	**RSF1**
Cytoplasmic FMR1-interacting protein 1	**CYFIP1**
Chromodomain-helicase-DNA-binding protein 4	**CHD4**
Isoform 10 of Abl interactor 1	**ABI1**
Uncharacterized protein C10orf88	**C10orf88**
BMP-2-inducible protein kinase	**BMP2K**
Nck-associated protein 1	**NCKAP1**
NTF2-related export protein 1	**NXT1**
Metastasis-associated protein MTA2	**MTA2**
Abl interactor 2 (Fragment)	**ABI2**
Protein Dr1	**DR1**
Cyclin-G-associated kinase	**GAK**
CCR4-NOT transcription complex subunit 8	**CNOT8**
Isoform 2 of Adaptin ear-binding coat-associated protein 1	**NECAP1**
Double-stranded RNA-binding protein Staufen homolog 2	**STAU2**

Prefoldin subunit 6	**PFDN6**
Zinc finger transcription factor Trps1 (Fragment)	**TRPS1**
Protein numb homolog (Fragment)	**NUMB**
CCR4-NOT transcription complex subunit 6-like (Fragment)	**CNOT6L**
Lon protease homolog, mitochondrial	**LONP1**
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 2	**SMARCD2**
UBX domain-containing protein 7	**UBXN7**
F-box/SPRY domain-containing protein 1	**FBXO45**
Pre-B-cell leukemia transcription factor 2	**PBX2**
Nucleolar and coiled-body phosphoprotein	**NOLC1**
Cyclic AMP-dependent transcription factor ATF-6 alpha	**ATF6**

B. ONLY Hoxa9 INTERACTING-PROTEINS

WD repeat-containing protein 6	**WDR6**
Translational activator GCN1	**GCN1L1**
Pentatricopeptide repeat domain-containing protein 3, mitochondrial	**PTCD3**
Casein kinase II subunit alpha'	**CSNK2A2**
DNA-directed RNA polymerase II subunit RPB3	**POLR2C**
28S ribosomal protein S6, mitochondrial	**MRPS6**
Thymidine kinase	**TK1**
Isoform 2 of WD repeat-containing protein 47	**WDR47**
DDB1- and CUL4-associated factor 10	**DCAF10**
Thymidine kinase, cytosolic	**TK1**
Anaphase-promoting complex subunit 1	**ANAPC1**
Probable cytosolic iron-sulfur protein assembly protein CIAO1	**CIAO1**
Ubiquitin-conjugating enzyme E2 O	**UBE2O**
Ribonuclease P protein subunit p30	**RPP30**
WD repeat domain phosphoinositide-interacting protein 3	**WDR45B**
Mitogen-activated protein kinase 1	**MAPK1**
Growth factor receptor-bound protein 2	**GRB2**
H/ACA ribonucleoprotein complex non-core subunit NAF1	**NAF1**
Paraneoplastic antigen Ma2	**PNMA2**
28S ribosomal protein S2, mitochondrial	**MRPS2**
Myc-associated zinc finger protein (Fragment)	**MAZ**
Transmembrane emp24 domain-containing protein 5	**TMED5**
Leucine-rich repeat-containing protein 1	**LRRC1**
RNA methyltransferase-like protein 1	**RNMTL1**
Schlafen family member 11	**SLFN11**
Protein SMG5	**SMG5**
2-oxoisovalerate dehydrogenase subunit beta, mitochondrial	**BCKDHB**

E3 ubiquitin-protein ligase HERC2	**HERC2**
rRNA 2'-O-methyltransferase fibrillarin	**FBL**
Nucleolar protein 56	**NOP56**
Isochorismatase domain-containing protein 2, mitochondrial (Fragment)	**ISOC2**
28S ribosomal protein S24, mitochondrial	**MRPS24**
Isoform 3 of FAD synthase	**FLAD1**
Probable serine carboxypeptidase CPVL	**CPVL**
CDKN2A-interacting protein	**CDKN2AIP**
HIG1 domain family member 1A, mitochondrial	**HIGD1A**
GTP-binding protein 1	**GTPBP1**
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A	**STT3A**
Tripartite motif-containing protein 26 (Fragment)	**TRIM26**
28S ribosomal protein S9, mitochondrial	**MRPS9**
Transmembrane emp24 domain-containing protein 1	**TMED1**
Interferon-inducible double stranded RNA-dependent protein kinase activator A	**PRKRA**
Peptidyl-prolyl cis-trans isomerase-like 4	**PPIL4**
Nucleolar complex protein 2 homolog	**NOC2L**
Nucleolar MIF4G domain-containing protein 1	**NOM1**
Microsomal glutathione S-transferase 3	**MGST3**
Isoform 3 of Protein spinster homolog 1	**SPNS1**
Lysine-rich nucleolar protein 1 (Fragment)	**KNOP1**
Probable ATP-dependent RNA helicase DDX41	**DDX41**
Isoform 2 of ATP-dependent RNA helicase DDX54	**DDX54**
Isoform 2 of Acidic leucine-rich nuclear phosphoprotein 32 family member B	**ANP32B**
ATP-dependent RNA helicase DDX55	**DDX55**
RISC-loading complex subunit TARBP2 (Fragment)	**TARBP2**
Casein kinase I isoform alpha (Fragment)	**CSNK1A1**
Isoform 2 of Replication factor C subunit 5	**RFC5**
Serine palmitoyltransferase 1	**SPTLC1**
Coatomer subunit gamma-2	**COPG2**
Isoform 2 of Histone deacetylase 6	**HDAC6**
Isoform 3 of PCI domain-containing protein 2]	**PCID2**
Isoform 2 of RNA-binding protein 28	**RBM28**
2-oxoisovalerate dehydrogenase subunit alpha, mitochondrial (Fragment)	**BCKDHA**
Isoform 2 of Anaphase-promoting complex subunit 7	**ANAPC7**
Protein FAM91A1	**FAM91A1**
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1	**PLOD1**
Serine/threonine-protein kinase A-Raf	**ARAF**
28S ribosomal protein S18c, mitochondrial	**MRPS18C**
STAGA complex 65 subunit gamma	**SUPT7L**
Protein phosphatase 1 regulatory subunit 3D	**PPP1R3D**
Mediator of RNA polymerase II transcription subunit 8	**MED8**
Isoform C of Isocitrate dehydrogenase [NAD] subunit beta, mitochondrial	**IDH3B**

Protein phosphatase 1F	**PPM1F**
General transcription factor 3C polypeptide 5	**GTF3C5**
Isoform 2 of ATPase WRNIP1	**WRNIP1**
Protein zer-1 homolog	**ZER1**
SART3 protein	**SART3**
DNA-directed RNA polymerase I subunit RPA2	**POLR1B**
Nischarin	**NISCH**
Anaphase-promoting complex subunit 10 (Fragment)	**ANAPC10**
Mediator of RNA polymerase II transcription subunit 15	**MED15**
Syntaxin-17	**STX17**
Casein kinase I isoform epsilon	**CSNK1E**
Isoform 4 of Interleukin-1 receptor-associated kinase 1	**IRAK1**
Ubiquitin carboxyl-terminal hydrolase 10	**USP10**
Polynucleotide 5'-hydroxyl-kinase NOL9	**NOL9**
Isoform 4 of Elongator complex protein 2	**ELP2**
Rab3 GTPase-activating protein non-catalytic subunit	**RAB3GAP2**
DNA-directed RNA polymerase II subunit RPB1	**POLR2A**
Mediator of RNA polymerase II transcription subunit 14	**MED14**
Dehydrogenase/reductase SDR family member 7B (Fragment)	**DHRS7B**
Target of rapamycin complex subunit LST8	**MLST8**
C-terminal-binding protein 1 (Fragment)	**CTBP1**
Acyl-coenzyme A thioesterase 8 (Fragment)	**ACOT8**
Isoform 3 of RNA-binding protein Musashi homolog 2	**MSI2**
Ferritin	**FTH1**
Transcription initiation factor IIA subunit 2	**GTF2A2**
Proteasome activator complex subunit 3 (Fragment)	**PSME3**
Cell division cycle protein 27 homolog (Fragment)	**CDC27**
FKBP8 isoform 1	**FKBP8**
Isoform 2 of Ras suppressor protein 1	**RSU1**
N-alpha-acetyltransferase 30 (Fragment)	**NAA30**
Isoform 3 of 28S ribosomal protein S11, mitochondrial	**MRPS11**
BRCA1-A complex subunit RAP80 (Fragment)	**UIMC1**
Cysteine desulfurase, mitochondrial (Fragment)	**NFS1**
E3 ubiquitin-protein ligase TRIM32 (Fragment)	**TRIM32**
Myelin expression factor 2 (Fragment)	**MYEF2**
Inositol polyphosphate 5-phosphatase K (Fragment)	**INPP5K**
Isoform 2 of TBC1 domain family member 10B	**TBC1D10B**
Isoform 2 of Zinc finger CCHC domain-containing protein 10	**ZCCHC10**
Striatin-4 (Fragment)	**STRN4**
Zinc finger protein 658B	**ZNF658B**
60S ribosomal protein L7-like 1	**RPL7L1**
Nuclear pore complex protein Nup205	**NUP205**
AP-3 complex subunit beta-2 (Fragment	**AP3B2**

mRNA turnover protein 4 homolog	**MRTO4**
Alpha-2-macroglobulin receptor-associated protein	**LRPAP1**
Protein FAM203B	**FAM203B**
ATPase family AAA domain-containing protein 1	**ATAD1**
Isoform 2 of 28S ribosomal protein S5, mitochondrial	**MRPS5**
Protein THEM6	**THEM6**
Isoform 2 of Probable ATP-dependent RNA helicase DDX47	**DDX47**
Copine-1	**CPNE1**
Protein CMSS1 (Fragment)	**CMSS1**
Mannose-1-phosphate guanyltransferase beta	**GMPPB**
Four and a half LIM domains protein 3	**FHL3**
Copine-3	**CPNE3**
Mannose-1-phosphate guanyltransferase alpha	**GMPPA**
Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial	**SDHB**
Sorting and assembly machinery component 50 homolog	**SAMM50**
S1 RNA-binding domain-containing protein 1	**SRBD1**
Isoform 2 of Transcription elongation factor SPT5	**SUPT5H**
Lysophosphatidylcholine acyltransferase 1	**LPCAT1**
Leucine-rich repeat-containing G-protein-coupled receptor 6 (Fragment)	**LGR6**
E3 ubiquitin-protein ligase TRIM56	**TRIM56**
NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial	**NDUFV1**
Isoform 2 of Interferon-induced, double-stranded RNA-activated protein	**EIF2AK2**
WD repeat-containing protein 75	**WDR75**
Fanconi anemia group I protein	**FANCI**
Tetratricopeptide repeat protein 27	**TTC27**
Leucine-rich repeat-containing protein 47	**LRRC47**
Isoform 2 of Pre-mRNA-processing factor 6	**PRPF6**
E3 ubiquitin-protein ligase HECTD3	**HECTD3**
Genetic suppressor element 1 (Fragment)	**GSE1**
Intraflagellar transport protein 81 homolog (Fragment)	**IFT81**
Ubiquitin-protein ligase E3C	**UBE3C**
Isoform 3 of Ribonuclease 3	**DROSHA**
Helicase SKI2W	**SKIV2L**
Isoform 3 of Protein VPRBP	**VPRBP**
Isoform 2 of Nucleoporin NUP188 homolog	**NUP188**
Mediator of RNA polymerase II transcription subunit 1	**MED1**

C. ONLY Nup98 INTERACTING-PROTEINS

MOB kinase activator 2	**MOB2**
Cold-inducible RNA-binding protein	**CIRBP**
Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 2	**ACAP2**
Isoform 3 of CAP-Gly domain-containing linker protein 1	**CLIP1**
Crk-like protein	**CRKL**
Survival of motor neuron-related-splicing factor 30	**SMNDC1**
Isoform B of Peptidyl-prolyl cis-trans isomerase E	**PPIE**
Bystin	**BYSL**
Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	**DLAT**
Methylcytosine dioxygenase TET2	**TET2**
Eukaryotic translation initiation factor 1A, Y-chromosomal	**EIF1AY**
Isoform 3 of Dynactin subunit 3	**DCTN3**
WD repeat-containing protein 5	**WDR5**
Succinate-semialdehyde dehydrogenase, mitochondrial	**ALDH5A1**
Peptidyl-prolyl cis-trans isomerase B	**PPIB**
Splicing regulatory glutamine/lysine-rich protein 1	**SREK1**
Isoform 4 of Dedicator of cytokinesis protein 7	**DOCK7**
Sjoegren syndrome/scleroderma autoantigen 1	**SSSCA1**
Peptidyl-prolyl cis-trans isomerase H (Fragment)	**PPIH**
Isoform 2 of PEST proteolytic signal-containing nuclear protein	**PCNP**
Isoform 4 of RNA-binding protein 42	**RBM42**
Actin-related protein T1	**ACTRT1**
La-related protein 7 (Fragment)	**LARP7**
Isoform 2 of Proteasome subunit alpha type-7	**PSMA7**
Actin-related protein 2/3 complex subunit 5	**ARPC5**
Guanine nucleotide exchange factor for Rab-3A (Fragment)	**RAB3IL1**
H/ACA ribonucleoprotein complex subunit 4	**DKC1**
Brain-specific angiogenesis inhibitor 1-associated protein 2-like protein 1	**BAIAP2L1**
Anoctamin-5	**ANO5**
U7 snRNA-associated Sm-like protein LSm10	**LSM10**
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	**NUCKS1**
Thioredoxin-dependent peroxide reductase, mitochondrial	**PRDX3**
60S ribosomal export protein NMD3	**NMD3**
SAP30-binding protein (Fragment)	**SAP30BP**
Tubulin--tyrosine ligase-like protein 12	**TTLL12**
Peroxisomal targeting signal 1 receptor (Fragment)	**PEX5**
Protein LYRIC	**MTDH**
Mammalian ependymin-related protein 1	**EPDR1**
Glutaredoxin-3	**GLRX3**
Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16	**DHX38**
Isoform Short of Tight junction protein ZO-1	**TJP1**
Rho guanine nucleotide exchange factor 10 (Fragment)	**ARHGEF10**
Isoform 2 of Pre-mRNA-processing factor 40 homolog A	**PRPF40A**

D. COMMON HOXA9 and NUP98-HOXA9 INTERACTING PROTEINS

Insulin-degrading enzyme	**IDE**
Casein kinase II subunit alpha	**CSNK2A1**
Isoform 7 of Protein LAP2	**ERBB2IP**
Transcription factor A, mitochondrial	**TFAM**
Mitogen-activated protein kinase kinase kinase kinase 4	**MAP4K4**
Protein RMD5 homolog A	**RMND5A**
Ribosome biogenesis protein BOP1	**BOP1**
Isoform 4 of Transcription factor Sp3	**SP3**
COP9 signalosome complex subunit 7a	**COPS7A**
Homeobox protein Hox-A9	**HOXA9**
COP9 signalosome complex subunit 4	**COPS4**
Isoform 2 of Glycogen [starch] synthase, muscle	**GYS1**
COP9 signalosome complex subunit 6	**COPS6**
COP9 signalosome complex subunit 2	**COPS2**
COP9 signalosome complex subunit 5	**COPS5**
COP9 signalosome complex subunit 1	**GPS1**
Isoform 2 of Armadillo repeat-containing protein 8	**ARMC8**
Replication protein A 32 kDa subunit	**RPA2**
Putative adenosylhomocysteinase 2	**AHCYL1**
NAD-dependent protein deacetylase sirtuin-1	**SIRT1**
E3 ubiquitin-protein ligase ARIH1	**ARIH1**
COP9 signalosome complex subunit 8	**COPS8**
Isoform 3 of Serine protease HTRA2, mitochondrial	**HTRA2**
Isoform 2 of Putative adenosylhomocysteinase 3	**AHCYL2**
Protein FAM98A	**FAM98A**
Isoform 2 of Uncharacterized protein C18orf25	**C18orf25**
Replication protein A 14 kDa subunit	**RPA3**
Homeobox protein Hox-D9	**HOXD9**
Muskelin	**MKLN1**
Importin subunit alpha-3	**KPNA3**
Isoform GN-1 of Glycogenin-1	**GYG1**
Cleavage and polyadenylation-specificity factor subunit 7 (Fragment)	**CPSF7**
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5	**SMARCA5**
Molybdopterin molybdenumtransferase	**GPHN**
RNA-binding protein 27	**RBM27**
Intraflagellar transport protein 46 homolog (Fragment)	**IFT46**
Mov10, Moloney leukemia virus 10, homolog (Mouse), isoform CRA_a	**MOV10**
Macrophage erythroblast attacher	**MAEA**
Isoform 2 of COP9 signalosome complex subunit 7b	**COPS7B**
Structural maintenance of chromosomes protein 3	**SMC3**
High mobility group protein HMG-I/HMG-Y	**HMGA1**
E3 ubiquitin-protein ligase RAD18	**RAD18**
E3 ubiquitin-protein ligase ARIH2	**ARIH2**
Isoform 4 of Nucleolar and spindle-associated protein 1	**NUSAP1**

Isoform HMG-Y of High mobility group protein HMG-I/HMG-Y	HMGA1
Isoform 6 of Cyclin-dependent kinase inhibitor 2A, isoform 4	CDKN2A
Phosphorylated adapter RNA export protein	PHAX
Vacuolar protein sorting-associated protein 18 homolog	VPS18
Glucose-induced degradation protein 4 homolog	GID4
Dr1-associated corepressor (Fragment)	DRAP1

E. COMMON NUP98 and NUP98-HOXA9 INTERACTING PROTEINS

Isoform 2 of UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit	OGT
Splicing factor, suppressor of white-apricot homolog	SFSWAP
Prefoldin subunit 5 (Fragment)	PFDN5
Pinin	PNN

F. COMMON HOXA9, NUP98 and NUP98-HOXA9 INTERACTING PROTEINS

Ran-binding protein 10	RANBP10
Histone deacetylase	HDAC2
Histone deacetylase 1	HDAC1
Exosome complex component CSL4	EXOSC1
Transcription elongation factor B (SIII), polypeptide 2 (18kDa, elongin B), isoform CRA_b	TCEB2

APPENDIX II

- Publications
- Oral communications and posters of the meetings attended

PUBLICATIONS

1. **Ana Rio-Machin,** Gonzalo Gomez-Lopez, Alba Maiques-Diaz, Maureen A. Powers, Sara Alvarez, Rocio N Salgado, Claudia Haferlach, Maria Jose Larrayoz' Maria Jose Calasanz and Juan C. Cigudosa. *New insights into the contribution of the two moieties of NUP98-HOXA9 provide a wider window of targeted therapy in acute myeloid leukemia.* Oct 2014. Submitted to *Blood*.

2. **Ana Rio-Machin**, Javier Muñoz, Fernando Garcia-Martinez, Maureen A. Powers and Juan C. Cigudosa. *The oncogene NUP98-HOXA9, more than a leukemic transcription factor.* Oct 2014. Submitted to *Haematologica*.

3. Alba Maiques-Diaz, Alvaro Eguileor, Mahesh Shestha, Gonzalo Gómez López, Mark Wunderlich, **Ana Rio-Machin**, Xavier Agirre , James C. Mulloy, Juan C. Cigudosa and Sara Alvarez. *Aberrant chromatin acetylation in MLL-AF9 leukemia mediates the response to HDAC inhibition.* Sep 2014. Submitted to *Plos One*.

4. Rocío N Salgado, **Ana Rio-Machin**, Juliane Menezes, C Guillén, J Sánchez-Calero, A Sánchez-López, F García-Sánchez, Sara Álvarez and Juan C Cigudosa. *GTF2I/RARA fusion gene: a novel RARA rearrangement in a variant acute promyelocytic leukemia resistent to ATRA* (in preparation).

5. Alberto Cascon, Iñaki Comino-Méndez, María Currás-Freixes, Aguirre De Cubas, Laura Contreras, Susan Richter, Veronika Mancikova, Andrés Pérez-Barrios, María Calatayud, Sharona Azriel, Rosa Villar-Vicente, Javier Aller, Fernando Setién, Juan Garcia, **Ana Rio-Machin**, Rocío Letón, Álvaro Gómez-Graña, Lucia Inglada-Pérez, María Apellániz-Ruiz, Giovanna Roncador, Manel Esteller, Cristina Rodríguez-Antona, Jorgina Satrústegui, Graeme Eisenhofer, Miguel Urioste, and Mercedes Robledo. *Whole-exome sequencing identifies MDH2 as a new familial paraganglioma gene. Oct 2014.* Provisionally accepted to JNCI. (IF: 15.161)

6. **Ana Rio-Machin**, Bibiana Ferreira, Travis Henry, Gonzalo Gómez-López, Xabier Agirre, Sara Alvarez, Sandra Rodriguez-Perales, Felipe Prosper, M José Calasanz, Joaquín Martínez, Rafael Fonseca, and Juan C. Cigudosa. *The downregulation of specific miRNAs in hyperdiploid multiple myeloma mimics the effect of the most frequent IgH translocations observed in the non-hyperdiploid subtype*. Leukemia. 2013 Apr;27(4):925-31 (IF: 10.164)

7. **Ana Rio-Machin,** Juliane Menezes, Agirre X, Ferreira BI, Acquadro F, Rodriguez-Perales S, Arribalzaga Juaristi K, Alvarez S, Cigudosa JC. *Abrogation of RUNX1 gene expression in* de novo *myelodysplastic syndrome with t(4;21)(q21;q22).* Haematologica. 2012 Apr;97(4):534-7 (IF: 6.424)