UNIVERSITÀ DEGLI STUDI DI MILANO
SCUOLA DI DOTTORATO IN
SCIENZE BIOLOGICHE E MOLECOLARI
Ph.D. School in Biomolecular Sciences

Relationship between NF-Y a general transcription factor and Ash2L a component of MLL complex

Dr. Andrea Fossati

TUTOR:

Prof. Roberto Mantovani

COORDINATORE DEL DOTTORATO:

Prof. Roberto Mantovani

***To Silvia**, ...Thanks for Everything.*

INDEX

INTRODUCTION

1. Chromatin structure

In eukaryotic cells, DNA is packaged with proteins, the histones, to form chromatin fibers.

The fundamental unit of chromatin is the nucleosome, formed by 146 bp. of DNA wrapped around four heterodimers of H2A-H2B and H3-H4 core histones. Histones are among the most conserved proteins in eukaryotes; they are formed by N- and C-terminal tails and a globular part, the histone-fold domain.

The histone tails have long been known to be modified by a plethora of post-translational modifications (PTMs) and it is now clear that these are marks of peculiar chromatin environments (Berger 2007, Ruthenburg 2007).

Some of them are associated with accessible, active chromatin, others with heterochromatin, either constitutive or facultative.

Histones, are highly alkaline proteins which package and order the DNA into structural units.

Histones are highly conserved and can be grouped into five classes: H1, H2A, H2B, H3, and H4.

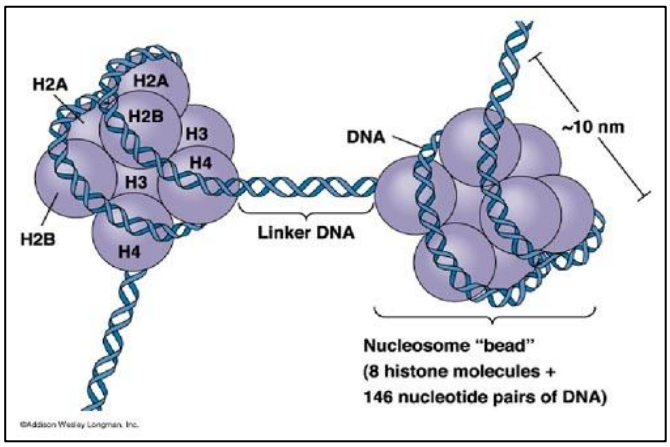


***Fig. 1**: Nucleosome structures*

These class are organized into two super-classes as follows:

- core histones – H2A, H2B, H3 and H4
- linker histones H1

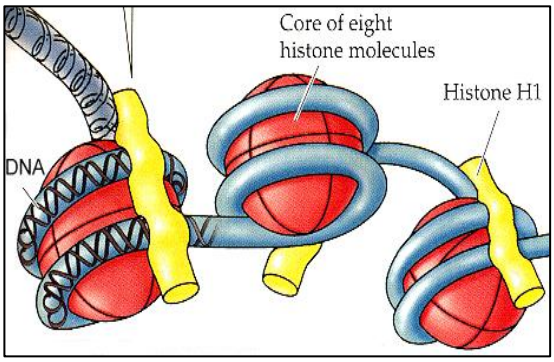


***Fig.2** Histone H1*

The nucleosome core is formed of two H2A-H2B dimers and H3-H4 tetramer in which, segments of DNA (146 bp) are wrapped around a protein core, the histone octamer (Fig 1).

The linker histone H1 binds the nucleosome and the entry and exit

sites of the DNA, thus locking DNA into place and allowing the formation of higher order structure (Fig 2).

2. The histone code

Histones tails are subject to post-traductional modifications, including methylation, acetylation, phosphorylation, sumorylation and ubiquitination.

These covalent histone modifications (Fig 3), by altering chromatin structure, function to modulate gene expression or other processes and allow the recruiting of additional effector proteins.

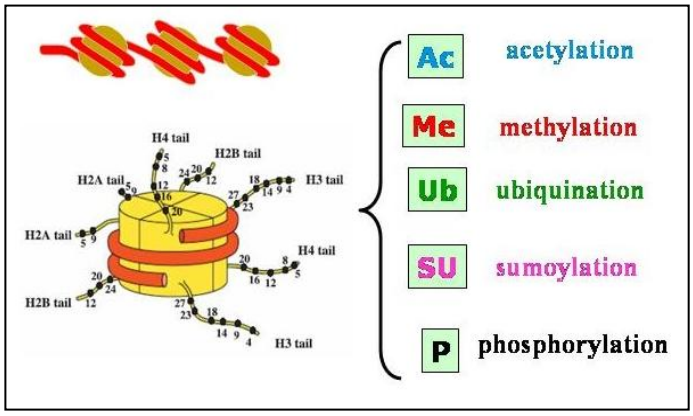


Fig.3 *Type of Histone tails modification.*

In others hands, histone post-translational modifications are marks of chromatin environments.

Some of them are associated with accessible chromatin, others with heterochromatin, either constitutive or facultative (Berger 2007).

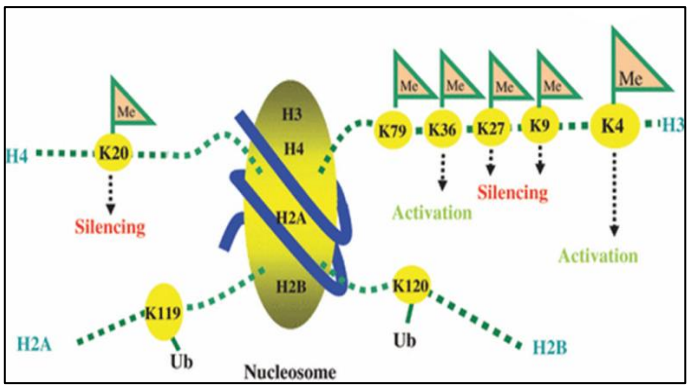


Fig.4 *methilation of histone tails*

Acetylations of H3 and H4, in particular, are believed to be hallmark of active areas of genomes, the Methylation, instead, is complex signals because in some residues is associated with 'open' or transcribed chromatin (H3- K4, H3-K36 and H3-K79) but in others cases (H3-K9, H3-K27 and H4-K20) are signposts of repression (Fig.4)

Furthermore the different number of methyl groups imposed on the same lysine (single, double or triple methylations) can be marks of different chromatin states.

The monoubiquitination of H2B at Lysine 120 (K123 in yeast) is one of the earliest events in the establishment of an active chromatin environment (Nakanishi 2009).

H3 methylations follow, on H3K4 and H3K79, in a hierarchy of events that leads to gene activation.

Methylation of H3K4 is highly regulated (Ruthenburg 2007), and generally present in active and poised promoters (Barski 2007).

3. NF-Y a general transcription factor

NF-Y is a complex composed of three subunits (Fig. 4): NF-YA (CBF-B, HAP2 in yeast), NF-YB (CBF-A, HAP3 in yeast) and NF-YC (CBF-C, HAP5 in yeast) (Fig 4).

The NF-YB-NF-YC subunits form a tight dimer, which offers a complex surface for NF-YA sub-unit association.

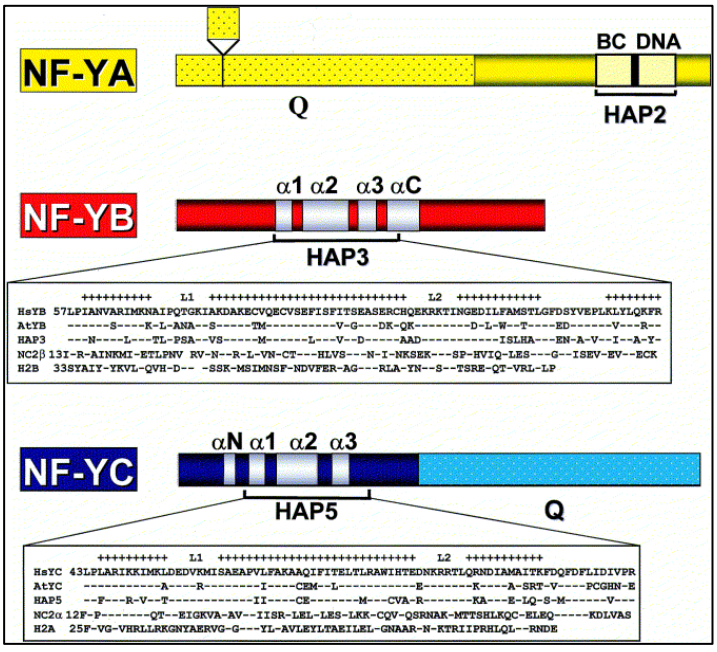


Fig.4 Schematic representation of the NF-Y sub-unit genes

The resulting trimer can then bind to DNA with high specificity and affinity: for most of the sites, the Kd is between 10^{-10} and 10^{-11}, among the highest of all transcription factors.

The transcriptional activation functions are in the large N-terminal of NF-YA and in the C-terminal of NF-YC.(Coustry 1995).

The yeast homology domains are indicated by brackets. White boxes in NF-YB and NF-YC indicate the position of the four a-helices of the histone fold domains. Below the NF-YB and NF-YC schemes are shown the sequences of the conserved domains of the human NF-YB and NF-YC in others organisms: A. thaliana At-YB/YC (Edwards 1998), S. cerevisiae HAP3/5, NC2β/a (Goppelt 1996) and Xenopus H2B/H2A and of the corresponding sequences of NF-YB/C, A. thaliana At-YB/YC, HAP3/5, human NC2b/ a and Xenopus H2B/H2A)

The" +" symbol indicates the position of the a-helices in the histone folds; the L1 and L2 are the loop regions.

The eterotrimer (NF-Y),when is folded, can recognize and bind DNA in sequence specific manner.

The sequence that NF-Y is able to recognize is the pentanucleotide named Double C Double A T Box, (CCAAT Box) (Fig.5), one of the most frequent and crucial promoter elements, that normally is located in the region of the core promoter of the genes.

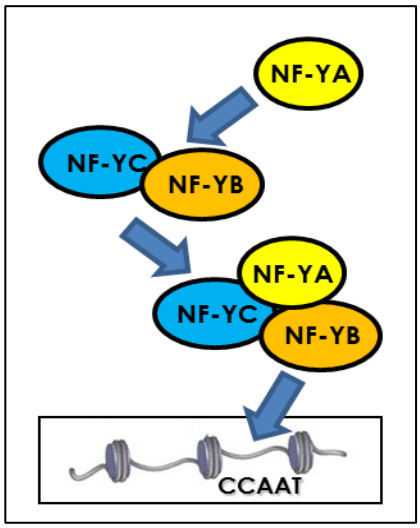


Fig.5 *Association of NF-Y subunits and binding to DNA*

In a statistical analysis of over 500 promoters, was found that the CCAAT box is one of the most ubiquitous elements, being present in 30% of eukaryotic promoters (Bucher, 1990).

With a strong bias in promoter position. Typically, the CCAAT box is found as a single copy element in the forward or reverse orientation between −60 and −100 of the major start site.

Although multiple sites are found, they are not the rule; their distance is variable, but in no cases are they separated by less than 27 base pairs.

Is known in literature that the bind of NF-Y to DNA (CCAAT Box) is important for the trascriptional activation (Mantovani , 1990; Romier, 2003), this function, as I said, is in the large N-terminal of NF-YA and in the C-terminal of NF-YC, terminals rich in glutammines and hydrophobic residues.

Furthermore in a previous work from our laboratory (my tesi's job) we have seen that the presence of H3K4me3 is dependent upon the binding of NF-Y (Donati, 2008), in NF-Y target genes.

4. Lysine methyl transferase (KMT) MLL, a chromatin remodelling complex.

As I said, Histone post-translational modifications are crucial in the regulation of the state of the chromatin (open/close).

In this contest methylation is one of major mechanism, and there are enzymes responsable of the apposition and others that removes the metyl groups.

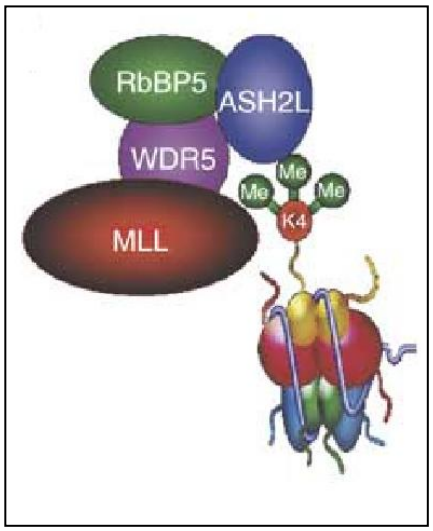


Fig.6 *KMT MLL*

The Histone methyltransferases (KMT) are enzymes, that catalyze the transfer of one to three methyl groups from the cofactor S-Adenosyl methionine to lysine and arginine residues of histone proteins.

Mixed lineage leukemias (MLLs) are an evolutionarily conserved trithorax family of human genes.

There are several MLL family proteins such as MLL1, MLL2, MLL3, MLL4, MLL5, Set1A and Set1B, and each possesses a locus specific histone H3 lysine 4 (H3K4)-methyltransferase activity and has critical roles in gene activation and epigenetics.

MLL1 is well known to be rearranged in myeloid and lymphoid leukemias in children and adults.

MLL1 (gene ALL1), is a very large complex and it's the major H3K4 methyltransferase, and is assembled in a complex that includes several proteins, like Menin, Ash2L, WDR5, RbBP5 and DPY30 (Nakamura, 2002; Steward , 2006).

ASH2L, RbBP5 and WDR5 are conserved subunits of MLL complexes with homology to the Cps40/Cps60, Cps50 and Cps30 subunits of COMPASS, respectively in yeast.

ASH2L differentially regulates MLL's catalysis of H3K4 trimethylation similarly to Cps40 and Cps60.

WDR5 is another very important protein of the MLLs complex and is required to maintain MLL complex integrity, including the stability of ASH2L within the complex.(Melissa M Steward, 2006)

5. MLL1: structure

The gene ALL1 is located on the 11 chromosome, includes a very large area of DNA over 88.000 bp. and includes 37 exon.

The 3,968 amino acid-containing MLL1 protein (430KDa) consists of an N-terminal (290 KDa) A-T hook DNA binding domain, a DNA methyltransferase-like domain with several continuous zinc fingers near to the center of the molecule and a conserved SET domain at its C-terminal domain (134 KDa) (Hsieh, 2003).

MLL is proteolytically cleaved into N-terminal p320 and C-terminal p180 fragments, the cleavage site(s) resides between aa 2253 and 2727 (CS1 and CS2) (James, 2003).

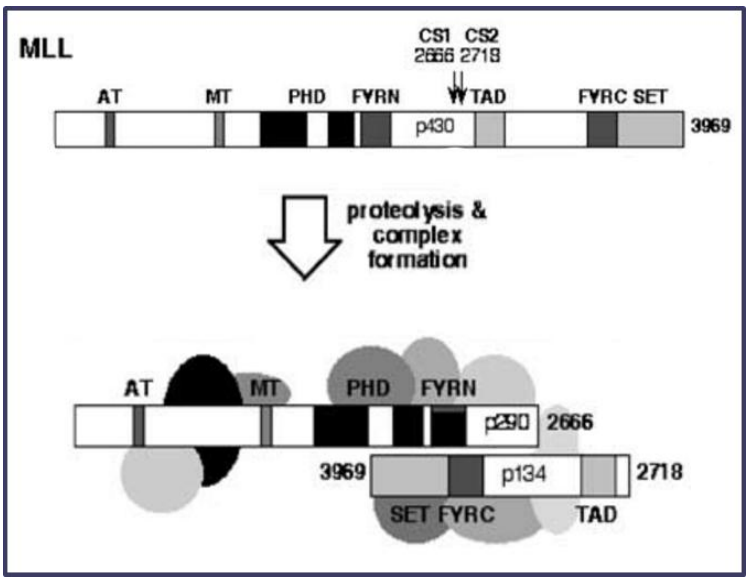


Fig.7 *clevage sites CS1 and CS2*

The large MLL1 protein is cut by an aspartic protease called taspase into an N-terminal (320kDa) fragment and a C-terminal (180kDa) moiety that are both core components of the MLL complex (***Fig. 7***). (Hsieh, 2003; Yokoyama, 2002).

The MLL-C subunit associates with proteins that help in preparing chromatin for efficient transcription. The WDR5 protein in turn recognizes the histone H3 lysine 4 methyl-mark introduced by MLL1 and it has therefore been suggested that WDR5 ensures the processitivity of histone modification.(Robert, 2009; Ruthenburg, 2006) And finally the proteins RBBP5 and ASH2L appear to be necessary for efficient methyltransferase activity by stabilizing an active conformation of MLL allowing allosteric control.

MLL1-N on the other side contains features essential for correct targeting of the MLL1 complex. At the outmost amino-terminal end of MLL1 a binding site for Menin, the product of the tumor suppressor gene multiple endocrine neoplasia is

present.(Caslini, 2007;- Milne, 2005) Menin and MLL1 form an interaction surface for LEDGF (lens epithelium derived growth factor) and LEDGF makes contact to chromatin via a PWWP domain.(Yokoyama, 2008)

In addition, MLL1-N codes for several AT-hooks, a minor groove DNA binding motif that preferentially recognizes DNA with distortions like bends or kinks.(Zeleznik, 1994) Further downstream a CxxC domain can be found. CxxC domains occur in proteins that discriminate the methylation status of DNA, and indeed, also the MLL1 CxxC moiety binds specifically to unmethylated CpG dinucleotides (Birke, 2002).

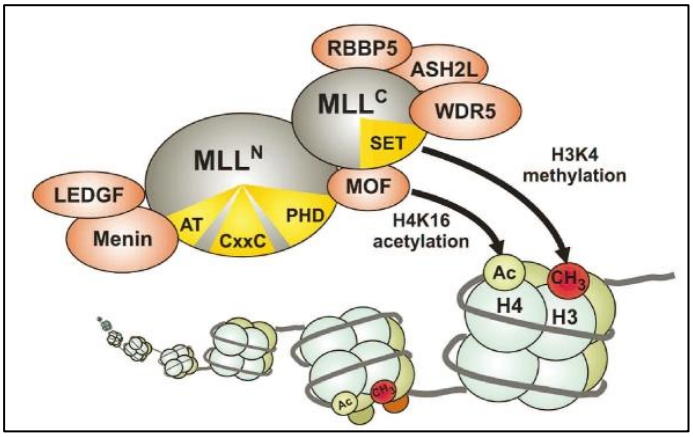


Fig.8 *The MLL1 complex. After post-transcriptional proteolytic processing amino-terminal and carboxy-terminal portions of MLL are incorporated in a macromolecular complex with histone methyltransferase and histone acetyltransferase function. Functional domains in MLL are indicated in yellow. AT = AT-hooks, a DNA binding domain, CxxC = motif recognizing unmethylated CpG dinucleotides, PHD = plant homeodomain, SET = histone methyltransferase active site. Proteins associated with MLL are explained in the text.*

6. MLL1, a high frequency translocations locus

The length of the sequence and the presence of several active structural motifs make the MLL1 locus prone to mutations.

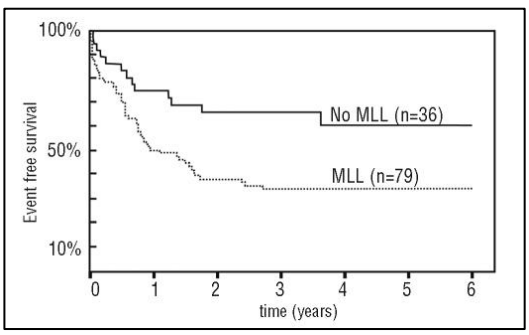


Fig. 9 *Event free survival of infants with ALL separated by MLL status*

More than 30 years ago, physicians realized that certain subsets of patients initially diagnosed with acute lymphoblastic (ALL) or acute myeloid leukemia (AML) fared far worse than others.

In the pediatric field, one of these high-risk leukemias stood out particularly amongst all remaining cases of childhood leukemia. A cohort of ALLs diagnosed in newborns and infants (younger than one year) fell into a group with similar clinical aspects and an extremely dismal prognosis.

With the advent of fluorescent activated cell sorting (FACS) it was revealed that the leukemic blasts of these aggressive leukemias frequently expressed surface markers of both the lymphoid and the myeloid lineage.

Sometimes even a complete lineage switch was observed during treatment and a leukemia initially diagnosed as ALL could relapse as AML.(Stass, 1984) Accordingly the term mixed lineage leukemia was coined (Mirro ,1985; Mirro 1986).

Even before this, cytogeneticists had noted that translocations affecting the locus 11q23, and in particular the translocation t(4;11), characterize a special subset of ALL that was associated with poor survival.(Morse, 1982; Mirro, 1985)

Soon thereafter it became clear that these translocations of the locus 11q23 are also typical for mixed lineage leukemia. Whereas treatment of non-mixed lineage leukemia in children has become the textbook success story of modern medicine with 5-year survival rates approaching 90%,(Pui, 2004) mixed lineage leukemia treatment seems to have hit a roadblock with hardly 40% of all infants surviving five years after diagnosis (Fig.9).(Pui, 2004;Tomizawa, 2007).

11q23 abnormalities occur in up to 70% of infant ALL, and in approximately 10% of all other ALL cases.(Armstrong, 2005).

Much has been speculated about the origins of the chromosomal aberrations that convert an innocuous chromatin modifier into a pernicious oncogene. Several lines of evidence point to a mishap in non-homologous end joining of double strand breaks as the most likely reason for 11q23 translocations.

For one, the characteristic peak of mixed lineage leukemia in patients treated with etoposide is highly suggestive for an involvement of DNA double strand lesions in the etiology of MLL fusions.

Etoposide inhibits topoisomerase II and therefore causes breaks in both DNA strands. Indeed, it could be shown that the locus 11q23 is particularly susceptible to this kind of assault in cells treated with Topo II inhibitors.(Hars, 2006; Libura, 2005) Alternatively, a break might be introduced at early stages of apoptotic DNA fragmentation that was later aborted and repaired.

The first and most striking property of MLL1 fusion proteins is their incredible diversity.

MLL has been found in 73 different translocations and 54 partner genes have been cloned (fig.10) (http://atlasgeneticsoncology.org/Genes/MLL.html; last update 5/08). Despite this variety most cases of mixed lineage leukemia present as a clinical entity and gene expression signatures in leukemic blasts do not separate MLL1 fusions according to the fusion partner (Ross, 2003; Armstrong, 2002).

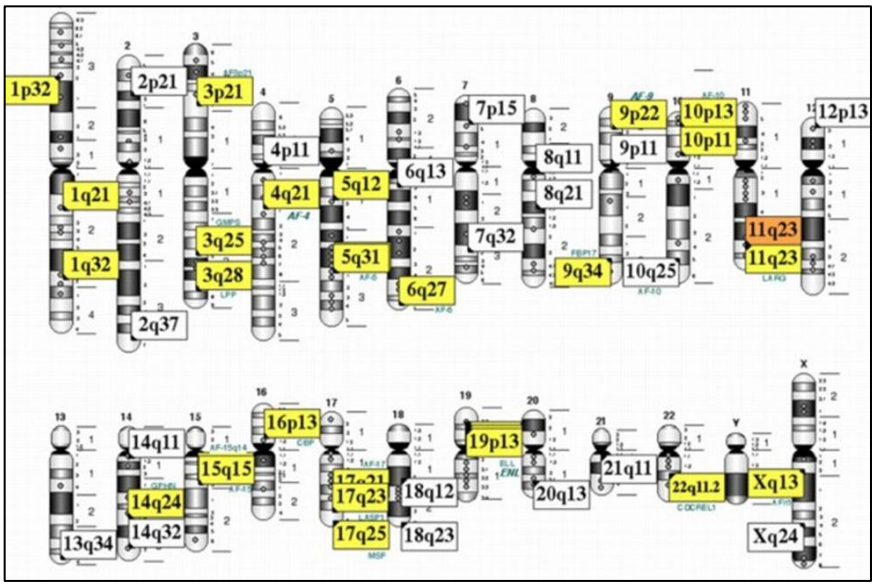


Fig.10 *common mutation in ALL1 gene*

Despite the amount of mutation discovered, only 6 frequent partner proteins (AF4, AF9, ENL, AF10, ELL, AF6) constitute the bulk (> 85%) of all clinical cases of mixed lineage leukemia (Burmeister, 2009).

This distinction is mirrored by the biology of the respective proteins. With the exception of AF6, all frequent MLL partners are nuclear while cytoplasmatic localization predominates amongst the rarely occurring MLL fusions.

7. The “common” nuclear fusion partners of MLL

Early reports showed that MLL1 fusions function as a novel type of general transcription factor that is able to indiscriminately activate many different promoters (Schreiner, 1999).

The search for a common MLL1 machinery revealed that the close homologs ENL and AF9 were both able to interact with other MLL fusion partners like AF4, the AF4-homolog AF5 and probably also with AF10.

In addition, it was realized that ENL could bind to histone H3, indicating a potential shared link of these proteins with chromatin modification.(Zeisig, 2005)
A breakthrough concerning the normal function of these proteins came from the purification of the ENL associated protein complex (EAP) (Bitoun, 2007; Mueller, 2007).

In this complex, ENL was not only linked with all members of the AF4 protein family that occur as MLL fusion partners (AF4, AF5q31, LAF4) but also with positive transcription elongation factor b (pTEFb) and the histone methyltransferase DOT1L.

DOT1L methylates histone H3 at lysine 79, a modification that is introduced also during transcriptional elongation (Steger, 2008).

8. RS 4:11 cell line

Normal MLL1 performs an important task necessary for the transcription of many genes. Because all domains within MLL-N thought to be involved in target selection are retained in the fusion proteins, it seems likely that MLL1 fusions will share many target loci with wild type MLL1.

This assumption has been confirmed for the clustered HOX homeobox genes that are under control of MLL1 as well as of MLL1 fusion proteins.

In addition to the HOX cluster, MLL-AF4 has been found on a genome-wide scale on more than 1,000 promoters that also showed a corresponding H3K79 methylation pattern as an indication for a functional interaction of MLL-AF4 with chromatin.

The MLL breakpoints are tightly clustered in an 8.3-kb genomic region containing exons 8 to 14, and the derivative 11 chromosome encodes the MLL fusion protein, which possesses ~1,300 aa of N-terminal MLL1. The C-terminal fusion partners of MLL are very diverse, ranging from putative transcription factors to cytoplasmic structural protein caratterized by the presence in the chimeric protein, of the C-terminal region of MLL1.

The RS4;11 cell line (human B cell precursor), was established from the bone marrow of a patient in relapse with an acute leukemia that was characterized by the t(4;11) chromosomal abnormality and precisely a chromosomal translocation between the fourth and eleventh chromosome.

9. Ash2L and others component of the MLL complex

The MLL/SET1 family consists of six members, Mixed Lineage Leukemia 1 (MLL1), MLL2 (ALR), MLL3 (HALR), MLL4, SET1A and SET1B, which, share a catalytic SET domain that has been shown to have H3K4 methyltransferase activity (Milne, 2002;Wysocka, 2003; Goo,2003; Lee, 2007).

MLL/SET1 proteins exist in multimeric complexes that contain three highly conserved subunits: Ash2l, RbBP5 and WDR5 (Patel, 2009), several biochemical studies in fact indicate that WDR5, RbBP5 and ASH2L proteins are integral members of each complex and could exist in a subcomplex in the absence of MLL.

Furthermore have been found to associate with human SET1, MLL and MLL2 and are predicted to interact with MLL3 and MLL4 (Melissa M Steward, 2006).

Recently, it had been reported that WDR5, RbBP5 and ASH2L are important for regulating the enzymatic activity complexes.

WDR5 is a central component of the MLL complex and is required for ASH2L–RbBP5–MLL complex formation and therefore proper H3K4 methylation.

Loss of ASH2L does not alter MLL core complex stability (MLL–RbBP5–WDR5),

The proteins Ash2l and RbBP5 are important for the enzymatic activity of the complex, the dimer formed by Ash2L and RbBP5 in fact has intrinsic histone methyltransferase activity that require the highly conserved SPRY domain of Ash2L and a short peptide of RbBP5 (Fang, 2007), Ash2L/RbBP5 is able to directly interact with substrates H3.

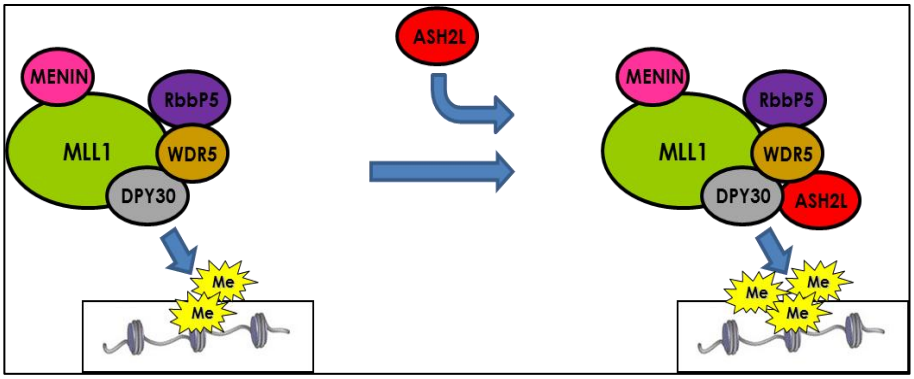


Fig.11 *Loss of ASH2L does not alter MLL core complex stability (MLL–RbBP5–WDR5), but ASH2L is required for proper H3K4 trimethylation by the MLL complex*

Ash2l, in particular, was shown to be critical for H3K4 trimethylation (fig11), the down-regulation of Ash2l leads to a genome-wide decrease in this epigenetic mark (Melissa M Steward, 2006).

PUBBLICATION

NF-Y rectruits Ash2L to impart H3K4 trimethylation on CCAAT promoters.

Andrea Fossati 1, Diletta Dolfini 1, Giacomo Donati 1°, and Roberto Mantovani 1*

Submitted to PLOS ONE	**01 September 2010**
Accepted	**26 Gennuary 2011**
Published	**21 March 2011**

1) Dipartimento di Scienze Biomolecolari e Biotecnologie. Università degli Studi di Milano. Via Celoria 26, 20133 Milano, Italy.

*Corresponding author.
Phone: 39-02-50315005;
Fax: 39-02-50315044;
E-mail: mantor@unimi.it

° Present address: Wellcome Trust Centre for Stem Cell Research, University of Cambridge, UK.

Background.

Different histone post-translational modifications (PTMs) are crucial in the regulation of chromatin, including methylations of H3 at Lysine 4 by the MLL complex. A relevant issue is how this is causally correlated to the binding of specific transcription factors (TFs) in regulatory regions. NF-Y is a TF that regulates 30% of mammalian promoters containing the widespread CCAAT element. We and others established that the presence of H3K4me3 is dependent upon the binding of NF-Y. Here, we investigate the mechanisms of H3K4me3 deposition by NF-Y.

Methods.

We employed Chromatin Immunoprecipitation in cells in which Ash2L and NF-Y subunits were knocked down by RNAi, to monitor the presence of histones PTMs and components of the MLL complex. We performed gene expression profiling of Ash2L-knocked down cells and analyzed the regulated genes. We performed ChIPs in leukemic cells in which MLL1 is devoid of the methyltransferase domain and fused to the AF4 gene.

Results.

Knock down of the Ash2L subunit of MLL leads to a decrease in global H3K4me3 with a concomitant increase in H3K79me2. Knock down of NF-Y subunits prevents promoter association of Ash2L, but not MLL1, nor WDR5, and H3K4me3 drops dramatically. Endogenous NF-Y and Ash2L specifically interact in vivo. Analysis of the promoters of Ash2L regulated genes, identified by transcriptional profiling, suggests that a handful TF binding sites are moderately enriched, among which the CCAAT box. Finally, leukemic cells carrying the MLL-AF4 translocation show a decrease of H3K4me3, absence of Ash2L and increase in H3K79me2, while NF-Y binding was not significantly affected.

Conclusions.

Three types of conclusions are reached: (i) H3K4 methylation is not absolutely required for NF-Y promoter association. (ii) NF-Y acts upstream of H3K4me3 deposition by recruiting Ash2L. (iii) There is a general cross-talk between H3K4me3 and H3K79me2 which is independent from the presence of MLL oncogenic fusions.

Introduction.

Histone post-translational modifications -PTMs- are marks of chromatin environments. Some of them are associated with accessible chromatin, others with heterochromatin, either constitutive or facultative (1). Specifically, monoubiquitination of H2B at Lysine 120 –K123 in yeast- is one of the earliest events in the establishment of an active chromatin environment (2). H3 methylations follow, on H3K4 and H3K79, in a hierarchy of events that leads to gene activation. Methylation of H3K4 is highly regulated (3) and generally present in active and poised promoters (4). The major H3K4 Methyltransferase is MLL (ALL1), the human homologue of the Drosophila Trithorax, assembled in a complex that includes Menin, Ash2L, WDR5, RbBP5, DPY30 and HCFs (5-7). The 4 MLL genes in humans, MLL1-4 contain a Set domain which mono-, di- and tri-methylates H3K4 (8); proteins within the complex are important to impart substrate specificity (9). Specifically, the complex is unable to tri-methylate H3K4 in the absence of Ash2L (6, 7). Moreover, MLL1 is involved in chromosomal translocations with a large cohort of >50 loci in aggressive myeloid and lymphoid leukemias (10).

In general, a key question is how histone modifying complexes are selectively and timely recruited to promoters, and binding of sequence-specific transcription factors offers a convenient explanation (11-16). NF-Y is a trimer composed of NF-YA, NF-YB and NF-YC (17), which regulates the CCAAT box, one of the most frequent and crucial promoter elements (18). A connection between NF-Y binding and H3K4 methylations was initially noticed on the promoters of the ER-stress response genes, prior to induction (19, 20). This was confirmed in genome-wide correlative ChIP on chip studies, since NF-Y and H3K4me3 locations overlapped significantly and correlated with expression (21). In cause-effect experiments, we and others noticed a parallel decrease in NF-Y binding, H3K4me3, H3K79me2 and transcription using a dominant negative NF-YA mutant (22-24). The reverse was not tested, namely whether H3K4me3 is important for NF-Y promoter association. This is a relevant point, since not all TFs are apparently equal in this regard: in a detailed correlative analysis, MYC binding was always associated with a specific context of histone marks, notably H3K4me3 and H3K79me2, but its removal –and comparison between myc+/+ and myc-/- cells- left the H3K4 pattern intact at target sites, whereas H4 acetylations

were substantially ablated. Therefore, it was concluded that these marks are required for MYC binding to E boxes (25). In a previous study focusing on cell cycle regulated promoters in single nucleosome ChIP assays, we established that H3K4 di-methylation is unaffected by NF-Y binding, which is instead involved in the transition to mono- and tri-methylation upon gene activation. We then focused on H3K4me1: NF-Y promotes it by recruiting the CoREST-KDM1 H3K4me2 demethylase complex through contacts between NF-Y and CoREST (24). Here, we report studies on the deposition of the H3K4me3 mark on CCAAT promoters. In particular, we first tested the hypothesis that H3K4me3 might be generally helpful in NF-Y promoter association, by eliminating Ash2L, the one subunit of the complex that is specifically required for the deposition of this mark.

RESULTS.

Knock down of Ash2L leads to decrease in H3K4me3, increase in H3K79me2 and selective reduction of NF-Y binding.

To study the role of H3K4me3 in NF-Y promoter association, we knocked down Ash2L by siRNA in HCT116 cells: Figure 1A shows that a substantial reduction of Ash2L –down to 30% of normal levels- could be achieved, while other subunits of MLL complexes, Menin and WDR5, were, if anything, increased (Fig. 1A).

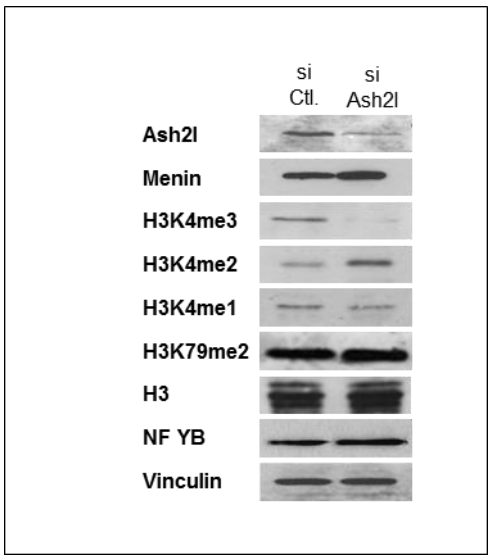


Fig. 1A. Effects of Ash2L knock down on H3K4me3, H3K79me2 and NF-Y binding.

A. Knock down of Ash2L in HCT116 human cells. Western blot analysis of the indicated proteins and histone PTMs in cells transfected with scramble and Ash2L siRNAs. On the right, the levels of Ash2L protein inactivation, based on three independent experiments.

Global levels of H3K4 methylations were controlled by Western blot analysis: H3K4me3 reduction by Ash2L siRNA was matched by an increase of H3K4me2, while H3K4me1 was unchanged (Fig. 1A, Lower Panels).

Next, we performed ChIP experiments with chromatin of HCT116 cells treated with control and Ash2L siRNAs, using antibodies against NF-Y, H3K4me3, H3K4me2, H3K79me2 and Ash2L; we analyzed a few promoters that are representatives of the different classes of CCAAT promoters -housekeeping, cell cycle and ER-stress- as well as of promoters devoid a of the CCAAT box -CCAAT-less- serving as controls. In parallel, we analyzed the transcriptional profile of the genes considered.

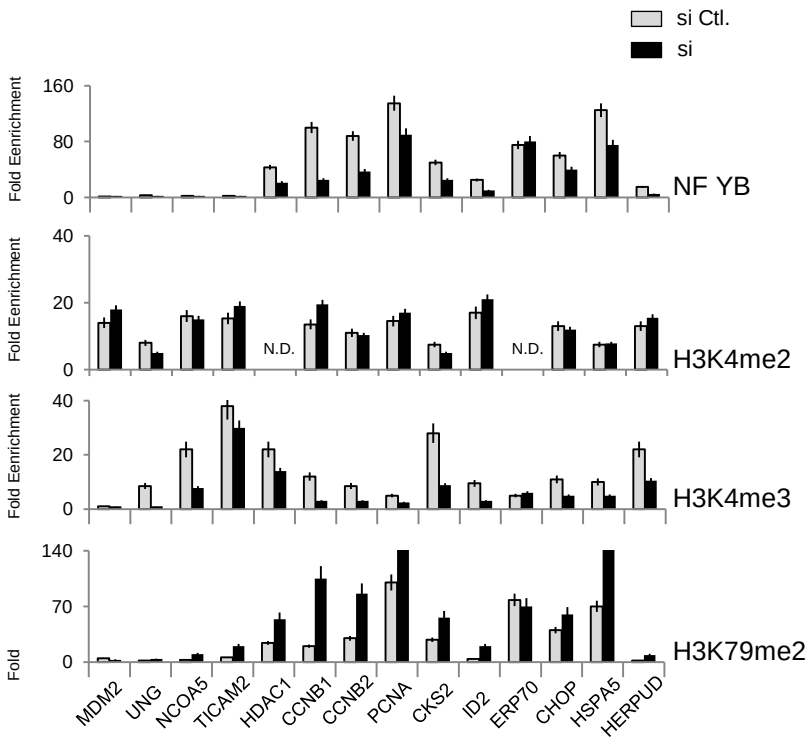


Figure 1B. ***Effects of Ash2L knock down on H3K4me3, H3K79me2 and NF-Y binding.***

A. Knock down of Ash2L in HCT116 human cells. Western blot analysis of the indicated proteins and histone PTMs in cells transfected with scramble and Ash2L siRNAs. On the right, the levels of Ash2L protein inactivation, based on three independent experiments. B. In the upper Panel, mRNAs levels of the indicated genes in HCT116 cells transfected with scramble (Grey bars) or Ash2L siRNAs (Black bars) were assessed by qRT-PCR. In the lower Panels, Chromatin immunoprecipitation (ChIP) analysis was performed with the indicated antibodies (NF-YB, H3K4me2, H3K4me3, H3K79me2) in the same cells transfected with scramble or Ash2L siRNA.

qPCR analysis was performed with primers centered in the core promoters of the indicated genes. CCAAT-less refers to genes with no CCAAT box in the promoters. Cell cycle and ER-stress are two categories of promoters with functionally important CCAAT boxes.

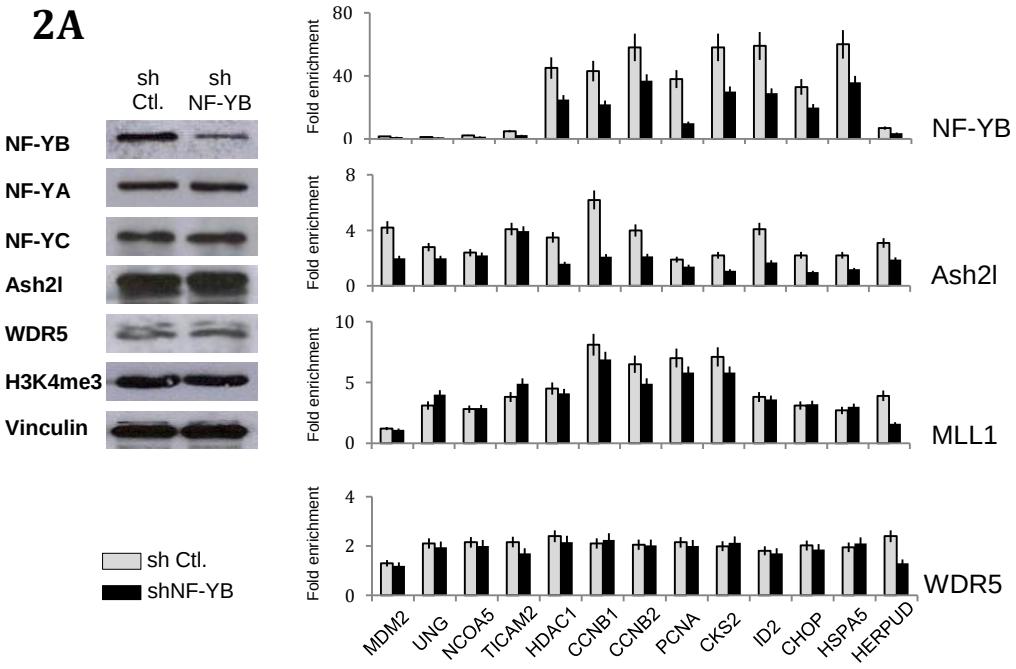
Fig. 1B shows a two to four-fold decrease of H3K4me3 on most promoters, with the exception of TICAM2 and ERP70.

Interestingly, H3K79me2 levels were concomitantly increased (2 to 3-fold): indeed, the greater the loss of H3K4me3, the higher the increase of H3K79me2.

NF-Y binding was decreased on some promoters -ERP70, CHOP, PCNA but not abolished. In most other promoters, however, the decrease in NF-Y binding was marginal. We conclude that the presence of H3K4me3 moderately influences NF-Y binding only on some promoters, and that indirect removal of H3K4me3 by inactivation of Ash2L leads to a compensatory positive effect on H3K79me2.

NF-Y recruits Ash2L on CCAAT-containing promoters.

Removal of the NF-Y trimer from cell cycle promoters by an NF-YA dominant negative mutant led to a substantial decrease of H3K4 tri-methylation (22-24), which is dependent from Ash2L, but not H3K4 di-methylation. One possibility is that NF-Y is important specifically for Ash2L recruitment. To establish this point, we decided to use an alternative system: knock down of the NF-YB (Fig. 2A) and of NF-YA (Fig. 2B) subunits in HCT116 cells by shRNA interference.



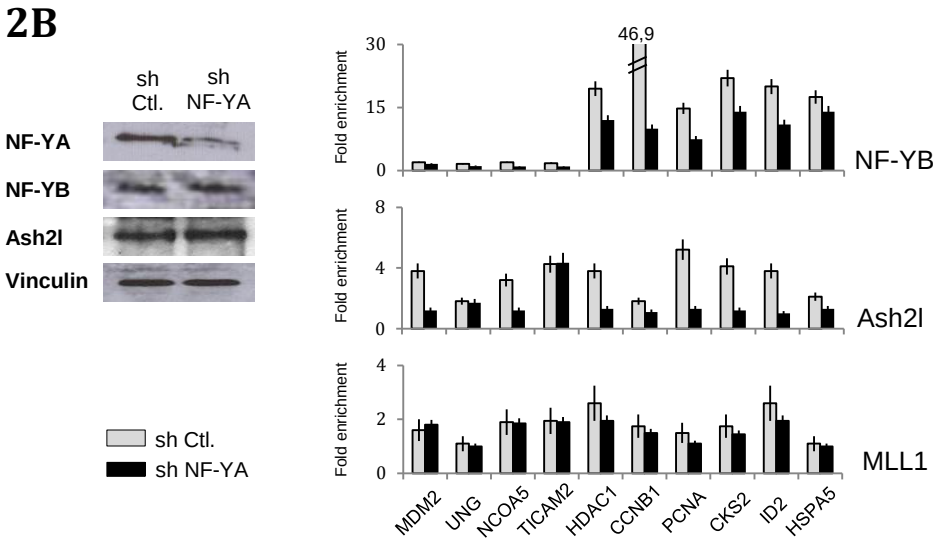


Figure 2. *NF-Y recruits Ash2L on CCAAT-containing promoters.*

A*. Knock down of the NF-YB subunit in HCT116 human cells by infection with control (GFP) or NF-YB shRNA producing Lentiviruses. Left Panels: Western blot analysis of the indicated proteins and histone PTMs, with a statistical evaluation of the degree of NF-YB protein knock down, measured based on three independent experiments. Right Panels: in the upper Panel, mRNAs levels of the indicated genes in HCT116 cells transfected with control GFP (Grey bars), or NF-YB shRNAs (Black bars) assessed by qRT-PCR. In the lower Panels, Chromatin immunoprecipitation (ChIP) analysis was performed with the indicated antibodies (NF-YB, H3K4me3, Ash2L, MLL1, WDR5) in the same cells. qPCR analysis was performed with primers centered in the core promoters of the genes. The same set of promoters of Figure 1 were analyzed.* ***B****. Same as A, except that cells were knocked down with shRNA for NF-YA.*

Western blot analysis of NF-Y subunits confirmed the selective reduction of the respective NF-Y subunits: to 30% for NF-YB and 27% for NF-YA. Ash2L, as well as the global levels of H3K4me3 were not affected. ChIP assays on the NF-Y-dependent promoters analyzed in Fig. 1 showed a parallel reduction of NF-YB, H3K4me3 and Ash2L, whereas the presence of MLL1 and WDR5 was unaffected. On the control CCAAT-less UNG, COA5 and TICAM2 promoters, no NF-Y binding and robust Ash2L association was evident (Fig. 2A and 2B); MDM2, instead, showed a drop in Ash2L: a possible interpretation is that an important NF-Y site in a distant region impacts on promoter organization affecting H3K4me3.

To ascertain whether there is a NF-Y-Ash2L interaction in vivo, we performed immunoprecipitations with NF-Y and Ash2L antibodies with HCT116 nuclear extracts, followed by Western blot analysis. Fig. 3 shows that Ash2L is immunoprecipitated with anti-NF-YB antibodies, whereas minimal levels of Menin and WDR5 were present in the bound fraction.

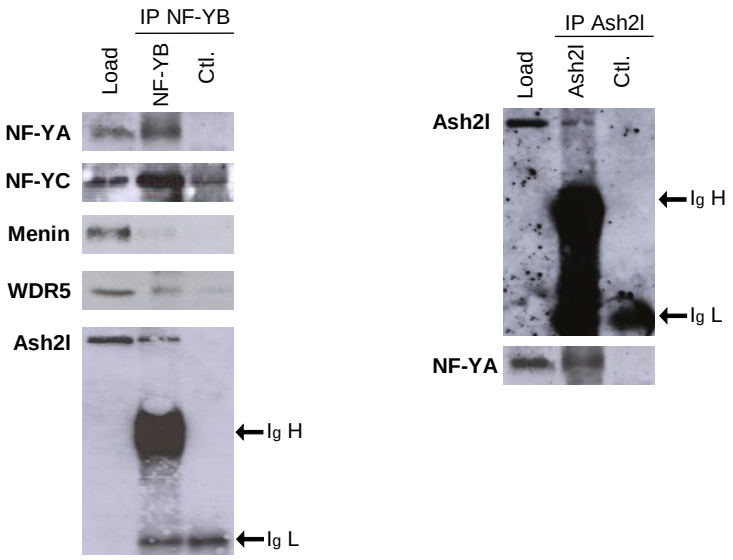


Figure 3. <u>Direct interaction between NF-Y and Ash2L.</u>

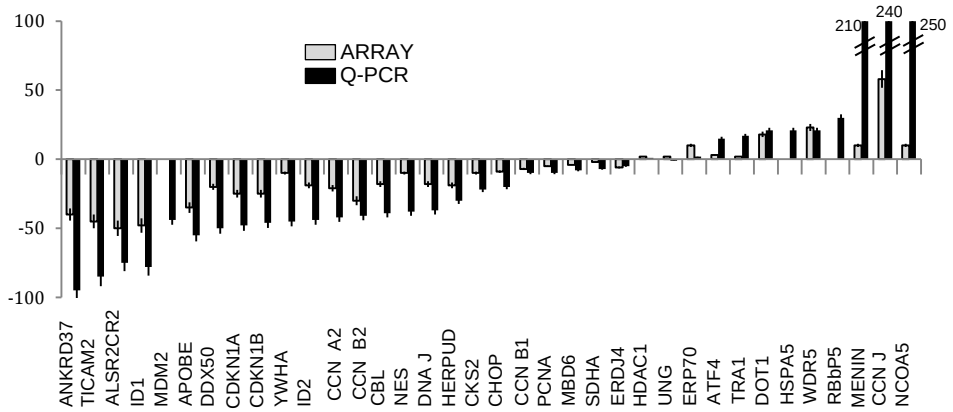
Western blot analysis of immunprecipitations with anti-NF-YB (Left Panel) and anti-Ash2L (Right Panel) antibodies, using acid-extracted nuclear extracts of HCT116 cells. Control IPs with anti-Flag antibodies were run in parallel. The Load lane refers to the starting nuclear extract material. Menin and WDR5 were marginally immunoprcipitated with anti-NF-YB, whereas Ash2L recovery was robust. The heavy (IgH) and light (IgL) chains of the IP antibodies are indicated.

The reciprocal was also true, as substantial amounts of NF-YA was in the Ash2L IP. Control IPs were negative for all proteins tested. Taken together, these data indicate that the recruitment of Ash2L is dependent upon NF-Y binding, that there is a NF-Y-Ash2L interaction in vivo, and that other subunits of the MLL complex are recruited independently from Ash2L.

To ascertain whether Ash2L is recruited preferentially on CCAAT promoters, we performed expression profiling of HCT116 cells after knock down of Ash2L by

siRNA, with the protocol shown above in Fig. 1. By setting the threshold at 1.35, 477 genes were down-regulated and 175 up-regulated (Fig. S1). We validated by qRT-PCR 32 of these genes, and the adherence to the gene expression data was almost complete (Fig. 4A).

A



B

P-Value < 10^{-6}

GO.ID	Biological process	Tot	Enr	p.value
GO:0006396	RNA processing	482	43	9.68E-12
GO:0006259	DNA metabolic process	474	42	2.04E-11
GO:0008380	RNA splicing	268	29	1.55E-10
GO:0016071	mRNA metabolic process	335	32	5.53E-10
GO:0006974	response to DNA damage stimulus	315	29	7.34E-09
GO:0006397	mRNA processing	298	27	3.09E-08
GO:0016192	vesicle-mediated transport	485	35	1.57E-07
GO:0006605	protein targeting	186	19	3.35E-07
GO:0051170	nuclear import	84	12	6.23E-07
GO:0009057	macromolecule catabolic process	454	32	8.18E-07
GO:0043285	Biopolymer catabolic process	372	28	9.30E-07
GO:0006281	DNA repair	257	22	1.21E-06
GO:0006913	nucleocytoplasmic transport	136	15	1.50E-06
GO:0051169	nuclear transport	138	15	1.82E-06
GO:0000375	RNA splicing, via transesterification reactions	148	15	4.58E-06
GO:0000398	nuclear mRNA splicing, via spliceosome	148	15	4.58E-06
GO:0006886	intracellular protein transport	383	27	4.91E-06
GO:0044419	interspecies interaction between organisms	231	19	9.66E-06

GO.ID	Cellular component	Tot	Enr	p.value
GO:0010008	endosome membrane	116	14	9.46E-07
GO:0044440	endosomal part	117	14	1.06E-06
GO:0005768	endosome	227	20	2.05E-06
GO:0005730	nucleolus	153	15	7.04E-06

GO.ID	Molecular function	Tot	Enr	p.value
GO:0005525	GTP binding	351	26	2.81E-06
GO:0019001	guanyl nucleotide binding	360	26	4.51E-06
GO:0032561	guanyl ribonucleotide binding	360	26	4.51E-06
GO:0022613	ribonucleoprotein complex biogenesis and assembly	210	18	8.55E-06

Figure 4 <u>Validation of Ash2L profiling experiments</u>.

A. Validation of the profiling data for selected genes whose expression was activated or repressed by Ash2L inactivation, as in the experiments shown in Figure 1, was performed by qRT-PCR. B. Gene Ontology analysis of the profiling experiments of Ash2L-dependent genes

The overall greater changes observed by qRT-PCR analysis suggest that additional genes below the threshold considered are affected by Ash2L interference. Analysis of Gene Ontology terms indicates that Ash2L targets different classes of genes, with DNA and RNA Metabolism being dominant in the Biological Process and Molecular Function categories (Fig. 4B). We then analyzed the regulated promoters to identify enriched Transcription Factor Binding Site -TFBS- (Figure S2): CCAAT was indeed at a top of a short list of sites, together with a few other, but the p values were relatively modest, an indication that there is no strong skewing toward a specific TFBS. We conclude that Ash2L recruitment on CCAAT promoters is NF-Y-dependent, but that the function of this MLL subunit is not restricted to CCAAT promoters.

Ash2L promoters association is affected in Mixed Lineage Leukemia cells.

The results shown above indicate that NF-Y lies upstream of Ash2L recruitment and H3K4me3 deposition. Previous data on cell cycle genes indicate that H3K4me2 is present independently from NF-Y, raising the possibility that this mark is actually required for NF-Y recruitment. To further evaluate this point, we used the cell line RS4-11, derived from a Mixed Lineage Leukemia, which contains a rearrangement of the MLL1 gene with AF4: as in other MLL1 fusions, it lacks the SET domain and it is therefore devoid of H3K4 methylating activity (10). Note that in MLL-AF10 regulated genes, this is compensated by higher levels of H3K79me2 (27, 28), which is also under the control of NF-Y (22, 24). Therefore, we first evaluated the global levels of the proteins and histone PTMs considered in this study in RS4-11, showing that they were largely similar to what is found in REH, a control leukemia cell line that harbours normal MLL1 alleles (Fig. 5A). We then analyzed by ChIPs NF-Y, Ash2L and H3K4 methylations on CCAAT promoters, in RS4-11 and REH. The levels of H3K4me3 were lower in RS4-11 compared to REH on most, but not all promoters (Fig. 5B). Higher levels of H3K79me2 were found, specifically in the promoters with low levels H3K4me3, which is consistent with previous results with MLL-AF10 fusions (26). The H3K4me1 and H3K4me2 levels were extremely low in RS4-11, with the exception of CKS2 and Cyclin B1 promoters. Importantly, the recovery of Ash2L were minimal from all promoters, while NF-Y binding was comparable in

the two cell lines, including in promoters with residual levels of H3K4 methylations. We take this as a further indication that H3K4 methylations are not strictly required for NF-Y association. In addition, it is clear that the presence of the oncogenic fusion protein affects the levels of H3K4 mono-, di- and tri-methylation.

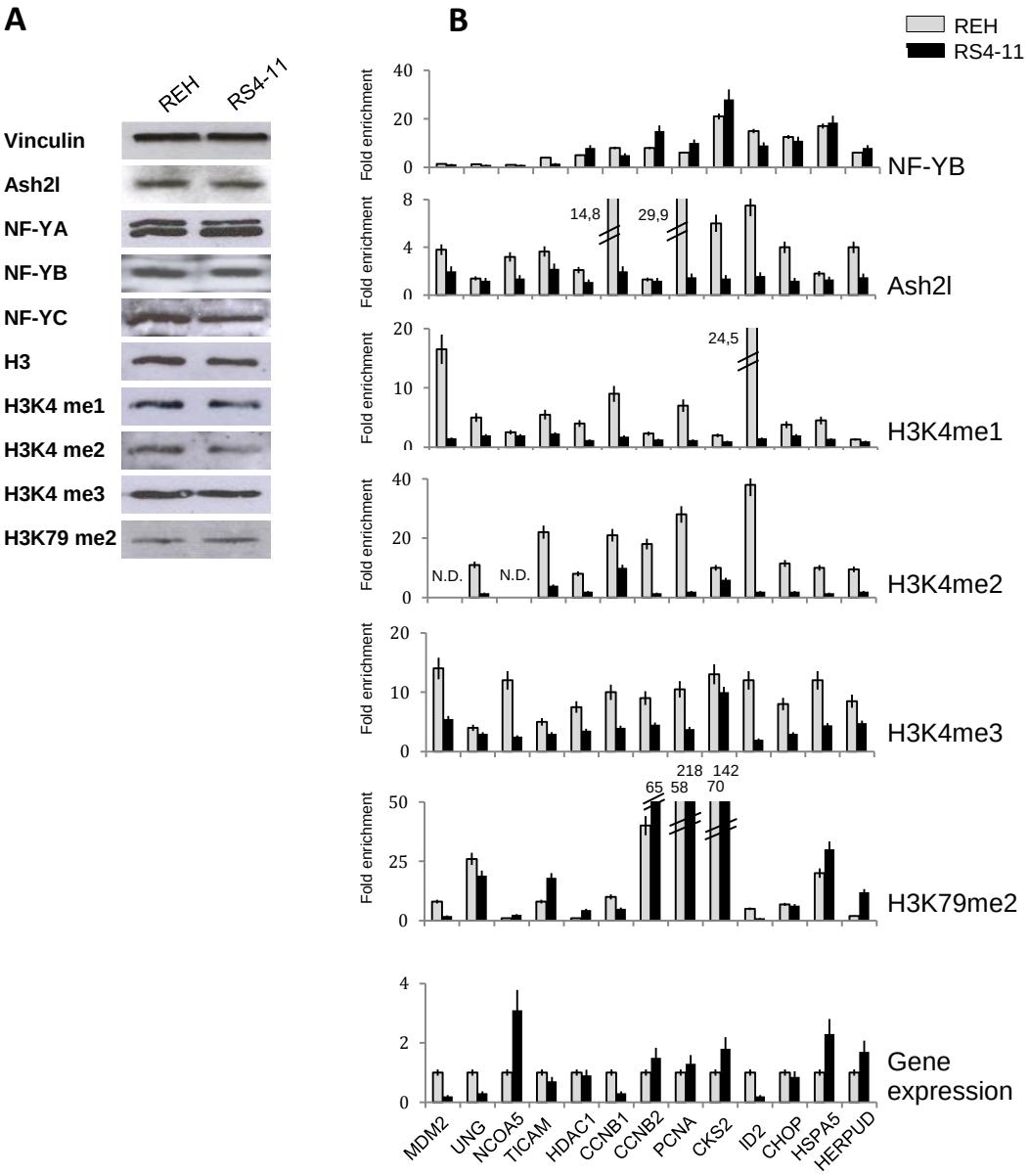


Figure 5. <u>MLL cells lack Ash2L promoter association.</u>

***A**. Western blot analysis of the levels of the indicated proteins (Ash2L, NF-Y subunits, global H3, H3K4me1, H3K4me2, H3K4me3, H3K79me2, and the internal control Vinculin) in MLL1-AF4 rearranged cells RS4-11 and in non MLL1 rearranged REH leukemic cells. **B**. In the upper Panel,*

mRNAs levels of the indicated genes in REH (Grey bars) and RS4-11 (Black bars) assessed by qRT-PCR. In the lower Panels, Chromatin immunoprecipitation (ChIP) analysis was performed with the indicated antibodies (NF-YB, Ash2L, H3K4mei, H3K4me2, H3K4me3, H3K79me2) in the same cells. qPCR analysis was performed with primers centered in the core promoters of the genes. The same set of promoters of Figure 1 were analyzed.

DISCUSSION.

NF-Y and H3K4 methylations.

A relevant question in chromatin studies is what determines the location of histone marks on genomes. The binding of TFs and cofactors to promoters are hallmarks of expression, by signalling to the Pol II machinery the appropriate positional coordinates. In general, it seems plausible that TFs are instrumental in determining the positions of specific histone PTMs. However, it is becoming increasingly clear that certain TFs do require a particular set of histone TFs to access regulatory regions: MYC binding, for example, happens only in sites in which H3K4 and H3K79 methylations are abundant (25). Hence, there must be a hierarchy in TF in establishing a certain chromatin environment, with some TFs being capable to recruit histone modifying enzymes for the benefit of additional TFs. There is a strong correlation in genomic locations of NF-Y and H3K4me3, and removal of NF-Y reduced the local, but not the global levels of H3K4me3 (19-24). Unlike other TFs, the presence of high levels of H3K4me3 is not strictly required for NF-Y recruitment: indeed, NF-Y is involved in the recruitment of Ash2L, indicating that it acts upstream of H3K4me3, and placing it at the heart of the local di- to tri-methylation transition. Furthermore, the results in MLL cells, in which most promoters we analyzed have residual levels of any H3K4 methylation, yet robust NF-Y association, indicate that none of the methylations of H3K4 is strictly required for binding of the trimer to DNA. The fact that H3K4 methylations are downstream of H2B monoubiquitination (2 and References therein) renders the H2B-like structure of NF-YB particularly relevant, since a parallel signalling can be envisaged.

The presence of the CCAAT box at the top of the list of the Ash2L regulated genes, together with few other TFBS with similar scores, is in line with the data. However, the relatively low statistical enrichments in TFBS analysis among

Ash2L-regulated promoters indicate that the preference is not absolute, and many other TFs can recruit Ash2L, as indicated by genetic screenings for interactors of AP2, Mef2c, PAX7 and Tbx1 (13-16). We confirm that removal of Ash2L leads to a drastic decrease in the global levels of H3K4me3, while H3K4me2 and H3K4me1 are not substantially changed. Whether Ash2L is the only subunit of the complex to direct the MLL activity toward trimethylation is currently unclear (9). As Ash2L is believed to be present in complexes containing different MLLs, we were surprised by the relatively low number of genes affected by Ash2L interference. There are technical explanations for this, such as the incomplete elimination of the protein and the fact that hybridization-derived profiling data are less sensitive to variation with respect to qRT-PCR, suggesting that many additional genes were indeed missed. In addition, the presence of partially redundant activities similar to Ash2L, such as Ash1 (29), should be considered.

Reciprocal regulation of H3K4me3 and H3K79me2.

MLL is a complex and genetically heterogeneous disease in which MLL1 fusion proteins alter gene expression; this is, in part, due to the partners, some of which - AF4, AF9 and ENL- have transcription activation domains required for transformation (10, 27). In cells bearing AF10 fusions, reduced H3K4me3 was reported to be "compensated" by high H3K79me2 levels. AF10 was shown to bind to hDOT1L, through a domain required for transformation (28). In accordance with these data, in our analysis of leukemic cells with an MLL1-AF4 fusion, the H3K4me3 decrease is also "compensated" by an increase in H3K79me2. However, in HCT116 cells, which carry a wt MLL1 configuration, the removal of Ash2L alone, by decreasing H3K4me3, is sufficient to increase H3K79me2, and there is an inverse correlation between these two marks on each of the promoters we analyzed. One could therefore hypothesize that it is the absence of Ash2L on promoters, rather than, or in addition to an AF4 or AF10-mediated recruitment of hDOT1L, which leads to high H3K79me2 levels. This finding could have consequences in the temporal deposition of the two marks, which lie downstream of H2BUb, suggesting that H3K79me2 acts upstream of H3K4me3. This matter is further complicated by the fact that mono- and di-methylations are also dramatically affected in MLL cells, which was somewhat expected, based on the assumption that the SET domain is absent from the MLL

fusions, but never tested. Importantly, in the absence of the N-terminal end of MLL1, Ash2L is not recruited onto promoters. These data indicate that the oncogenic potential might be influenced by the variation of composition of the MLL complex as a whole, as well as by the presence of the fusion partner.

A link between the MLL complex and NF-Y in cellular transformation?

In 23 acute lymphoblastic leukemias with MLL translocations, two signatures correlating with prognosis were found (30). Top rank genes were HspCBF, an NF-Y coactivator (31) in the poor prognosis group, and CDP, a negative regulator of CCAAT activity (32), in the cohort with good prognosis. Importantly, genes with two CCAAT boxes, predicted to be down-regulated by CDP, were found underexpressed in the latter group. (ii) The CCAAT box was repeatedly reported, along with E2F sites, in genes specifically overexpressed in tumors (33-35). Notably, de novo motifs discovery in leukemias pointed at three sites: the expected E2F-NF-Y duo and p53 (36). (iii) Analysis of gene expression profiles of MLL-AF9 leukemias identified a haematopoietic stem cell -HSCs- signature that confers self-renewal properties (10), with some targets, such as HOXA members, important for tumor growth. Interestingly, NF-Y was shown to be a potent HSC self-renewal regulator, by activating HOX4 paralogues (37). Notably, the short form of NF-YA appears to be crucial (38, 39). Thus, MLL fusions and NF-Y could work together in reactivation of a self-renewal program. The absence of significant levels of H3K4 methylations is coupled to high levels of H3K79me2, a mark also dependent from the presence of NF-Y (24): we can imagine that NF-Y is also involved in the recruitment of the hDOT1 complex, even without the presence of the AF4/AF10 fusion partner. NF-Y-mediated recruitment of MLL complexes with an altered composition on growth-promoting genes could impair profoundly regulation, via alteration of the local histone PTMs. Further biochemical and in vivo ChIP work is required to shed light on this hypothesis.

ACKNOWLEDGEMENTS.

We thank E. Canaani (Weizman I., Il), G. Cazzaniga, A. Biondi (F. Tettamanti, I) for gift of MLL reagents and cell lines, C. Imbriano (U. of Modena, I) for gift of shRNA vectors.

MATERIALS AND METHODS.

Cell cultures and transfections. HCT116 cells were cultured in DMEM supplemented with 10% FCS, 1% penicillin and streptomycin, and L-glutamine. All transfections were carried out using Lipofectamine 2000 (Invitrogen, USA). The Ash2L siRNA was purchased by (Dharmacon, USA). Scrambled siRNA was used as a negative control (Ambion, USA). NF-YA and NF-YB shRNA vectors were purchased from Sigma; the control was a similar vector expressing shRNA against GFP.

RT-PCR analysis, Nuclear and Acid extracts preparation and Western blot analysis. Total RNAs were extracted using an RNA-Easy kit (Qiagen, D). 1µg of each RNA was retrotranscribed (Promega, USA). Normalization of the cDNAs were performed with GAPDH control. The RT-PCR primers used are listed in Supplementary 3.

Nuclear extracts were prepared according to standard procedures (20). Acid extracts were prepared collecting cells in 5-10 volumes of Lysis Buffer H (10mM Hepes pH 7.9, 150mM MgCl2, 10mM KCl, 0.5 mM DTT, 1.5 mM PMSF); Perchloric acid 0.2 M was added and cells were kept on ice for 30 minutes, followed by centrifugation for 10 minutes at 11000g at 4°C; the supernatants were stocked at -80°C. 15 µg of nuclear and acid extracts were used in 12% SDS-PAGE. Proteins were transferred to nitrocellulose membranes and immunoblotted using the antibodies of interest. The protein-antibody complexes were detected using horseradish peroxidase-conjugated secondary antibodies (GE Healthcare, UK) and the chemioluminescence system (Genespin, I).

Chromatin Immunoprecipitation. ChIP assays were performed as previously described (Donati et al, 2007).

Immunoprecipitations were performed with ProtG-Sepharose (KPL, USA) and 3µg of the following antibodies: NF-YB (Genespin, I); H3 (Abcam 1791); H4K4me3 (Abcam 8580, Active Motif 39159); H4K4me2 (Abcam 7766); H4K4me1 (Abcam 8895); H3K79me2 (Abcam 3594), Ash2L (Active Motif 39099).

The MLL antibody was a kind gift of E. Canaani (Weizman I., Il).

The ChIP-PCR primers are listed in Figure S3. Quantitative Real Time PCR was performed using SYBR green IQ reagent (Biorad, USA) in the iCycler IQ. The relative sample enrichment was calculated with the following formula: 2 ΔCtx -2 ΔCtb, where ΔCt x = Ct input-Ct sample and ΔCt b = Ct input-Ct control Ab.

REFERENCES.

1) Berger SL. (2007). The complex language of chromatin regulation during transcription. Nature 447:407-412.
2) Nakanishi S, Lee JS, Gardner KE, Gardner JM, Takahashi YH et al. (2009). Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. J Cell Biol. 186:371-317.
3) Ruthenburg AJ, Allis CD, Wysocka, J. (2007). Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. Mol. Cell 25:15-30.
4) Barski A, Cuddapah S, Cui K, Roh TY, Schones DE et al. (2007). High-resolution profiling of histone methylations in the human genome. Cell 129:823-837.
5) Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T et al. (2002). ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol Cell 10:1119-1128.
6) Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW et al. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nat Struct Mol Biol. 13:713-9.
7) Steward MM, Lee JS, O'Donovan A, Wyatt M, Bernstein BE et al., (2006). Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. Nat Struct Mol Biol. 13:852-854.
8) Wang P, Lin C, Smith ER, Guo H, Sanderson BW et al. (2009). Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. Mol Cell Biol. 29:6074-6085.

9) Patel A, Dharmarajan V, Vought VE, Cosgrove MS. (2009). On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. J Biol Chem. 284:24242-4256.

10) Krivtsov AV, Armstrong SA. (2007). MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer 7:823-833.

11) Wynder C, Hakimi MA, Epstein JA, Shilatifard A, Shiekhattar R. (2005). Recruitment of MLL by HMG-domain protein iBRAF promotes neural differentiation. Nat Cell Biol. 7:1113-1117.

12) Dreijerink KM, Mulder KW, Winkler GS, Höppener JW, Lips CJ, et al. (2006). Menin links estrogen receptor activation to histone H3K4 trimethylation. Cancer Res. 66:4929-4935.

13) McKinnell IW, Ishibashi J, Le Grand F, Punch VG, Addicks GC et al. (2008). Pax7 activates myogenic genes by recruitment of a histone methyltransferase complex. Nat Cell Biol. 10:77-84.

14) Rampalli S, Li L, Mak E, Ge K, Brand M et al., (2007). p38 MAPK signaling regulates recruitment of Ash2L-containing methyltransferase complexes to specific genes during differentiation. Nat Struct Mol Biol. 14:1150-1156.

15) Tan CC, Walsh MJ, Gelb BD. (2009). Fgfr3 is a transcriptional target of Ap2delta and Ash2l-containing histone methyltransferase complexes. PLoS One 4:e8535.

16) Stoller JZ, Huang L, Tan CC, Huang F, Zhou DD et al. (2010). AshL interacts with Tbx1 and is required during early embryogenesis. Exp Biol Med. 235:569-576.

17) Romier C, Cocchiarella F, Mantovani R, Moras D. (2003). The crystal structure of the NF-YB/NF-YC heterodimer gives insight into transcription regulation and DNA binding and bending by transcription factor NF-Y. J. Biol. Chem. 278:1336-1345.

18) Dolfini D., Zambelli F., Pavesi G. Mantovani R. (2009). A perspective of promoter architecture from the CCAAT box. Cell Cycle 8:4127-4237

19) Baumeister P, Luo S, Skarnes WC, Sui G, Seto E, Shi Y. Lee AS. (2005) Endoplasmic reticulum stress induction of the Grp78/BiP promoter: activating mechanisms mediated by YY1 and its interactive chromatin modifiers. *Mol Cell Biol.* **25**: 4529-40.

20) Donati, G., Imbriano, C., and Mantovani, R. (2006). Dynamic recruitment of transcription factors and epigenetic changes on the ER stress response gene promoters. Nucleic Acids Res. *34*, 3116-3127.

21) Ceribelli M, Dolfini D, Merico D, Gatta R, Viganò MA et al. (2008). The histone like NF-Y is a bifunctional transcription factor. Mol. Cell. Biol. 28:2047-2058.

22) Donati G, Gatta R, Dolfini D, Fossati A, Ceribelli M et al. (2008). An NF-Y-dependent switch of positive and negative histone methyl marks on CCAAT promoters. PLoS ONE 3:e2066

23) Gurtner A, Fuschi P, Magi F, Colussi C, Gaetano C et al. (2008). NF-Y dependent epigenetic modifications discriminate between proliferating and postmitotic tissue. PLoS One 3:e2047.

24) Gatta R. and Mantovani R. (2008). NF-Y substitutes H2A-H2B on active cell cycle promoters: recruitment of CoRest-KDM1 and fine tuning of H3 methylations. Nucleic Acids Res. 36:6592-6607.

25) Guccione, E., Martinato, F., Finocchiaro, G., Luzi, L., Tizzoni, L., Dall' Olio, V., Zardo, G., Nervi, C., Bernard, L., and Amati, B. (2006). Myc-binding-site recognition in the human genome is determined by chromatin context. Nat. Cell Biol. *8*, 764-770.

26) Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S et al. (2008). H3K79 methylation profiles define murine and human MLL-AF4 leukemias. Cancer Cell 14:355-368.

27) Okada Y, Feng Q, Lin Y, Jiang Q, Li Y et al. (2005). hDOT1L links histone methylation to leukemogenesis. Cell 121:167-178.

28) Neff T, Armstrong SA. (2009). Chromatin maps, histone modifications and leukemia. Leukemia 23:1243-51.

29) Gregory GD, Vakoc CR, Rozovskaia T, Zheng X, Patel S, Nakamura T, Canaani E, Blobel GA. (2007). Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. Mol. Cell. Biol. 27:8466-79.

30) Tsutsumi S, Taketani T, Nishimura K, Ge X, Taki T et al. (2003). Two distinct gene expression signatures in pediatric acute lymphoblastic leukemia with MLL rearrangements. Cancer Res. 63:4882-4887.

31) Imbriano C, Bolognese F, Gurtner A, Piaggio G, Mantovani R. (2001). HSP-CBF is an NF-Y-dependent coactivator of heat shock promoters CCAAT boxes. J. Biol. Chem. 276, 26332-26339.

32) Sansregret L, Nepveu A. (2008). The multiple roles of CUX1: insights from mouse models and cell-based assays. Gene 412:84-94.

33) Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Barrette TR, Ghosh D et al. (2005). Mining for regulatory programs in the cancer transcriptome. Nat Genet. 37:579-83.

34) Tabach Y, Milyavsky M, Shats I, Brosh R, Zuk O et al. (2005).e promoters of human cell cycle genes integrate signals from two tumor suppressive pathways during cellular transformation. Mol Syst Biol. 1:2005.0022.

35) Sinha S, Adler AS, Field Y, Chang HY, Segal E. (2008). Systematic functional characterization of cis-regulatory motifs in human core promoters. Genome Res. 18:477-488.

36) Goodarzi H, Elemento O, Tavazoie S. (2009). Revealing global regulatory perturbations across human cancers. Mol Cell 36:900-911.

37) Zhu J, Zhang Y, Joe GJ, Pompetti R, Emerson SG. (2005). NF-Ya activates multiple hematopoietic stem cell (HSC) regulatory genes and promotes HSC self-renewal. Proc Natl Acad Sci USA. 102:11728-11733.

38) van Wageningen S, Nikoloski G, Vierwinden G, Knops R, van der Reijden BA et al. (2008). The transcription factor nuclear factor Y regulates the proliferation of myeloid progenitor cells. Haematologica 93:1580-1582.

39) Domashenko AD, Danet-Desnoyers G, Aron A, Carroll MP, Emerson SG. (2010). TAT-mediated transduction of NF-Ya peptide induces the ex vivo proliferation and engraftment potential of human hematopoietic progenitor cells. Blood, in press.

SUPPLEMENTARY

Figure S1. List of Ash2L-regulated genes in HCT116 cells.

Genes whose expression is altered upon Ash2L knock down.

A. Genes downregulated after ASH2L silencing

SYMBOL	ILMN_GENE	CHROMOSOME	DEFINITION
ABCA5	ILMN_3630	17	ATP-binding cassette, sub-family A (ABC1), member 5 (ABCA5).
ABHD2	ILMN_24898	15	abhydrolase domain containing 2 (ABHD2).
ACAD8	ILMN_19744	11	acyl-Coenzyme A dehydrogenase family, member 8 (ACAD8).
ACSBG2	ILMN_138348	19	acyl-CoA synthetase bubblegum family member 2 (ACSBG2).
ACSL4	ILMN_177778	X	acyl-CoA synthetase long-chain family member 4 (ACSL4).
ACSS2	ILMN_9910	20	acyl-CoA synthetase short-chain family member 2 (ACSS2).
ACTA2	ILMN_6588	12	actin, alpha 2, smooth muscle, aorta (ACTA2).
ADAM19	ILMN_12727	5	ADAM metallopeptidase domain 19 (meltrin beta) (ADAM19).
ADAMTS14	ILMN_176528	10	ADAM metallopeptidase with thrombospondin type 1 motif, 14 (ADAMTS14).
ADK	ILMN_29192	10	adenosine kinase (ADK), transcript variant ADK-long.
ADPRHL1	ILMN_2061	13	ADP-ribosylhydrolase like 1 (ADPRHL1).
AFAP1L2	ILMN_23732	10	actin filament associated protein 1-like 2 (AFAP1L2).
AKAP11	ILMN_13368	13	A kinase (PRKA) anchor protein 11 (AKAP11).
AKR1B10	ILMN_6435	7	aldo-keto reductase family 1, member B10 (aldose reductase) (AKR1B10).
ALAS1	ILMN_20926	3	aminolevulinate, delta-, synthase 1 (ALAS1).
ALDH1A2	ILMN_17630	15	aldehyde dehydrogenase 1 family, member A2 (ALDH1A2).
ALS2CR14	ILMN_947	2	amyotrophic lateral sclerosis 2 (juvenile).
ALS2CR2	ILMN_24336	19	amyotrophic lateral sclerosis 2 (juvenile).
AMPD3	ILMN_13634	11	adenosine monophosphate deaminase (isoform E) (AMPD3).
ANGPTL4	ILMN_31891	7	angiopoietin-like 4 (ANGPTL4), transcript variant 3.
ANKRD22	ILMN_7804	1	ankyrin repeat domain 22 (ANKRD22).
ANKRD23	ILMN_32220	2	ankyrin repeat domain 23 (ANKRD23).
ANKRD37	ILMN_2423		ankyrin repeat domain 37 (ANKRD37).
ANKRD5	ILMN_9924	20	ankyrin repeat domain 5 (ANKRD5).
AP3M1	ILMN_5879	10	adaptor-related protein complex 3, mu 1 subunit (AP3M1).
APAF1	ILMN_886	12	apoptotic peptidase activating factor 1 (APAF1).
APOBEC3B	ILMN_5305	22	apolipoprotein B editing enzyme, catalytic polypeptide-like 3B (APOBEC3B).
APOL2	ILMN_19374	22	apolipoprotein L, 2 (APOL2), transcript variant alpha.
ARHGDIB	ILMN_9074	11	Rho GDP dissociation inhibitor (GDI) beta (ARHGDIB).
ARHGEF16	ILMN_3096	1	Rho guanine exchange factor (GEF) 16 (ARHGEF16).
ARL4C	ILMN_15416	2	ADP-ribosylation factor-like 4C (ARL4C).
ARMCX3	ILMN_15168	X	armadillo repeat containing, X-linked 3 (ARMCX3).
ASH2L	ILMN_14109	2	ash2 (absent, small, or homeotic)-like (Drosophila) (ASH2L).
ASPHD1	ILMN_11756	16	aspartate beta-hydroxylase domain containing 1 (ASPHD1).
ATL3	ILMN_20827	2	atlastin 3 (ATL3).
ATM	ILMN_16722	11	ataxia telangiectasia mutated (ATM).
ATP2A3	ILMN_26846	17	ATPase, Ca++ transporting, ubiquitous (ATP2A3).
BAI1	ILMN_25426	8	brain-specific angiogenesis inhibitor 1 (BAI1).
BAX	ILMN_11763	19	BCL2-associated X protein (BAX), transcript variant sigma.
BBS9	ILMN_165104	7	Bardet-Biedl syndrome 9 (BBS9).
BCL2L11	ILMN_16692	2	BCL2-like 11 (apoptosis facilitator) (BCL2L11), transcript variant 9.
BCL6	ILMN_540	3	B-cell CLL/lymphoma 6 (zinc finger protein 51) (BCL6).
BIRC7	ILMN_21032	20	baculoviral IAP repeat-containing 7 (livin) (BIRC7).
BLVRB	ILMN_11803	10	biliverdin reductase B (flavin reductase (NADPH)) (BLVRB).
BTBD3	ILMN_22115	20	BTB (POZ) domain containing 3 (BTBD3).
BTRC	ILMN_25556	10	beta-transducin repeat containing (BTRC).
C10ORF10	ILMN_8623	8	chromosome 10 open reading frame 10 (C10orf10).
C10orf11	ILMN_17394	10	chromosome 10 open reading frame 11 (C10orf11).
C10orf12	ILMN_15899	10	chromosome 10 open reading frame 12 (C10orf12).
C10orf4	ILMN_549	10	chromosome 10 open reading frame 4 (C10orf4).
C10orf61	ILMN_7866	10	chromosome 10 open reading frame 61 (C10orf61).
C14orf143	ILMN_13424	14	chromosome 14 open reading frame 143 (C14orf143).

C14orf151	ILMN_6769	14	chromosome 14 open reading frame 151 (C14orf151).
C14ORF151	ILMN_6769	3	chromosome 14 open reading frame 151 (C14orf151).
C14orf21	ILMN_9819	14	chromosome 14 open reading frame 21 (C14orf21).
C14orf45	ILMN_24020	14	chromosome 14 open reading frame 45 (C14orf45).
C15orf38	ILMN_18724	15	chromosome 15 open reading frame 38 (C15orf38).
C15ORF41	ILMN_18248	14	chromosome 15 open reading frame 41 (C15orf41).
C16orf35	ILMN_42546	16	chromosome 16 open reading frame 35 (C16orf35).
C19orf45	ILMN_16641	19	chromosome 19 open reading frame 45 (C19orf45).
C1ORF66	ILMN_20031	17	chromosome 1 open reading frame 66 (C1orf66).
C1orf84	ILMN_10202	1	chromosome 1 open reading frame 84 (C1orf84).
C20orf107	ILMN_1548	20	chromosome 20 open reading frame 107 (C20orf107).
C20orf195	ILMN_7222	20	chromosome 20 open reading frame 195 (C20orf195).
C2ORF15	ILMN_16952	11	chromosome 2 open reading frame 15 (C2orf15).
C3orf18	ILMN_9603	3	chromosome 3 open reading frame 18 (C3orf18).
C3orf63	ILMN_25103	3	chromosome 3 open reading frame 63 (C3orf63).
C4orf36	ILMN_5377	4	chromosome 4 open reading frame 36 (C4orf36).
C4orf38	ILMN_24921	4	chromosome 4 open reading frame 38 (C4orf38).
C5orf34	ILMN_179972	5	chromosome 5 open reading frame 34 (C5orf34).
C5orf39	ILMN_17443	5	chromosome 5 open reading frame 39 (C5orf39).
C6ORF166	ILMN_1311	8	chromosome 6 open reading frame 166 (C6orf166).
C6ORF57	ILMN_23605	10	chromosome 6 open reading frame 57 (C6orf57).
C8orf13	ILMN_27702	8	chromosome 8 open reading frame 13 (C8orf13).
C8ORF40	ILMN_20841	3	chromosome 8 open reading frame 40 (C8orf40).
C9ORF119	ILMN_34116	1	chromosome 9 open reading frame 119 (C9orf119).
C9orf165	ILMN_13417	9	chromosome 9 open reading frame 165 (C9orf165).
C9orf66	ILMN_3927	9	chromosome 9 open reading frame 66 (C9orf66).
C9orf84	ILMN_17236	9	chromosome 9 open reading frame 84 (C9orf84).
C9ORF9	ILMN_12570		chromosome 9 open reading frame 9 (C9orf9).
CALB2	ILMN_11415	10	calbindin 2, 29kDa (calretinin) (CALB2), transcript variant CALB2c.
CALN1	ILMN_27637	7	calneuron 1 (CALN1).
CATSPER1	ILMN_5805	11	cation channel, sperm associated 1 (CATSPER1).
CCDC110	ILMN_15423	4	coiled-coil domain containing 110 (CCDC110).
CCDC117	ILMN_170970	22	coiled-coil domain containing 117 (CCDC117).
CCDC34	ILMN_2645	11	coiled-coil domain containing 34 (CCDC34).
CCDC89	ILMN_4285	11	coiled-coil domain containing 89 (CCDC89).
CCNB2	ILMN_15254	3	cyclin B2 (CCNB2).
CCNT2	ILMN_11186	2	cyclin T2 (CCNT2), transcript variant a.
CD300LG	ILMN_20387	17	CD300 molecule-like family member g (CD300LG).
CD68	ILMN_5188	9	CD68 molecule (CD68), transcript variant 1.
CD70	ILMN_19983	19	CD70 molecule (CD70).
CDA	ILMN_137423	16	cytidine deaminase (CDA).
CDC25B	ILMN_137183	20	cell division cycle 25B (CDC25B), transcript variant 4.
CDH24	ILMN_791	14	cadherin-like 24 (CDH24).
CDK6	ILMN_3062	19	cyclin-dependent kinase 6 (CDK6).
CDKN2B	ILMN_183997	9	cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4) (CDKN2B).
CDS2	ILMN_18323	19	CDP-diacylglycerol synthase (phosphatidate cytidylyltransferase) 2 (CDS2).
CENPH	ILMN_18041	5	centromere protein H (CENPH).
CEP192	ILMN_7783	18	centrosomal protein 192kDa (CEP192).
CEP57	ILMN_27141	11	centrosomal protein 57kDa (CEP57).
CGI-96	ILMN_12816	22	CGI-96 protein (CGI-96).
CHEK2	ILMN_4865	22	CHK2 checkpoint homolog (S. pombe) (CHEK2).
CITED4	ILMN_15271	11	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain.
CLEC2D	ILMN_15059	12	C-type lectin domain family 2, member D (CLEC2D).
CLPB	ILMN_2118	11	ClpB caseinolytic peptidase B homolog (E. coli) (CLPB).
CMTM1	ILMN_8242	16	CKLF-like MARVEL transmembrane domain containing 1 (CMTM1).
CNOT7	ILMN_137418	8	CCR4-NOT transcription complex, subunit 7 (CNOT7).
CNTNAP5	ILMN_12389	2	contactin associated protein-like 5 (CNTNAP5).
COL17A1	ILMN_5067	10	collagen, type XVII, alpha 1 (COL17A1).
COX17	ILMN_19252	9	COX17 cytochrome c oxidase assembly homolog (S, cerevisiae) (COX17).

CPA4	ILMN_21403	1	carboxypeptidase A4 (CPA4).
CPEB2	ILMN_3402	4	cytoplasmic polyadenylation element binding protein 2 (CPEB2).
CPZ	ILMN_14396	4	carboxypeptidase Z (CPZ).
CSNK1E	ILMN_24481	22	casein kinase 1, epsilon (CSNK1E).
CSNK1G2	ILMN_17274	14	casein kinase 1, gamma 2 (CSNK1G2).
CYB561	ILMN_9883	17	cytochrome b-561 (CYB561).
CYB5D2	ILMN_26268	11	cytochrome b5 domain containing 2 (CYB5D2).
CYBRD1	ILMN_4649	X	cytochrome b reductase 1 (CYBRD1).
CYP4F2	ILMN_18195	19	cytochrome P450, family 4, subfamily F, polypeptide 2 (CYP4F2).
DAG1	ILMN_16432	3	dystroglycan 1 (dystrophin-associated glycoprotein 1) (DAG1).
DBP	ILMN_1724	19	D site of albumin promoter (albumin D-box) binding protein (DBP).
DCP2	ILMN_4971	17	DCP2 decapping enzyme homolog (S, cerevisiae) (DCP2).
DCTN3	ILMN_2126	9	dynactin 3 (p22) (DCTN3).
DDI2	ILMN_27420	1	DDI1, DNA-damage inducible 1, homolog 2 (S. cerevisiae) (DDI2).
DDIT4	ILMN_13176	2	DNA-damage-inducible transcript 4 (DDIT4).
DDX12	ILMN_34429	12	PREDICTED: DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 12.
DDX60	ILMN_17673	4	DEAD (Asp-Glu-Ala-Asp) box polypeptide 60 (DDX60).
DEFB131	ILMN_166014	4	defensin, beta 131 (DEFB131).
DHRS3	ILMN_12432	X	dehydrogenase/reductase (SDR family) member 3 (DHRS3).
DICER1	ILMN_24464	14	Dicer1, Dcr-1 homolog (Drosophila) (DICER1).
DKFZp434K191	ILMN_28495	22	hypothetical protein DKFZp434K191 (DKFZp434K191).
DKFZp686O24166	ILMN_19977	11	hypothetical protein DKFZp686O24166 (DKFZp686O24166).
DKK1	ILMN_22862	15	dickkopf homolog 1 (Xenopus laevis) (DKK1).
DNAJC16	ILMN_139021	1	DnaJ (Hsp40) homolog, subfamily C, member 16 (DNAJC16).
DNASE1	ILMN_14833	16	deoxyribonuclease I (DNASE1).
DNASE2B	ILMN_27910	1	deoxyribonuclease II beta (DNASE2B).
DPH1	ILMN_21347	17	DPH1 homolog (S. cerevisiae) (DPH1).
DSCAM	ILMN_1163	21	Down syndrome cell adhesion molecule (DSCAM).
DTX2	ILMN_21612	7	deltex homolog 2 (Drosophila) (DTX2).
DUSP2	ILMN_21154	2	dual specificity phosphatase 2 (DUSP2).
ECE2	ILMN_175313	3	endothelin converting enzyme 2 (ECE2).
EFEMP2	ILMN_137851	11	EGF-containing fibulin-like extracellular matrix protein 2 (EFEMP2).
EGR2	ILMN_10721	10	early growth response 2 (Krox-20 homolog, Drosophila) (EGR2).
EIF4EBP2	ILMN_16446	6	eukaryotic translation initiation factor 4E binding protein 2 (EIF4EBP2).
EIF5A2	ILMN_4591	3	eukaryotic translation initiation factor 5A2 (EIF5A2).
EPB41	ILMN_15301	1	erythrocyte membrane protein band 4.1 (elliptocytosis 1, RH-linked) (EPB41).
EPM2A	ILMN_16230	6	epilepsy, progressive myoclonus type 2A, Lafora disease (laforin) (EPM2A).
EPS8L1	ILMN_15325	19	EPS8-like 1 (EPS8L1).
ERC1	ILMN_32989	12	ELKS/RAB6-interacting/CAST family member 1 (ERC1), transcript variant gamma.
ERGIC1	ILMN_7272	5	endoplasmic reticulum-golgi intermediate compartment (ERGIC) 1 (ERGIC1).
FAM100B	ILMN_22874	6	family with sequence similarity 100, member B (FAM100B).
FAM127A	ILMN_27078	3	family with sequence similarity 127, member A (FAM127A).
FAM129B	ILMN_8005	9	family with sequence similarity 129, member B (FAM129B).
FAM53B	ILMN_165220	10	family with sequence similarity 53, member B (FAM53B).
FAM73A	ILMN_4385	1	family with sequence similarity 73, member A (FAM73A).
FBXO16	ILMN_1262	3	F-box protein 16 (FBXO16).
FBXO17	ILMN_20838	19	F-box protein 17 (FBXO17).
FBXO24	ILMN_3577	7	F-box protein 24 (FBXO24).
FBXO27	ILMN_14974	19	F-box protein 27 (FBXO27).
FBXO27	ILMN_14974	7	F-box protein 27 (FBXO27).
FBXW8	ILMN_4038	12	F-box and WD repeat domain containing 8 (FBXW8).
FGFR3	ILMN_22960	4	fibroblast growth factor receptor 3.
FHIT	ILMN_15082	3	fragile histidine triad gene (FHIT).
FLJ10374	ILMN_3627	19	hypothetical protein FLJ10374 (FLJ10374).
FLJ30058	ILMN_29759	X	hypothetical protein FLJ30058 (FLJ30058).
FLJ35767	ILMN_28580	6	FLJ35767 protein (FLJ35767).
FLJ38717	ILMN_13488	6	FLJ38717 protein (FLJ38717).
FLJ39061	ILMN_172404		hypothetical protein FLJ39061 (FLJ39061).
FRYL	ILMN_138098	4	FRY-like (FRYL).

FRZB	ILMN_29091	2	frizzled-related protein (FRZB).
FUSIP1	ILMN_30145	1	FUS interacting protein (serine/arginine-rich) 1 (FUSIP1).
GALE	ILMN_24444	1	UDP-galactose-4-epimerase (GALE).
GCUD2	ILMN_19354	1|NT_113874.1	gastric cancer up-regulated-2 (GCUD2).
GLB1	ILMN_23626	3	galactosidase, beta 1 (GLB1), transcript variant 179423.
GLUL	ILMN_25881	1	glutamate-ammonia ligase (glutamine synthetase) (GLUL), transcript variant 3.
GNAT1	ILMN_21210	3	guanine nucleotide binding protein (G protein).
GOLGA6B	ILMN_6462	15	golgi autoantigen, golgin subfamily a, 6B (GOLGA6B).
GPNMB	ILMN_25675	17	glycoprotein (transmembrane) nmb (GPNMB), transcript variant 2.
GPR161	ILMN_22837	1	G protein-coupled receptor 161 (GPR161).
GPX1	ILMN_10376	3	glutathione peroxidase 1 (GPX1).
GRB10	ILMN_2830	7	growth factor receptor-bound protein 10 (GRB10).
GRINL1A	ILMN_20762	15	glutamate receptor, ionotropic, N-methyl D-aspartate-like 1A (GRINL1A).
H1FX	ILMN_26614		H1 histone family, member X (H1FX).
H3F3B	ILMN_26885		H3 histone, family 3B (H3,3B) (H3F3B).
HACL1	ILMN_24469	9	2-hydroxyacyl-CoA lyase 1 (HACL1).
HDAC7A	ILMN_29995	12	histone deacetylase 7A (HDAC7A), transcript variant 3.
HESX1	ILMN_23312	3	HESX homeobox 1 (HESX1).
HESX1	ILMN_23312	11	HESX homeobox 1 (HESX1).
HEXIM2	ILMN_12244		hexamthylene bis-acetamide inducible 2 (HEXIM2).
HIST1H2AM	ILMN_26622	6	histone cluster 1, H2am (HIST1H2AM).
HIST1H2BM	ILMN_27454	6	histone cluster 1, H2bm (HIST1H2BM).
HIST2H2BE	ILMN_28293	1	histone cluster 2, H2be (HIST2H2BE).
HLA-DMB	ILMN_2252	6	major histocompatibility complex, class II, DM beta (HLA-DMB).
HMGCL	ILMN_18301	3	3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase.
HNRNPA3P1	ILMN_5777	10	heterogeneous nuclear ribonucleoprotein A3 pseudogene 1 (HNRNPA3P1)
HOXA4	ILMN_138508	7	homeobox A4 (HOXA4).
HOXB1	ILMN_138553	17	homeobox B1 (HOXB1).
HOXC6	ILMN_172478	12	homeobox C6 (HOXC6).
HS,12876	ILMN_71296	10	602659965F1 NCI_CGAP_Skn3 cDNA clone IMAGE:4802969 5, sequence
HS,279842	ILMN_84472	17	HSPC157 protein, (cDNA clone IMAGE:6672800), partial cds
HS,326560	ILMN_86232	7	PREDICTED: LOC440151 (LOC440151),
HS,386232	ILMN_89152	9	RST21348 Athersys RAGE Library cDNA, sequence
HS,568690	ILMN_120871	12	PM0-ST0264-161199-001-b06 ST0264 cDNA, sequence
HS,568741	ILMN_120922	7	TC125227 Human breast cancer tissue.
HS,569162	ILMN_121343	6	AV681673 GKB cDNA clone GKBABD06 5, sequence
HS,574590	ILMN_126771	19	DA728582 NT2RM2 cDNA clone NT2RM2002174 5, sequence
HS.120300	ILMN_75629		HESC4_33_E11.g1_A037 NIH_MGC_262 cDNA clone IMAGE:7474271 5.
HS.123214	ILMN_75922		hd27b07.y1 Human Retina cDNA.
HS.125056	ILMN_76084		cDNA FLJ36663 fis, clone UTERU2002826
HS.144479	ILMN_78077		AGENCOURT_13979145 NIH_MGC_179 cDNA clone IMAGE:30367627 5.
HS.145444	ILMN_78206		cDNA FLJ11494 fis, clone HEMBA1001942
HS.167721	ILMN_80111		cDNA FLJ37425 fis, clone BRAWH2001530
HS.168162	ILMN_80119		AGENCOURT_13779116 NIH_MGC_184 cDNA clone IMAGE:30349586 5.
HS.23459	ILMN_71835		PREDICTED: hypothetical LOC388727 (LOC388727),
HS.245405	ILMN_83114		UI-H-CO0-ara-d-07-0-UI.s1 NCI_CGAP_Sub9 cDNA clone IMAGE:3105827 3.
HS.412918	ILMN_90317		cDNA FLJ32550 fis, clone SPLEN1000056
HS.445843	ILMN_93075		602617110F1 NIH_MGC_79 cDNA clone IMAGE:4730811 5.
HS.518527	ILMN_99459		PREDICTED: hypothetical LOC389189 (LOC389189),
HS.527515	ILMN_100905		AGENCOURT_13631433 NIH_MGC_184 cDNA clone IMAGE:30327753 5.
HS.539765	ILMN_104729		601823962F1 NIH_MGC_79 cDNA clone IMAGE:4043678 5.
HS.548213	ILMN_109801		tx54c03.x1 NCI_CGAP_Lu24 cDNA clone IMAGE:2273380 3.
HS.561526	ILMN_114927		full-length cDNA clone CS0DN003YC08 of Adult brain of (human)
HS.569175	ILMN_121356		DA697821 NT2NE2 cDNA clone NT2NE2019092 5, sequence
HS.569953	ILMN_122134		MR3-FN0206-070201-014-b05 FN0206 cDNA, sequence
HS.571028	ILMN_123209		17000531973183 GRN_ES cDNA 5, sequence
HS.571741	ILMN_123922		ny62g05.s1 NCI_CGAP_GCB1 cDNA clone IMAGE:1282904 3, sequence
HS.572064	ILMN_124245		UI-E-EJ1-ajh-l-06-0-UI.r1 UI-E-EJ1 cDNA clone UI-E-EJ1-ajh-l-06-0-UI 5.
HS.575085	ILMN_127266		DB336481 TESTI2 cDNA clone TESTI2007279 3, sequence

HS.98330	ILMN_74482		PREDICTED: hypothetical LOC388227 (LOC388227),
ID1	ILMN_28002	5	inhibitor of DNA binding 1, dominant negative helix-loop-helix protein (ID1).
IGSF8	ILMN_5527	1	immunoglobulin superfamily, member 8 (IGSF8).
IL15RA	ILMN_4552	10	interleukin 15 receptor, alpha (IL15RA).
IL18BP	ILMN_10368	11	interleukin 18 binding protein (IL18BP), transcript variant D.
IL6R	ILMN_22419	1	interleukin 6 receptor (IL6R).
INPP5A	ILMN_137853		PREDICTED: inositol polyphosphate-5-phosphatase, 40kDa (INPP5A).
KAZALD1	ILMN_28149	10	Kazal-type serine peptidase inhibitor domain 1 (KAZALD1).
KCNH6	ILMN_171630	17	potassium voltage-gated channel, subfamily H (eag-related), member 6 (KCNH6).
KCNJ1	ILMN_19663	11	potassium inwardly-rectifying channel, subfamily J, member 1 (KCNJ1).
KCNN4	ILMN_11994		potassium intermediate/small conductance calcium-activated channel.
KIAA0247	ILMN_2972	14	KIAA0247 (KIAA0247).
KIAA1107	ILMN_33304		PREDICTED: KIAA1107 (KIAA1107).
KIAA1539	ILMN_29031	17	KIAA1539 (KIAA1539).
KLF9	ILMN_2670	6	Kruppel-like factor 9 (KLF9).
KLHDC9	ILMN_13602	1	kelch domain containing 9 (KLHDC9).
KLK8	ILMN_21619	19	kallikrein-related peptidase 8 (KLK8).
KRT15	ILMN_3189		keratin 15 (KRT15).
KRT9	ILMN_15527	17	keratin 9 (epidermolytic palmoplantar keratoderma) (KRT9).
LEMD1	ILMN_28526	19	LEM domain containing 1 (LEMD1).
LEPREL1	ILMN_11123	3	leprecan-like 1 (LEPREL1).
LILRB3	ILMN_14901	19	leukocyte immunoglobulin-like receptor, subfamily B (withTM and ITIM domains).
LIMA1	ILMN_6603	12	LIM domain and actin binding 1 (LIMA1).
LIPH	ILMN_17820	3	lipase, member H (LIPH).
LMTK3	ILMN_35173	20	PREDICTED: lemur tyrosine kinase 3 (LMTK3).
LOC144383	ILMN_41654		PREDICTED: similar to Interferon-induced transmembrane protein 3.
LOC144383	ILMN_41654	19	PREDICTED: similar to Interferon-induced transmembrane protein 3.
LOC150383	ILMN_11584	22	similar to RIKEN cDNA 2210021J22 (LOC150383).
LOC286310	ILMN_34458	9	PREDICTED: lipocalin 1-like 1 (LOC286310), misc RNA.
LOC338758	ILMN_37634	12	PREDICTED: hypothetical protein LOC338758 (LOC338758).
LOC388344	ILMN_34544	17	PREDICTED: similar to ribosomal protein L13, transcript variant 1 (LOC388344).
LOC390561	ILMN_138198	15	PREDICTED: similar to hect domain and RLD 2 (LOC390561).
LOC390637	ILMN_16995	15	similar to RIKEN cDNA D330012F22 gene (LOC390637).
LOC400214	ILMN_42653		PREDICTED: hypothetical gene supported by BX248296 (LOC400214).
LOC401677	ILMN_36492		PREDICTED: similar to eukaryotic translation elongation factor 1 alpha 2.
LOC441294	ILMN_169689	7	similar to CTAGE6 (LOC441294).
LOC442597	ILMN_42881		PREDICTED: hypothetical LOC442597 (LOC442597).
LOC51035	ILMN_13595	15	SAPK substrate protein 1 (LOC51035).
LOC641978	ILMN_31151		PREDICTED: similar to general transcription factor II I (LOC641978).
LOC642035	ILMN_33714	5	PREDICTED: hypothetical protein LOC642035 (LOC642035).
LOC642393	ILMN_41363	1	PREDICTED: similar to mitochondrial ribosomal protein L20, transcript variant 2.
LOC642726	ILMN_39041	4	PREDICTED: hypothetical protein LOC642725, transcript variant 1 (LOC642726).
LOC642969	ILMN_31159	12	PREDICTED: similar to Phosphoglycerate mutase 1.
LOC643272	ILMN_32441	10	PREDICTED: hypothetical protein LOC643272 (LOC643272).
LOC644889	ILMN_33330	11	PREDICTED: similar to large subunit ribosomal protein L36a (LOC644889).
LOC645676	ILMN_41958	1	PREDICTED: hypothetical protein LOC645676, transcript variant 1 (LOC645676).
LOC653604	ILMN_40991	1	PREDICTED: similar to H3 histone, family 2 isoform 2 (LOC653604).
LOC654126	ILMN_40336		PREDICTED: similar to leucine rich repeat containing 37B.
LOC729776	ILMN_102997	9	PREDICTED: hypothetical protein LOC729776 (LOC729776).
LOC730083	ILMN_32599	16	PREDICTED: similar to exonuclease domain containing 1 (LOC730083).
LOC731950	ILMN_138552		PREDICTED: similar to slit (Drosophila) homolog 2 (LOC731950).
LRRC1	ILMN_13909	6	leucine rich repeat containing 1 (LRRC1).
LRRC29	ILMN_11967	16	leucine rich repeat containing 29 (LRRC29).
LRRC59	ILMN_26734	17	leucine rich repeat containing 59 (LRRC59).
LRRFIP2	ILMN_5592	3	leucine rich repeat (in FLII) interacting protein 2 (LRRFIP2).
LSM6	ILMN_8596	4	LSM6 homolog, U6 small nuclear RNA associated (S. cerevisiae) (LSM6).
LYZ	ILMN_4879	12	lysozyme (renal amyloidosis) (LYZ).
M6PRBP1	ILMN_10971	4	mannose-6-phosphate receptor binding protein 1 (M6PRBP1).
MACF1	ILMN_3751	1	microtubule-actin crosslinking factor 1 (MACF1).

MADD	ILMN_9428	11	MAP-kinase activating death domain (MADD), transcript variant 5.
MAGEH1	ILMN_8979	X	melanoma antigen family H, 1 (MAGEH1).
MAP2	ILMN_38764	2	microtubule-associated protein 2 (MAP2).
MAP3K3	ILMN_426	17	mitogen-activated protein kinase kinase kinase 3 (MAP3K3).
MAPK14	ILMN_17267	6	mitogen-activated protein kinase 14 (MAPK14).
MAPK7	ILMN_9862	17	mitogen-activated protein kinase 7 (MAPK7), transcript variant 3.
MAX	ILMN_2124	14	MYC associated factor X (MAX), transcript variant 3.
ME3	ILMN_24802	11	malic enzyme 3, NADP(+)-dependent, mitochondrial (ME3).
MGAT5	ILMN_21616	2	mannosyl (alpha-1,6-)-glycoprotein beta-1,6-N-acetyl-glucosaminyltransferase.
MGC3207	ILMN_138056	19	hypothetical protein MGC3207 (MGC3207).
MGC59937	ILMN_13120	9	Similar to RIKEN cDNA 2310002J15 gene (MGC59937).
MGC59937	ILMN_13120	19	Similar to RIKEN cDNA 2310002J15 gene (MGC59937).
MGC70863	ILMN_10471	22	similar to RPL23AP7 protein (MGC70863).
MIB2	ILMN_139150		PREDICTED: mindbomb homolog 2 (Drosophila) (MIB2).
MIDN	ILMN_6472	14	midnolin (MIDN).
MLYCD	ILMN_2280	16	malonyl-CoA decarboxylase(MLYCD), nucl. gene encoding mitochondrial protein.
MMP7	ILMN_9188	11	matrix metallopeptidase 7 (matrilysin, uterine) (MMP7).
MMP9	ILMN_28136	20	matrix metallopeptidase 9.
MRPL40	ILMN_21771		mitochondrial ribosomal protein L40 (MRPL40).
MUC1	ILMN_162845	1	mucin 1, cell surface associated (MUC1), transcript variant 5.
MUC4	ILMN_164899	3	mucin 4, cell surface associated (MUC4).
MYEOV	ILMN_4623	8	myeloma overexpressed gene (in a subset of t(11;14).
MZF1	ILMN_8368	19	myeloid zinc finger 1 (MZF1).
NAT5	ILMN_43222	20	N-acetyltransferase 5 (NAT5), transcript variant 3.
NAT6	ILMN_29898	17	N-acetyltransferase 6 (NAT6).
NBPF10	ILMN_137199	1	PREDICTED: neuroblastoma breakpoint family.
NBR2	ILMN_1966	17	neighbor of BRCA1 gene 2 (NBR2).
NDUFS2	ILMN_9109	1	NADH dehydrogenase (ubiquinone) Fe-S protein 2, 49kDa.
NEDD4	ILMN_27071	15	neural precursor cell expressed, developmentally down-regulated 4 (NEDD4).
NEIL1	ILMN_18060	15	nei endonuclease VIII-like 1 (E. coli) (NEIL1).
NEK11	ILMN_10349	3	NIMA (never in mitosis gene a)- related kinase 11 (NEK11).
NEK2	ILMN_14211	19	NIMA (never in mitosis gene a)-related kinase 2 (NEK2).
NOPE	ILMN_17897	15	neighbor of Punc E11 (NOPE).
NSMCE2	ILMN_15101	10	non-SMC element 2, MMS21 homolog (S, cerevisiae) (NSMCE2).
NT5E	ILMN_28610	10	5'-nucleotidase, ecto (CD73) (NT5E).
NT5M	ILMN_14877	17	5',3'-nucleotidase, mitochondrial (NT5M).
NUBP2	ILMN_10701	2	nucleotide binding protein 2 (MinD homolog, E, coli) (NUBP2).
OLFML2B	ILMN_572	1	olfactomedin-like 2B (OLFML2B).
OR1J1	ILMN_13947	9	olfactory receptor, family 1, subfamily J, member 1 (OR1J1).
OR2A20P	ILMN_13267	7	olfactory receptor, family 2, subfamily A, member 20 pseudogene.
OR51B2	ILMN_20840		olfactory receptor, family 51, subfamily B, member 2 (OR51B2).
OR51F1	ILMN_19096	11	olfactory receptor, family 51, subfamily F, member 1 (OR51F1).
OR51I1	ILMN_10084	11	olfactory receptor, family 51, subfamily I, member 1 (OR51I1).
P2RY2	ILMN_15883	11	purinergic receptor P2Y, G-protein coupled, 2 (P2RY2).
PAG1	ILMN_174074	8	phosphoprotein associated with glycosphingolipid microdomains 1 (PAG1).
PARD6G	ILMN_138732		PREDICTED: par-6 partitioning defective 6 homolog gamma (C. elegans).
PAX8	ILMN_6456	2	paired box 8 (PAX8), transcript variant PAX8D.
PBRM1	ILMN_16253	3	polybromo 1 (PBRM1), transcript variant 4.
PCOLCE	ILMN_13969	1	procollagen C-endopeptidase enhancer (PCOLCE).
PCTK1	ILMN_11214	X	PCTAIRE protein kinase 1 (PCTK1).
PDE4C	ILMN_29696	19	phosphodiesterase 4C, cAMP-specific.
PERP	ILMN_4512	16	PERP, TP53 apoptosis effector (PERP).
PGBD4	ILMN_6577	15	piggyBac transposable element derived 4 (PGBD4).
PIGX	ILMN_21837	3	phosphatidylinositol glycan anchor biosynthesis, class X (PIGX).
PLA2G6	ILMN_13517	22	phospholipase A2, group VI (cytosolic, calcium-independent) (PLA2G6).
PLEKHH2	ILMN_28132	2	pleckstrin homology domain containing, family H (with MyTH4 domain).
PMEPA1	ILMN_24935	20	prostate transmembrane protein, androgen induced 1 (PMEPA1).
PMP22	ILMN_9212	17	peripheral myelin protein 22 (PMP22).
PODXL	ILMN_24120	7	podocalyxin-like (PODXL).

POLD4	ILMN_14887	17	polymerase (DNA-directed), delta 4 (POLD4).
POLR2J4	ILMN_9922		polymerase (RNA) II (DNA directed) polypeptide J, 13.3kDa pseudogene.
PORCN	ILMN_19819	X	porcupine homolog (Drosophila) (PORCN), transcript variant A.
PPAP2B	ILMN_5681	1	phosphatidic acid phosphatase type 2B (PPAP2B).
PPARA	ILMN_7270	22	peroxisome proliferator-activated receptor alpha (PPARA), transcript variant 3.
PPARG	ILMN_22381	3	peroxisome proliferator-activated receptor gamma (PPARG).
PPFIBP1	ILMN_21261	12	PTPRF interacting protein, binding protein 1 (liprin beta 1) (PPFIBP1).
PPM1A	ILMN_10552	14	protein phosphatase 1A (formerly 2C), magnesium-dependen.
PPP1R14A	ILMN_20916	19	protein phosphatase 1, regulatory (inhibitor) subunit 14A (PPP1R14A).
PPP1R1C	ILMN_36531	2	protein phosphatase 1, regulatory (inhibitor) subunit 1C (PPP1R1C).
PPP2R5C	ILMN_18075	22	protein phosphatase 2, regulatory subunit B', gamma isoform (PPP2R5C).
PQLC3	ILMN_138926	6	PQ loop repeat containing 3 (PQLC3).
PRKAG2	ILMN_671	7	protein kinase, AMP-activated, gamma 2 non-catalytic subunit.
PRNPIP	ILMN_38371		PREDICTED: prion protein interacting protein, transcript variant 4 (PRNPIP).
PRR11	ILMN_137089	20	proline rich 11 (PRR11).
PRSS22	ILMN_21409	16	protease, serine, 22 (PRSS22).
PTCH1	ILMN_18640	9	patched homolog 1 (Drosophila) (PTCH1), transcript variant 1c'.
PTGS2	ILMN_29986	1	prostaglandin-endoperoxide synthase 2.
PTH2	ILMN_2034	17	parathyroid hormone 2 (PTH2).
PTHLH	ILMN_4025	12	parathyroid hormone-like hormone (PTHLH), transcript variant 3.
PTPN14	ILMN_27079	1	protein tyrosine phosphatase, non-receptor type 14 (PTPN14).
PTPRM	ILMN_19957	14	protein tyrosine phosphatase, receptor type, M (PTPRM).
RAB11FIP1	ILMN_4635	8	RAB11 family interacting protein 1 (class I) (RAB11FIP1).
RAB11FIP4	ILMN_427	17	RAB11 family interacting protein 4 (class II) (RAB11FIP4).
RAP2B	ILMN_178464	3	RAP2B, member of RAS oncogene family (RAP2B).
RAPSN	ILMN_849	11	receptor-associated protein of the synapse (RAPSN).
RASSF2	ILMN_10884	20	Ras association (RalGDS/AF-6) domain family 2 (RASSF2).
RASSF6	ILMN_15686	4	Ras association (RalGDS/AF-6) domain family member 6 (RASSF6).
RBM9	ILMN_13392	11	RNA binding motif protein 9 (RBM9), transcript variant 3.
RCAN3	ILMN_26881	1	RCAN family member 3 (RCAN3).
RGS12	ILMN_19624	4	regulator of G-protein signaling 12 (RGS12).
RHOF	ILMN_1762	7	ras homolog gene family, member F (in filopodia) (RHOF).
RILP	ILMN_15991	17	Rab interacting lysosomal protein (RILP).
RNASE4	ILMN_16267	12	ribonuclease, RNase A family, 4 (RNASE4), transcript variant 3.
RNF128	ILMN_28439	X	ring finger protein 128 (RNF128).
ROPN1B	ILMN_37645		PREDICTED: ropporin, rhophilin associated protein 1B.
RPL28	ILMN_10642	1	ribosomal protein L28 (RPL28).
RPS26P10	ILMN_40627	8	PREDICTED: ribosomal protein S26 pseudogene 10 (RPS26P10).
RRAS	ILMN_23748		related RAS viral (r-ras) oncogene homolog (RRAS).
RTN3	ILMN_20904	11	reticulon 3 (RTN3).
RUNX3	ILMN_16236	1	runt-related transcription factor 3 (RUNX3).
SAMD13	ILMN_10578	22	sterile alpha motif domain containing 13 (SAMD13).
SCARNA9	ILMN_25861	19	small Cajal body-specific RNA 9 (SCARNA9) on chromosome 11,
SCEL	ILMN_15617	13	sciellin (SCEL).
SCNN1A	ILMN_5697	1	sodium channel, nonvoltage-gated 1 alpha (SCNN1A).
SDCCAG1	ILMN_3136	6	serologically defined colon cancer antigen 1 (SDCCAG1).
SDPR	ILMN_11513	2	serum deprivation response (phosphatidylserine binding protein) (SDPR).
SENP6	ILMN_14173	6	SUMO1/sentrin specific peptidase 6 (SENP6).
SEPT10	ILMN_5056	2	septin 10 (SEPT10).
SEPT8	ILMN_36453	5	septin 8 (SEPT8), transcript variant 4.
SERPINB1	ILMN_10210	1	serpin peptidase inhibitor, clade B (ovalbumin), member 1 (SERPINB1).
SH2D1A	ILMN_9529	X	SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome).
SHD	ILMN_13682	19	Src homology 2 domain containing transforming protein D (SHD).
SHFM1	ILMN_26583	12	split hand/foot malformation (ectrodactyly) type 1 (SHFM1).
SLC22A18	ILMN_20563	11	solute carrier family 22 (organic cation transporter), member 18 (SLC22A18).
SLC22A9	ILMN_172906	11	solute carrier family 22 (organic anion transporter), member 9 (SLC22A9).
SLC25A16	ILMN_165967	10	solute carrier family 25 (mitochondrial carrier; Graves disease autoantigen).
SLC2A12	ILMN_19964	6	solute carrier family 2 (facilitated glucose transporter), member 12 (SLC2A12).
SLC2A12	ILMN_19964	1	solute carrier family 2 (facilitated glucose transporter), member 12 (SLC2A12).

SLC2A3	ILMN_15812	11	solute carrier family 2 (facilitated glucose transporter), member 3 (SLC2A3).
SLC30A5	ILMN_24834	5	solute carrier family 30 (zinc transporter), member 5 (SLC30A5).
SLC35F5	ILMN_22210	2	solute carrier family 35, member F5 (SLC35F5).
SMG1	ILMN_9076	16	PI-3-kinase-related kinase SMG-1 (SMG1).
SMG7	ILMN_29503	1	Smg-7 homolog, nonsense mediated decay factor (C. elegans) (SMG7).
SMYD5	ILMN_26626	2	SMYD family member 5 (SMYD5).
SNORD73A	ILMN_669	4	small nucleolar RNA, C/D box 73A (SNORD73A) on chromosome 4.
SOCS3	ILMN_167297	17	suppressor of cytokine signaling 3 (SOCS3).
SP8	ILMN_24886	7	Sp8 transcription factor (SP8).
SPAG1	ILMN_12212	8	sperm associated antigen 1 (SPAG1).
SPDYC	ILMN_1637	11	speedy homolog C (Drosophila) (SPDYC).
SPOCD1	ILMN_6339	1	SPOC domain containing 1 (SPOCD1).
SPPL3	ILMN_30223	12	signal peptide peptidase 3 (SPPL3).
SRC	ILMN_6831	20	v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian) (SRC).
ST8SIA3	ILMN_16138	18	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 3 (ST8SIA3).
STAG2	ILMN_14082	X	stromal antigen 2 (STAG2), transcript variant 4.
STAP2	ILMN_136955	19	signal transducing adaptor family member 2 (STAP2).
STEAP1	ILMN_139257		PREDICTED: six transmembrane epithelial antigen of the prostate 1 (STEAP1).
STEAP2	ILMN_18795	7	six transmembrane epithelial antigen of the prostate 2 (STEAP2).
STK40	ILMN_25410	1	serine/threonine kinase 40 (STK40).
SUMO1	ILMN_14949	2	SMT3 suppressor of mif two 3 homolog 1 (S. cerevisiae) (SUMO1).
SUZ12P	ILMN_38242	17	PREDICTED: suppressor of zeste 12 homolog pseudogene.
TACSTD2	ILMN_4004	1	tumor-associated calcium signal transducer 2 (TACSTD2).
TAF1C	ILMN_4122	16	TATA box binding protein (TBP)-associated factor, RNA polymerase I, C.
TALDO1	ILMN_138767		PREDICTED: transaldolase 1 (TALDO1).
TCEA3	ILMN_27218	1	transcription elongation factor A (SII), 3 (TCEA3).
TCOF1	ILMN_18418	5	Treacher Collins-Franceschetti syndrome 1 (TCOF1).
TCP11L2	ILMN_21615	12	t-complex 11 (mouse)-like 2 (TCP11L2).
TFPI	ILMN_1429	7	tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor).
TGM2	ILMN_8134	20	transglutaminase 2 (C polypeptide, protein-glutamine-gamma-glutamyltransferase).
TH	ILMN_138010	11	tyrosine hydroxylase (TH), transcript variant 3.
TIAM2	ILMN_9891	6	T-cell lymphoma invasion and metastasis 2 (TIAM2).
TICAM2	ILMN_24482		toll-like receptor adaptor molecule 2 (TICAM2).
TIMM17A	ILMN_16007	1	translocase of inner mitochondrial membrane 17 homolog A (yeast).
TM4SF18	ILMN_5533	18	transmembrane 4 L six family member 18 (TM4SF18).
TMEM117	ILMN_21932	12	transmembrane protein 117 (TMEM117).
TMEM207	ILMN_2922	3	transmembrane protein 207 (TMEM207).
TMEM45A	ILMN_30168	19	transmembrane protein 45A (TMEM45A).
TMEM64	ILMN_3155	8	transmembrane protein 64 (TMEM64).
TMEM87B	ILMN_20699	2	transmembrane protein 87B (TMEM87B).
TNFRSF6B	ILMN_14212	19	tumor necrosis factor receptor superfamily, member 6b, decoy.
TOMM40L	ILMN_42128	1	translocase of outer mitochondrial membrane 40 homolog (yeast)-like.
TRIM16	ILMN_139304		PREDICTED: tripartite motif-containing 16 (TRIM16).
TSC22D3	ILMN_9893		TSC22 domain family, member 3 (TSC22D3), transcript variant 2.
TSGA10	ILMN_16441	2	testis specific, 10 (TSGA10).
TSPAN1	ILMN_7052	1	tetraspanin 1 (TSPAN1).
TTC35	ILMN_16778	8	tetratricopeptide repeat domain 35 (TTC35).
TTC9C	ILMN_5250	3	tetratricopeptide repeat domain 9C (TTC9C).
TUBD1	ILMN_16764	17	tubulin, delta 1 (TUBD1).
UBE1C	ILMN_22726	3	ubiquitin-activating enzyme E1C (UBA3 homolog, yeast) (UBE1C).
UBE2V1	ILMN_12143	20	ubiquitin-conjugating enzyme E2 variant 1 (UBE2V1), transcript variant 3.
UPK3B	ILMN_24270	7	uroplakin 3B (UPK3B).
USP30	ILMN_5598	12	ubiquitin specific peptidase 30 (USP30).
USP54	ILMN_18622	10	ubiquitin specific peptidase 54 (USP54).
VAC14	ILMN_30132		PREDICTED: Vac14 homolog (S. cerevisiae) (VAC14).
VAMP3	ILMN_19403	3	vesicle-associated membrane protein 3 (cellubrevin) (VAMP3).
VEGFA	ILMN_5181	6	vascular endothelial growth factor A (VEGFA), transcript variant 3.
VGF	ILMN_9112	2	VGF nerve growth factor inducible (VGF).
VIM	ILMN_676	1	vimentin (VIM).

WHSC1	ILMN_27418	4	Wolf-Hirschhorn syndrome candidate 1 (WHSC1), transcript variant 8.
YAP1	ILMN_19290	20	Yes-associated protein 1, 65kDa (YAP1).
ZBTB44	ILMN_30202	11	zinc finger and BTB domain containing 44 (ZBTB44).
ZC3H14	ILMN_22941	14	zinc finger CCCH-type containing 14 (ZC3H14).
ZDHHC11	ILMN_138235	5	zinc finger, DHHC-type containing 11 (ZDHHC11).
ZFP36	ILMN_1557		zinc finger protein 36, C3H type, homolog (mouse) (ZFP36).
ZKSCAN5	ILMN_27210	7	zinc finger with KRAB and SCAN domains 5 (ZKSCAN5).
ZNF197	ILMN_1508	3	zinc finger protein 197 (ZNF197).
ZNF566	ILMN_16016	19	zinc finger protein 566 (ZNF566).
ZNF773	ILMN_1981	19	zinc finger protein 773 (ZNF773).
ZSCAN1	ILMN_19352	19	zinc finger and SCAN domain containing 1 (ZSCAN1).

B. Genes upregulated after ASH2L silencing

SYMBOL	ILMN_GENE	CHROMOSOME	DEFINITION
ACP2	ILMN_3044	11	acid phosphatase 2, lysosomal (ACP2).
AKAP8	ILMN_10271	19	A kinase (PRKA) anchor protein 8 (AKAP8).
ALCAM	ILMN_21054	3	activated leukocyte cell adhesion molecule (ALCAM).
AQR	ILMN_166517	15	aquarius homolog (mouse) (AQR), mRNA.
AXIN2	ILMN_26857	17	axin 2 (conductin, axil) (AXIN2).
B3GNT2	ILMN_138549	2	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2 (B3GNT2).
BCLAF1	ILMN_3336	6	BCL2-associated transcription factor 1 (BCLAF1), transcript variant 2.
BRPF1	ILMN_17423	3	bromodomain and PHD finger containing, 1 (BRPF1), transcript variant 1.
C14orf173	ILMN_41230	14	chromosome 14 open reading frame 173 (C14orf173), transcript variant 2, mRNA.
C18orf32	ILMN_26126	18	chromosome 18 open reading frame 32 (C18orf32), mRNA.
CAND1	ILMN_22065	12	cullin-associated and neddylation-dissociated 1 (CAND1).
CASC3	ILMN_28416	17	cancer susceptibility candidate 3 (CASC3).
CCDC86	ILMN_27103	11	coiled-coil domain containing 86 (CCDC86).
CEBPA	ILMN_27029	19	CCAAT/enhancer binding protein (C/EBP), alpha (CEBPA).
CENTD3	ILMN_5090	5	centaurin, delta 3 (CENTD3).
CEP72	ILMN_10995	5	centrosomal protein 72kDa (CEP72), mRNA.
CHAC2	ILMN_1763	2	ChaC, cation transport regulator homolog 2 (E. coli) (CHAC2), mRNA.
CHPF	ILMN_22953	2	chondroitin polymerizing factor (CHPF).
CHTF18	ILMN_28360	16	CTF18, chromosome transmission fidelity factor 18 homolog (S, cerevisiae) (CHTF18)
CLCN7	ILMN_8600	16	chloride channel 7 (CLCN7).
CNTNAP1	ILMN_6876	17	contactin associated protein 1 (CNTNAP1), mRNA.
COL6A1	ILMN_138363	21	collagen, type VI, alpha 1 (COL6A1).
CPLX1	ILMN_30247	4	complexin 1 (CPLX1).
CPS1	ILMN_15726	2	carbamoyl-phosphate synthetase 1, mitochondrial (CPS1).
CTPS	ILMN_18906	1	CTP synthase (CTPS).
CYB5R4	ILMN_137297	6	cytochrome b5 reductase 4 (CYB5R4).
CYP2D6	ILMN_27062	22	cytochrome P450, family 2, subfamily D, polypeptide 6 (CYP2D6), transcript variant 2
DDX3X	ILMN_139078	X	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked (DDX3X), mRNA.
DDX3X	ILMN_139078	X	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked (DDX3X).
DEGS1	ILMN_6374	1	degenerative spermatocyte homolog 1, lipid desaturase (Drosophila) (DEGS1),
DMAP1	ILMN_138222	1	DNA methyltransferase 1 associated protein 1 (DMAP1), transcript variant 2.
DNMT1	ILMN_17904	19	DNA (cytosine-5-)-methyltransferase 1 (DNMT1).
DUS3L	ILMN_3805	19	dihydrouridine synthase 3-like (S, cerevisiae) (DUS3L).
DUSP4	ILMN_181455	8	dual specificity phosphatase 4 (DUSP4), transcript variant 2, mRNA.
E2F6	ILMN_14185	2	E2F transcription factor 6 (E2F6), mRNA.
EFNB2	ILMN_3827	13	ephrin-B2 (EFNB2).
EIF4G1	ILMN_5831	3	eukaryotic translation initiation factor 4 gamma, 1 (EIF4G1), transcript variant 1, mRNA.
EVPL	ILMN_1544	17	envoplakin (EVPL).
FASTKD5	ILMN_10466	20	FAST kinase domains 5 (FASTKD5).
FGF11	ILMN_8195	17	fibroblast growth factor 11 (FGF11), mRNA.
FGF19	ILMN_18897	11	fibroblast growth factor 19 (FGF19).
FGF9	ILMN_1771	13	fibroblast growth factor 9 (glia-activating factor) (FGF9).
FGFR3	ILMN_22960	4	fibroblast growth factor receptor 3 (achondroplasia, thanatophoric dwarfism) (FGFR3),

FKBP5	ILMN_16562	6	FK506 binding protein 5 (FKBP5).
FOXC1	ILMN_23624	6	forkhead box C1 (FOXC1).
FOXF1	ILMN_11804	16	forkhead box F1 (FOXF1), mRNA.
FOXO3	ILMN_15283	6	forkhead box O3 (FOXO3), transcript variant 1.
GCC1	ILMN_21168	7	GRIP and coiled-coil domain containing 1 (GCC1),
GGA2	ILMN_17168	16	golgi associated, gamma adaptin ear containing, ARF binding protein 2 (GGA2),
GNB1	ILMN_26098	1	guanine nucleotide binding protein (G protein), beta polypeptide 1 (GNB1),
GNS	ILMN_6937	12	glucosamine (N-acetyl)-6-sulfatase (Sanfilippo disease IIID) (GNS),
HDAC2	ILMN_28766	6	histone deacetylase 2 (HDAC2),
HYOU1	ILMN_659	11	hypoxia up-regulated 1 (HYOU1),
ICMT	ILMN_137648	1	isoprenylcysteine carboxyl methyltransferase (ICMT),
IDH3A	ILMN_3303	15	isocitrate dehydrogenase 3 (NAD+) alpha (IDH3A),
INPP5E	ILMN_11866	9	inositol polyphosphate-5-phosphatase, 72 kDa (INPP5E),
ISG20L1	ILMN_12401	15	interferon stimulated exonuclease gene 20kDa-like 1 (ISG20L1),
KCNH3	ILMN_11930	12	potassium voltage-gated channel, subfamily H (eag-related), member 3 (KCNH3),
KIAA0913	ILMN_16257	10	KIAA0913 (KIAA0913),
KIAA1712	ILMN_5346	4	KIAA1712 (KIAA1712).
KIAA1737	ILMN_24671	14	KIAA1737 (KIAA1737),
KLHDC5	ILMN_1962	12	kelch domain containing 5 (KLHDC5),
KTELC1	ILMN_15488	3	KTEL (Lys-Tyr-Glu-Leu) containing 1 (KTELC1),
KTI12	ILMN_27125	1	KTI12 homolog, chromatin associated (S, cerevisiae) (KTI12),
LCOR	ILMN_173510	10	ligand dependent nuclear receptor corepressor (LCOR),
LEP	ILMN_10827	7	leptin (obesity homolog, mouse) (LEP).
LMF2	ILMN_11132	22	lipase maturation factor 2 (LMF2),
LMNB2	ILMN_24712	19	lamin B2 (LMNB2),
LPCAT1	ILMN_15076	5	lysophosphatidylcholine acyltransferase 1 (LPCAT1),
LRRC14	ILMN_29237	8	leucine rich repeat containing 14 (LRRC14),
LSM14A	ILMN_19241	19	LSM14A, SCD6 homolog A (S, cerevisiae) (LSM14A),
LYRM2	ILMN_30197	6	LYR motif containing 2 (LYRM2),
LZTR1	ILMN_18977	22	leucine-zipper-like transcription regulator 1 (LZTR1),
MANEAL	ILMN_11393	1	mannosidase, endo-alpha-like (MANEAL), transcript variant 1.
MAP6D1	ILMN_7455	3	MAP6 domain containing 1 (MAP6D1),
MED22	ILMN_22202	9	mediator complex subunit 22 (MED22), transcript variant c,
MEPCE	ILMN_8017	7	methylphosphate capping enzyme (MEPCE),
MGEA5	ILMN_11399	10	meningioma expressed antigen 5 (hyaluronidase) (MGEA5),
MKX	ILMN_28370	10	mohawk homeobox (MKX),
MLSTD2	ILMN_11547	11	male sterility domain containing 2 (MLSTD2),
MRPS18C	ILMN_12033	4	mitochondrial ribosomal protein S18C (MRPS18C).
MSX1	ILMN_137891	4	msh homeobox 1 (MSX1),
MTCH2	ILMN_2631	11	mitochondrial carrier homolog 2 (C. elegans) (MTCH2).
NADK	ILMN_29863	1	NAD kinase (NADK),
NCOA5	ILMN_3004	20	nuclear receptor coactivator 5 (NCOA5).
NLRP2	ILMN_17259	19	NLR family, pyrin domain containing 2 (NLRP2),
NOL5A	ILMN_13841	20	nucleolar protein 5A (56kDa with KKE/D repeat) (NOL5A),
NR6A1	ILMN_2143	9	nuclear receptor subfamily 6, group A, member 1 (NR6A1), transcript variant 2.
NRBP2	ILMN_27830	8	nuclear receptor binding protein 2 (NRBP2),
NT5DC2	ILMN_20328	3	5'-nucleotidase domain containing 2 (NT5DC2),
NUDT16L1	ILMN_20270	16	nudix (nucleoside diphosphate linked moiety X)-type motif 16-like 1 (NUDT16L1).
NXF1	ILMN_11773	11	nuclear RNA export factor 1 (NXF1), transcript variant 1,
OGDHL	ILMN_15789	10	oxoglutarate dehydrogenase-like (OGDHL),
PAQR3	ILMN_20371	4	progestin and adipoQ receptor family member III (PAQR3),
PFAS	ILMN_17615	17	phosphoribosylformylglycinamidine synthase (FGAR amidotransferase) (PFAS),
POLR3B	ILMN_11759	12	polymerase (RNA) III (DNA directed) polypeptide B (POLR3B),
PPAP2A	ILMN_5193	5	phosphatidic acid phosphatase type 2A (PPAP2A), transcript variant 2.
PPP4R1	ILMN_6519	18	protein phosphatase 4, regulatory subunit 1 (PPP4R1), transcript variant 2,
PPRC1	ILMN_10445	10	peroxisome proliferator-activated receptor gamma, coactivator-related 1 (PPRC1),
PRELID1	ILMN_28597	5	PRELI domain containing 1 (PRELID1),
PSMA7	ILMN_13260	20	proteasome (prosome, macropain) subunit, alpha type, 7 (PSMA7), transcript variant 2.
PTDSS1	ILMN_4241	8	phosphatidylserine synthase 1 (PTDSS1),

PTK7	ILMN_11421	6	PTK7 protein tyrosine kinase 7 (PTK7), transcript variant PTK7-2,
PYGB	ILMN_21544	20	phosphorylase, glycogen; brain (PYGB),
RASSF1	ILMN_11841	3	Ras association (RalGDS/AF-6) domain family 1 (RASSF1), transcript variant B.
RBM12	ILMN_22052	20	RNA binding motif protein 12 (RBM12), transcript variant 1,
RBM14	ILMN_16867	11	RNA binding motif protein 14 (RBM14),
RBM16	ILMN_3973	6	RNA binding motif protein 16 (RBM16),
RBM38	ILMN_20092	20	RNA binding motif protein 38 (RBM38), transcript variant 2,
RCBTB2	ILMN_15980	13	regulator of chromosome condensation (RCC1) and BTB (POZ)
RIC8B	ILMN_3462	12	resistance to inhibitors of cholinesterase 8 homolog B (C. elegans) (RIC8B).
RIMS3	ILMN_21581	1	regulating synaptic membrane exocytosis 3 (RIMS3),
RNF145	ILMN_27136	5	ring finger protein 145 (RNF145),
RNF19A	ILMN_10959	8	ring finger protein 19A (RNF19A), transcript variant 2.
RNF219	ILMN_38012	13	ring finger protein 219 (RNF219),
RRP1B	ILMN_14288	21	ribosomal RNA processing 1 homolog B (S, cerevisiae) (RRP1B),
RRS1	ILMN_9627	8	RRS1 ribosome biogenesis regulator homolog (S, cerevisiae) (RRS1),
SAR1B	ILMN_16595	5	SAR1 gene homolog B (S, cerevisiae) (SAR1B), transcript variant 1,
SCRN2	ILMN_2736	17	secernin 2 (SCRN2).
SETMAR	ILMN_17510	3	SET domain and mariner transposase fusion gene (SETMAR),
SF4	ILMN_9581	19	splicing factor 4 (SF4), transcript variant c,
SFPQ	ILMN_5703	1	splicing factor proline/glutamine-rich(polypyrimidine tract binding protein associated).
SFRS6	ILMN_24964	20	splicing factor, arginine/serine-rich 6 (SFRS6),
SFRS7	ILMN_7620	2	splicing factor, arginine/serine-rich 7, 35kDa (SFRS7),
SH3GLB2	ILMN_14480	9	SH3-domain GRB2-like endophilin B2 (SH3GLB2),
SHRM	ILMN_16821	4	shroom (SHRM),
SIGIRR	ILMN_18194	11	single immunoglobulin and toll-interleukin 1 receptor (TIR) domain (SIGIRR).
SLC25A10	ILMN_24086	17	solute carrier family 25 (mitochondrial carrier; dicarboxylate transporter).
SLC41A1	ILMN_2825	1	solute carrier family 41, member 1 (SLC41A1),
SLC4A2	ILMN_6956	7	solute carrier family 4, anion exchanger, member 2
SNAPC4	ILMN_3646	9	small nuclear RNA activating complex, polypeptide 4, 190kDa (SNAPC4),
SNORD13	ILMN_135987	8	small nucleolar RNA, C/D box 13 (SNORD13) on chromosome 8,
SPEN	ILMN_3876	1	spen homolog, transcriptional regulator (Drosophila) (SPEN),
SRPRB	ILMN_2452	3	signal recognition particle receptor, B subunit (SRPRB),
STC2	ILMN_28725	5	stanniocalcin 2 (STC2),
STK35	ILMN_27051	20	serine/threonine kinase 35 (STK35),
TEAD2	ILMN_4452	19	TEA domain family member 2 (TEAD2),
TEX2	ILMN_27579	17	testis expressed 2 (TEX2),
TLN1	ILMN_18029	9	talin 1 (TLN1),
TLR5	ILMN_18399	1	toll-like receptor 5 (TLR5).
TMEM132A	ILMN_41781	11	transmembrane protein 132A (TMEM132A), transcript variant 2,
TMEM177	ILMN_22670	2	transmembrane protein 177 (TMEM177),
TMEM184B	ILMN_10219	22	transmembrane protein 184B (TMEM184B),
TMEM200A	ILMN_19059	6	transmembrane protein 200A (TMEM200A),
TMEM41A	ILMN_28063	3	transmembrane protein 41A (TMEM41A),
TMEM50B	ILMN_138521	21	transmembrane protein 50B (TMEM50B),
TMEM63A	ILMN_932	1	transmembrane protein 63A (TMEM63A).
TOP1MT	ILMN_15321		PREDICTED: topoisomerase (DNA) I, mitochondrial (TOP1MT),
TRRAP	ILMN_18258	7	transformation/transcription domain-associated protein (TRRAP),
TRSPAP1	ILMN_23632	1	tRNA selenocysteine associated protein 1 (TRSPAP1), transcript variant 2.
TSEN2	ILMN_2707	3	tRNA splicing endonuclease 2 homolog (S, cerevisiae) (TSEN2),
TSPAN17	ILMN_18131	5	tetraspanin 17 (TSPAN17), transcript variant 1,
TTC15	ILMN_21565	2	tetratricopeptide repeat domain 15 (TTC15),
TTL	ILMN_14027	2	tubulin tyrosine ligase (TTL),
TUBB2A	TUBB	6	tubulin, beta 2A (TUBB2A),
TUBB4	ILMN_23388	19	tubulin, beta 4 (TUBB4),
TXLNA	ILMN_1892	1	taxilin alpha (TXLNA),
TYK2	ILMN_29800	19	tyrosine kinase 2 (TYK2),
UBE2G2	ILMN_22176	21	ubiquitin-conjugating enzyme E2G 2 (UBC7 homolog, yeast), transcript variant 1,
UBE2I	ILMN_27920	16	ubiquitin-conjugating enzyme E2I (UBC9 homolog, yeast), transcript variant 4.
UBP1	ILMN_25221	3	upstream binding protein 1 (LBP-1a) (UBP1),

USP38	ILMN_28204	4	ubiquitin specific peptidase 38 (USP38),
VASN	ILMN_138591	16	vasorin (VASN),
YWHAH	ILMN_27184	22	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein.
ZBED1	ILMN_14030	X	zinc finger, BED-type containing 1 (ZBED1),
ZBED4	ILMN_8641	22	zinc finger, BED-type containing 4 (ZBED4),
ZBED5	ILMN_17879	11	zinc finger, BED-type containing 5 (ZBED5),
ZFP36L2	ILMN_169693	2	zinc finger protein 36, C3H type-like 2 (ZFP36L2),
ZNF342	ILMN_7933	19	zinc finger protein 342 (ZNF342),
ZNF35	ILMN_15315	3	zinc finger protein 35 (ZNF35),
ZNF512	ILMN_5859	2	zinc finger protein 512 (ZNF512),
ZNF518B	ILMN_3143	4	zinc finger protein 518B (ZNF518B),
ZNF559	ILMN_11090	19	zinc finger protein 559 (ZNF559),
ZNF696	ILMN_21311	8	zinc finger protein 696 (ZNF696),

Figure S2. Promoters of genes downregoleted after ASH2 silencing, were analyzed with Pscan alghortim (Zambelli F. et al.Nucleic Acids Res. 2009 Jul 1;37) using the matrices of the JASPAR database.

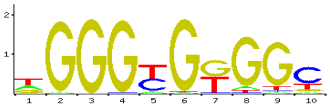
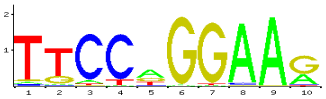
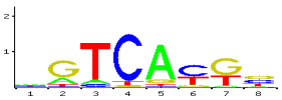
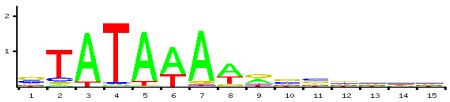
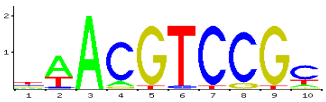
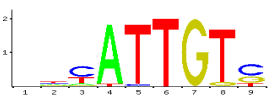
TF_NAME	LOGO	MATRIX_ID	Z_SCORE	P_VALUE
NF-YA		Dolfini et al. Ref. 18	2.65185	0.00389626
Klf4		MA0039.2	2.24908	0.0120964
Stat3		MA0144.1	2.07726	0.0186067
Pax2		MA0067.1	2.07831	0.0186399
TBP		MA0108.2	1.93156	0.0264203
Myf		MA0055.1	1.78919	0.0364266
MIZF		MA0131.1	1.6672	0.0474185
Sox17		MA0078.1	1.55573	0.0595298

Figure S3. List of primers used in q-RT-PCR and ChIPs.

ChIP-PCR PRIMERS

MDM2	GGTTGACTCAGCTTTTCCTCTTG
MDM2	GGAAAATGCATGGTTTAAATAGCC
UNG	CGCGCGCCTATAATCCTAGC
UNG	GCTTGATGGCTCACGTCTGT
NCOA5	GTGGTCCGGAGGTTACAGGAC
NCOA5	GAGCACATTCCCTCCTCCCTA
TICAM2	AGCGCTCTACCACCAACGATT
TICAM2	GACAAGAGTCTCGCCCTGTCA
HDAC1	CCTCTCCGGGCTGCCCTTG
HDAC1	CCTCCTGGGCTCACCTATAGC
ERP70	CTCACGTTAGGGCTCGGAGTTT
ERP70	GGAAAAACCCACGGAAGTCGT
CHOP	CTCGTGACCCAAAGCCACTTC
CHOP	GGACCCCAAAACCTACCAATCAG
CCN B1	TGTCACCTTCCAAAGGCCACTA
CCN B1	AGAAGAGCCAGCCTAGCCTCAG
CCN B2	AGAGGCGTCCTACGTCTGCTTT
CCN B2	ATTCAAATACCGCGTCGCTTG
HSPA5	AGGGGAGGACCTGAACGGTTAC
HSPA5	TGTTGTCTCGGCCAGTATCGAG
PCNA	GAGTCAAAGAGGCGGGGAGAC
PCNA	CTTGCGGGGAAGACTTTAGGG
CKS2	CCCCGTGACGTACCTATCTT
CKS2	ACAACTCGCCGGAGACTAAC

HERPUD	ATTGGGCCACGTTGGGAGAGT
HERPUD	CAACGACAGTTCACGTCTCTGG
ID2	TCTTGATAGACGTGCCACCTTCC
ID2	TAACGGACCTCACGGGACTGA

RT-PCR PRIMERS

ANKRD37	AATCCACATGACCAAGCGAGA
ANKRD37	TAAGTCAGTGGGCGTGAGAGG
TICAM2	AAGAGAAGCTCAAGGCCGAAG
TICAM2	TCCAAGGCAGAAGAGGAAAAC
ALSR2CR2	GCCATCAGCAAGCAGTTTATTG
ALSR2CR2	AAGCAGGAGGCAACAGTGAAA
ID1	TCCGCTCAGCACCCTCAAC
ID1	CGCTTCAGCGACACAAGATG
MDM2	CAGCTTCGGAACAAGAGACC
MDM2	GGCACGCCAAACAAATCTCC
APOBEC3B	CTGCTTCTCCTGGGGCTGT
APOBEC3B	GACATCCCTGGCGGTACAC
DDX50	ATAGCTCAAGCACGGACAGG
DDX50	GCCACGCTGAGTTTCCTAGT
CDKN1A	CTGGAGACTCTCAGGGTCGAA
CDKN1A	GGATTAGGGCTTCCTCTTGGA
CDKN1B	CCACGAAGAGTTAACCCGGG
CDKN1B	GTCTGCTCCACAGAACCGGC
YWHA	ACTTTTGGTACATTGTGGCTTCAA
YWHA	CCGCCAGGACAAACCAGTAT
ID2	CCTCAACACGGATATCAGCA
ID2	AGAACACCCTGGGAAGATGA
CCNA2	TATTGCTGGAGCTGCCTTTC
CCNA2	CTCTGGTGGGTTGAGGAGAG
CBL	GCTGGTTGTCTCTGGATGGT
CBL	CCCCTGACTCATGAGGTTCT

NES	GCAGCAGCTGGCGCACCTCAAGA
NES	GCCAGGTGTTTGCAGCCGGGAGT
DNAJ	GAGTGGAACCCGGAGACATTG
DNAJ	GATACTGCGGCATCCCTTCAC
HERPUD1	CTACTCCTCCCTGAGCAGATT
HERPUD1	GGTTGGGGTCTTAGTTTCAG

CONSIDERATIONS

This project started after two considarations:

First consideration:
In literature, is known that ASH2L, a component of MLL complexes (a very important histone methyltransferase) is essential for the enzymatic activity of the complex (Steward 2006).
In absence of Ash2l in fact, the methyltransferases are able to mono and dimetylates the lisyne four of H3 but cannot trimethylates the histone tail.

Second consideration:
In my previous work, we have seen that the recruitment and the binding to DNA of NF-Y is a very important event for the transcriptional activation.
Furthermore in absence of NF-Y on core promoter of the genes (NF-Y's targets) the level of trimethylation of H3K4 are lower.

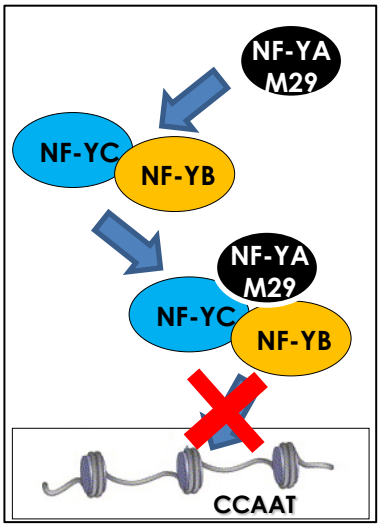


Fig. 12 *The NF-YA m29 mutant substantially removed the trimer from promoters*

To obtain the above described data, we performed a adenoviral infection.
We infected NIH 3T3 cell line with adenovirus in order to express a dominant- negative mutant form of NF-YA and GFP and NF-YA wt as controls.
The dominant negative named M29 is a mutated form of the NF-YA sub-unit that competes with NF-YA wt to form the eterotrimer.
The mutant, crippled in the DNA binding subdomain, is able to associate with the Histone Fold Motive of NF-YB/NF-YC dimer, rendering the complex incapable of binding to a CCAAT box (Fig.12)

We performed experiment of chromatin immunoprecipitation (ChIP) , with the infected cells and checked if the binding of NF-Y to DNA was correlated with the level of trimethylation of the H3K4.
As expected, in the cells infected with the M29 mutant, the data obtained show that the reduction of binding of NF-Y is correlated with lower levels of the trimethylation of the lisyne 4 of H3 histone in NF-Y's targets (Fig.23).

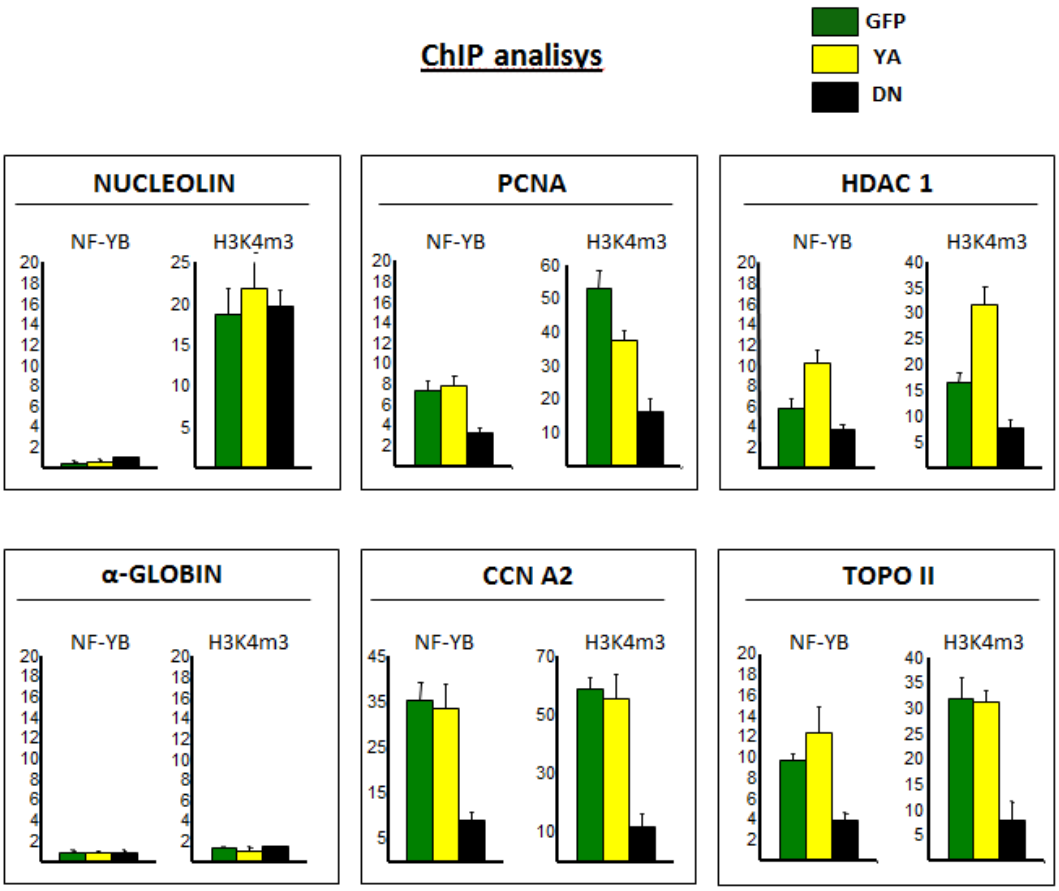


Fig. 23: Rappresentation of Semi-quantitative PCR. Values are represented as fold enrichment over the control (anti-Flag).
NUCLEOLIN and αGLOBIN are the controls of the experiment, the first is a transcribted gene but not target of NF-Y, the second is a not transcribted gene.
PCNA, HDAC1, CCN A2 and TOPO2 are transcripted genes and all targets of NF-Y.

In green, the cells infected with the adenovirus of GFP, in yellow the cells infected with NF-YA WT and in black the cells infected with NF-YA M29

Another important consideration regards the lysine 79 of H3.
Is known in literature in fact that the presence of MLL's translocation (RS 4:11 cell lines) is correlated with high level of lysine 79 dimethylation of this PTM.
The explanation of this is not completely clear but has been hypothesized that this phenomenon could be related to the expression of the chimeric protein formed by MLL and AF4.

We have confirmed this fact with experiment of chromatin immunoprecipitation realized on the RS4:11 cell lines (see fig.5) but after the experiment of chromatin immunoprecipitation in Ash2l silenced HCT 116 cell lines, we have shown that the same pattern of lysine 79 dimethylation is also present in this cells (see fig.1b).

It's Important to note that in HCT 116 cell lines, the chromosomal translocation between the chromosome 4 and the chromosome 11 is not present. As a result, the loci of MLL and AF-4 are WT and the chimeric protein (MLL-AF4) is not transcribed.

APPENDIX

PLoS One. 2008 Apr 30;3(4):e2066.

An NF-Y-dependent switch of positive and negative histone methyl marks on CCAAT promoters.

Donati G, Gatta R, Dolfini D, **Fossati A**, Ceribelli M, Mantovani R.

Dipartimento di Scienze Biomolecolari e Biotecnologie, Università di Milano, Milano, Italy

Received February 7, 2008; Accepted March 11, 2008.

BIBLIOGRAPHY

Armstrong SA, Look AT.
Molecular genetics of acute lymphoblastic leukemia.
J Clin Oncol. 2005 Sep 10;23(26):6306-15. Review.

Armstrong SA, Staunton JE, Silverman LB, Pieters R, den Boer ML, Minden MD, Sallan SE, Lander ES, Golub TR, Korsmeyer SJ.
MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia.
Nat Genet. 2002 Jan;30(1):41-7. Epub 2001 Dec 3.

Berger SL.
The complex language of chromatin regulation during transcription.
Nature. 2007 May 24;447(7143):407-12.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K.
High-resolution profiling of histone methylations in the human genome.
Cell. 2007 May 18;129(4):823-37.

Birke M, Schreiner S, García-Cuéllar MP, Mahr K, Titgemeyer F, Slany RK.
The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation.
Nucleic Acids Res. 2002 Feb 15;30(4):958-65.

Bitoun E, Oliver PL, Davies KE.
The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling.
Hum Mol Genet. 2007 Jan 1;16(1):92-106. Epub 2006 Nov 29.

Bucher, P.
Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences.
J Mol Biol. 1990 Apr 20;212(4):563-78.

Burmeister T, Meyer C, Schwartz S, Hofmann J, Molkentin M, Kowarz E, Schneider B, Raff T, Reinhardt R, Gökbuget N, Hoelzer D, Thiel E, Marschalek R.

The MLL recombinome of adult CD10-negative B-cell precursor acute lymphoblastic leukemia: results from the GMALL study group.
Blood. 2009 Apr 23;113(17):4011-5. Epub 2009 Jan 14.

Caslini C, Yang Z, El-Osta M, Milne TA, Slany RK, Hess JL.
Interaction of MLL amino terminal sequences with menin is required for transformation.
Cancer Res. 2007 Aug 1;67(15):7275-83.

Coustry, F., Maity, S.N. and de Crombrugghe
Studies on transcription activation by the multimeric CCAAT-binding factor CBF
J Biol Chem. 1995 Jan 6;270(1):468-75.

Donati G, Gatta R, Dolfini D, Fossati A, Ceribelli M, Mantovani R.
An NF-Y-dependent switch of positive and negative histone methyl marks on CCAAT promoters.
PLoS One. 2008 Apr 30;3(4):e2066.

Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW, Verdine GL, Allis CD, Roeder RG.
Regulation of MLL1 H3K4 methyltransferase activity by its core components.
Nat Struct Mol Biol. 2006 Aug;13(8):713-9. Epub 2006 Jul 30.

Edwards, D., Murrey, J.A.H., Smith, A.G..
Multiple genes coding the conserved CCAAT-box transcription complex are expressed in Arabidopsis.
Plant Physiol. 1998 Jul;117(3):1015-22.

Fang Cao, Yong Chen, Tomasz Cierpicki, Yifan Liu, Venkatesha Basrur1, Ming Lei, Yali Dou

An Ash2L/RbBP5 Heterodimer Stimulates the MLL1 Methyltransferase Activity through Coordinated Substrate Interactions with the MLL1 SET Domain

PLoS One. 2010 Nov 23;5(11):e14102.

Goo YH, Sohn YC, Kim DH, Kim SW, Kang MJ, Jung DJ, Kwak E, Barlev NA, Berger SL, Chow VT, Roeder RG, Azorsa DO, Meltzer PS, Suh PG, Song EJ, Lee KJ, Lee YC, Lee JW.

Activating signal cointegrator 2 belongs to a novel steady-state complex that contains a subset of trithorax group proteins.

Mol Cell Biol. 2003 Jan;23(1):140-9.

Lee JH, Tate CM, You JS, Skalnik DG.

Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex.

J Biol Chem. 2007 May 4;282(18):13419-28. Epub 2007 Mar 13.

Goppelt, A., Steltzer, G., Lottspeich, F., Meisterernst, M.

A mechanism for repression of class II gene transcription through specific binding of NC2 to TBP-promoter complexes via heterodim-eric histone fold domains.

EMBO J. 1996 Jun 17;15(12):3105-16.

Hars ES, Lyu YL, Lin CP, Liu LF.

Role of apoptotic nuclease caspase-activated DNase in etoposide-induced treatment-related acute myelogenous leukemia.

Cancer Res. 2006 Sep 15;66(18):8975-9.

Hsieh JJ, Cheng EH, Korsmeyer SJ.

Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression.

Cell. 2003 Oct 31;115(3):293-303.

Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P, Korsmeyer SJ.
Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization.
Mol Cell Biol. 2003 Jan;23(1):186-94.

Krivtsov AV, Armstrong SA.
MLL translocations, histone modifications and leukaemia stem-cell development
Nat Rev Cancer. 2007 Nov;7(11):823-33.

James J.-D. Hsieh,[1] Patricia Ernst,[1] Hediye Erdjument-Bromage,[2] Paul Tempst,[2] and Stanley J. Korsmeyer[1*]
Proteolytic Cleavage of MLL Generates a Complex of N- and C-Terminal Fragments That Confers Protein Stability and Subnuclear Localization.
Molecular and Cellular Biology, January 2003, p. 186-194, Vol. 23, No. 1
0270-7306/03/$08.00+0 DOI: 10.1128/MCB.23.1.186-194.2003

Libura J, Slater DJ, Felix CA, Richardson C.
Therapy-related acute myeloid leukemia-like MLL rearrangements are induced by etoposide in primary human CD34+ cells and remain stable after clonal expansion.
Blood. 2005 Mar 1;105(5):2124-31. Epub 2004 Nov 4.

Melissa M Steward[1], Jung-Shin Lee[1], Aisling O'Donovan[2], Matt Wyatt[1], Bradley EBernstein[2] & Ali Shilatifard[1, 3]
Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes
Nature Structural & Molecular Biology - **13**, 852 - 854 (2006)
Published online: 6 August 2006; | doi:10.1038/nsmb1131

Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, Hess JL.
MLL targets SET domain methyltransferase activity to Hox gene promoters.
Mol Cell. 2002 Nov;10(5):1107-17.

Milne TA, Hughes CM, Lloyd R, Yang Z, Rozenblatt-Rosen O, Dou Y, Schnepp RW, Krankel C, Livolsi VA, Gibbs D, Hua X, Roeder RG, Meyerson M, Hess JL.
Menin and MLL cooperatively regulate expression of cyclin-dependent kinase inhibitors.
Proc Natl Acad Sci U S A. 2005 Jan 18;102(3):749-54. Epub 2005 Jan 7.

Melissa M Steward, Jung-Shin Lee1, Aisling O'Donovan2,Matt Wyatt1, Bradley E Bernstein & Ali Shilatifard
Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes
Nat Struct Mol Biol. 2006 Sep;13(9):852-4. Epub 2006 Aug 6.

Mantovani R.
The molecular biology of the CCAAT-binding factor NF-Y.
Gene. 1999 Oct 18;239(1):15-27. Review.

Mirro J, Kitchingman G, Williams D, Lauzon GJ, Lin CC, Callihan T, Zipf TF.
Clinical and laboratory characteristics of acute leukemia with the 4;11 translocation.
Blood. 1986 Mar;67(3):689-97.

Mirro J, Zipf TF, Pui CH, Kitchingman G, Williams D, Melvin S, Murphy SB, Stass S.
Acute mixed lineage leukemia: clinicopathologic correlations and prognostic significance.
Blood. 1985 Nov;66(5):1115-23.

Morse HG, Heideman R, Hays T, Robinson A.
4;11 translocation in acute lymphoblastic leukemia: a specific syndrome.
Cancer Genet Cytogenet. 1982 Oct;7(2):165-72.

Mueller D, Bach C, Zeisig D, Garcia-Cuellar MP, Monroe S, Sreekumar A, Zhou R, Nesvizhskii A, Chinnaiyan A, Hess JL, Slany RK.

A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification.
Blood. 2007 Dec 15;110(13):4445-54. Epub 2007 Sep 12.

Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, Wassell R, Dubois G, Mazo A, Croce CM, Canaani E.
ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation.
Mol Cell. 2002 Nov;10(5):1119-28

Nakanishi S, Lee JS, Gardner KE, Gardner JM, Takahashi YH, Chandrasekharan MB, Sun ZW, Osley MA, Strahl BD, Jaspersen SL, Shilatifard A.
Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1.
J Cell Biol. 2009 Aug 10;186(3):371-7.

Patel A, Dharmarajan V, Vought VE, Cosgrove MS. (2009).
On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex.
J Biol Chem. 2009 Sep 4;284(36):24242-56. Epub 2009 Jun 2545x

Pui CH, Sandlund JT, Pei D, Campana D, Rivera GK, Ribeiro RC, Rubnitz JE, Razzouk BI, Howard SC, Hudson MM, Cheng C, Kun LE, Raimondi SC, Behm FG, Downing JR, Relling MV, Evans WE.
Improved outcome for children with acute lymphoblastic leukemia: results of Total Therapy Study XIIIB at St Jude Children's Research Hospital.
Blood. 2004 Nov 1;104(9):2690-6. Epub 2004 Jul 13.

Robert K. Slany
The molecular biology of mixed lineage leukemia
Department of Genetics, University Erlangen, Erlangen, Germany
Correspondence: Robert K. Slany, Genetics, University Erlangen, Staudtstrasse 5, 91058 Erlangen, Haematologica. 2009 July; 94(7): 984–993.

Romier C, Cocchiarella F, Mantovani R, Moras D.

The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y.
J Biol Chem. 2003 Jan 10;278(2):1336-45. Epub 2002 Oct 24.

Ross ME, Zhou X, Song G, Shurtleff SA, Girtman K, Williams WK, Liu HC, Mahfouz R, Raimondi SC, Lenny N, Patel A, Downing JR.
Classification of pediatric acute lymphoblastic leukemia by gene expression profiling.
Blood. 2003 Oct 15;102(8):2951-9. Epub 2003 May 1.

Ruthenburg AJ, Allis CD, Wysocka J.
Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark.
Mol Cell. 2007 Jan 12;25(1):15-30.

Ruthenburg AJ, Wang W, Graybosch DM, Li H, Allis CD, Patel DJ, Verdine GL.
Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex.
Nat Struct Mol Biol. 2006 Aug;13(8):704-12. Epub 2006 Jul 9.

Schreiner SA, García-Cuéllar MP, Fey GH, Slany RK.
The leukemogenic fusion of MLL with ENL creates a novel transcriptional transactivator.
Leukemia. 1999 Oct;13(10):1525-33.

Stass S, Mirro J, Melvin S, Pui CH, Murphy SB, Williams D.
Lineage switch in acute leukemia.
Blood. 1984 Sep;64(3):701-6.

Steger DJ, Lefterova MI, Ying L, Stonestrom AJ, Schupp M, Zhuo D, Vakoc AL, Kim JE, Chen J, Lazar MA, Blobel GA, Vakoc CR.
DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells.
Mol Cell Biol. 2008 Apr;28(8):2825-39. Epub 2008 Feb 19.

Steward MM, Lee JS, O'Donovan A, Wyatt M, Bernstein BE et al., (2006).
Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes.
Nat Struct Mol Biol. 2006 Sep;13(9):852-4. Epub 2006 Aug 6.

Tomizawa D, Koh K, Sato T, Kinukawa N, Morimoto A, Isoyama K, Kosaka Y, Oda T, Oda M, Hayashi Y, Eguchi M, Horibe K, Nakahata T, Mizutani S, Ishii E.
Outcome of risk-based therapy for infant acute lymphoblastic leukemia with or without an MLL gene rearrangement, with emphasis on late effects: a final report of two consecutive studies, MLL96 and MLL98, of the Japan Infant Leukemia Study Group.
Leukemia. 2007 Nov;21(11):2258-63. Epub 2007 Aug 9.

Yokoyama A, Kitabayashi I, Ayton PM, Cleary ML, Ohki M.
Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties.

Blood. 2002 Nov 15;100(10):3710-8. Epub 2002 Jun 28.

Yokoyama A, Cleary ML.
Menin critically links MLL proteins with LEDGF on cancer-associated target genes.
Cancer Cell. 2008 Jul 8;14(1):36-46.

Wysocka J, Myers MP, Laherty CD, Eisenman RN, Herr W.
Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1.
Genes Dev. 2003 Apr 1;17(7):896-911.

Zeisig DT, Bittner CB, Zeisig BB, García-Cuéllar MP, Hess JL, Slany RK.
The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin.
Oncogene. 2005 Aug 18;24(35):5525-32.

Zeleznik-Le NJ, Harden AM, Rowley JD.
11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene.
Proc Natl Acad Sci U S A. 1994 Oct 25;91(22):10610-4.