



UNIVERSITÀ DEGLI STUDI DI MILANO

Scuola di dottorato in Molecular and Cell Biology

*XXIX Ciclo*

# "DISSECTING THE CONTRIBUTION OF THE HISTONE DEMETHYLASE LSD1 IN RETINOIC ACID-INDUCED DIFFERENTIATION OF APL CELLS"

**E. Ceccacci**

PhD Thesis

**Scientific supervisor: Giulio Pavesi**

**Scientific tutor: Saverio Minucci**

Academic year: 2015-2016

SSD: BIO/11

Thesis performed at Saverio Minucci Lab, Dept. of Experimental Oncology at European Institute of Oncology

# TABLE OF CONTENTS

# ABSTRACT

Acute promyelocytic leukemia (APL) is a cytogenetically distinct subtype of acute myeloid leukemia, characterized by the chromosomal translocation t(15;17) that involves the retinoic acid receptor (RAR) gene and leads to the production of the fusion protein PML-RARα. In the past it has been successfully treated with all-trans retinoic acid at high doses to differentiate the leukemic blast. The fusion protein indeed retains the capability to binds DNA with an even stronger affinity and recruits repressive co-factors, making the cells insensitive to physiological concentrations of retinoic acid.

In the last few years, also the lysine-specific demethylase (LSD1) protein has emerged as important target for the epigenetic therapy of cancer.

We found that both pharmacological inhibition and knock down of LSD1 are able to sensitize NB4 cells - a cells line derived from an APL patient - to lower (physiological) doses of retinoic acid (RA) causing growth arrest and differentiation without degradation of the fusion protein. In order to elucidate the role of LSD1 in this mechanism, we characterized the LSD1 genomic distribution in acute myeloid leukemia by ChIP-seq experiment and performed RNA-seq and ChIP-seq for H3K4me1/me2/me3 and H3K27ac in all the four treatments (DMSO as control, RA low, RA high, LSD1 inhibition and cotreatment of LSD1 inhibitor and RA low).

Results of RNA-seq analyses show significant changes in gene expression only after co-treatment (RAlow + LSD1i) and RA high, in line with experimental

evidence of phenotype. Moreover, in addition to an high overlap between genes expressed in both co-treatment and RAhigh, we observe a significative number of cotreatment-specific expressed genes that suggest a putative synergistic and stronger effect of the cotreatment compare to the retinoic acid high alone. In parallel, the data on histone modifications obtained through ChIP-seq experiments show a significant increase of the di-metylation of H3K4 after treatment with inhibitor. This mainly occurs in regions marked with peaks of LSD1 and associated with genes involved in cell differentiation (genes over-expressed in co-treatment, but not in other single treatments). In the same regions we observe the presence of H3K27 acetylation after treatment with retinoic acid but not after LSD1 inhibition. Only after treatment with both drugs the regions acquire the two histone marks, and we can observe an effective phenotype of differentiation, correlating with the observed change in gene expression. An hypothesis might be that there are regulatory regions linked by LSD1, which undergo a kind of "pre-mark" in histone modifications after treatment with the inhibitor (gain of H3K4me2) or with retinoic acid (gain of H3K27ac), which is necessary but not sufficient to determine a change in expression, found only after co-treatment in the simultaneous presence of both epigenetic modifications. Overall the combination of the LSD1 inhibition and RA low bypasses the block of PML-RAR fusion protein activating a different pathway of genes compared to RA high with stronger effects on the differentiation of the cells.

Taken together our results contribute to understand the role of LSD1 in the RA-induced differentiation of leukemic cells, suggest new therapeutic strategies for the intervention in APL and potentially other leukemias, and highlight the

importance of combination therapies as new potent weapon in the cancer treatments.

# INTRODUCTION

## EPIGENETICS

Epigenetics is defined as heritable cell traits that are not linked to changes in the DNA sequence and comprise the mechanism by which the chromatin associated proteins and post-translational modification of histone (PTMs) regulate transcription. It is well known that all the cells within a human body contain the same DNA sequences, and the differentiation of all these cells and the acquiring of distinct functions and morphologies are due to epigenetic changes and its fine regulation. Thus epigenetics doesn't involve change of the sequence of the DNA but most of all of its spatial organization; epigenetic regulators and transcription factors organize the genome into accessible or not accessible regions, which determine the correct activation of different transcriptional program in each cell type. The identity of each cells, determined by the expression of unique gene patterns, must be remembered and passed to the daughter cells through epigenetic mechanisms. Thus epigenetic is essential for the correct maintaining of cell identity and is determinant for many fundamental process such as proliferation, development, differentiation and genome integrity. These regulations can be mediated through several mechanism like DNA methylation, ATP-dependent nucleosome remodeling, replacement of canonical histones with histone variants, post–translational modification of histone (PTMs), non-coding RNA (ncRNAs). Since each cell type has its unique epigenome, complex organisms have multiple epigenomes, depending on the tissue type and development stage.

Because of the key role of epigenetic mechanism in the control of several biological processes, it is not surprising that chromatin alteration may lead to the onset and progression of many diseases, first of all cancer. Unlike genetic alteration, epigenetic alterations are generally reversible and for this reason drugs against epigenetic target (epi.drugs) are considered as a new and promising field for cancer therapy, with already some compounds approved.

## CHROMATIN STRUCTURE AND COMPONENT

Chromatin structure can be viewed as a series of superimposed layers: at the root there is the DNA sequence and its direct chemical modification by cytosine methylation; the DNA is then folded into nucleosome, composed by 147 bp of DNA wrapped around the histone octamer, formed by a tetramer of H3-H4 histone and a couple of H2A-H2B histone dimers. The primary structure of chromatin is represented by the assembly of core nucleosomes with the fragment of DNA linker (so called "beads on a string"), while the presence of the histone H1 which binds the DNA linker, leads the structure to achieve a higher level of compaction, called "30nm fiber". The degree of chromatin wrapping, specific for cell type and differential stage, could be divided in two main subtypes or environments: euchromatin and heterochromatin. Both the environment are enriched and also depleted of certain characteristic histone modification, with regions of demarcation between heterochromatin and euchromatin. These boundary elements are enriched of specific factors such as CTCF, that play a role in the maintaining of the boundary, H3K9me1 and the histone variant H2A.Z, while are depleted of histone acetylation. Euchromatin represent regions accessible and generally active transcribed while heterochromatin includes

highly condensed regions with genes transcriptionally repressed, and with mainly repetitive DNA sequences. In addiction to these well-known structures three-dimensional models of chromatin are now reached an increased precision and suggested that there are additional sophisticated level of genome regulation through ulterior order of organization and nuclear compartmentalization.

The chromatin structure and condensed level are fining regulated and highly dynamic and are due to the combined action of DNA methylation, histone modification and chromatin remodeling. The wide range of histone modification occurs not only at the N-terminal tails, including acethylation, methylation, phosphorylation, ubiquination, SUMOylation, crotonylation and others (Zentner and Henikoff, 2013), but also in the core of the histones and in the C-terminal regions (Bannister and Kouzarides, 2011).

The enzyme responsible for the fine regulation of all these histone modification can be divided in “writers”, the enzyme responsible for the addition of chemical groups on eithers histone tails or DNA itself, while the proteins that recognize these specific epigenetic marks are called “readers”. Then, since all the epigenetic modifications are not permanent, another class of enzyme called “erasers” can remove them.

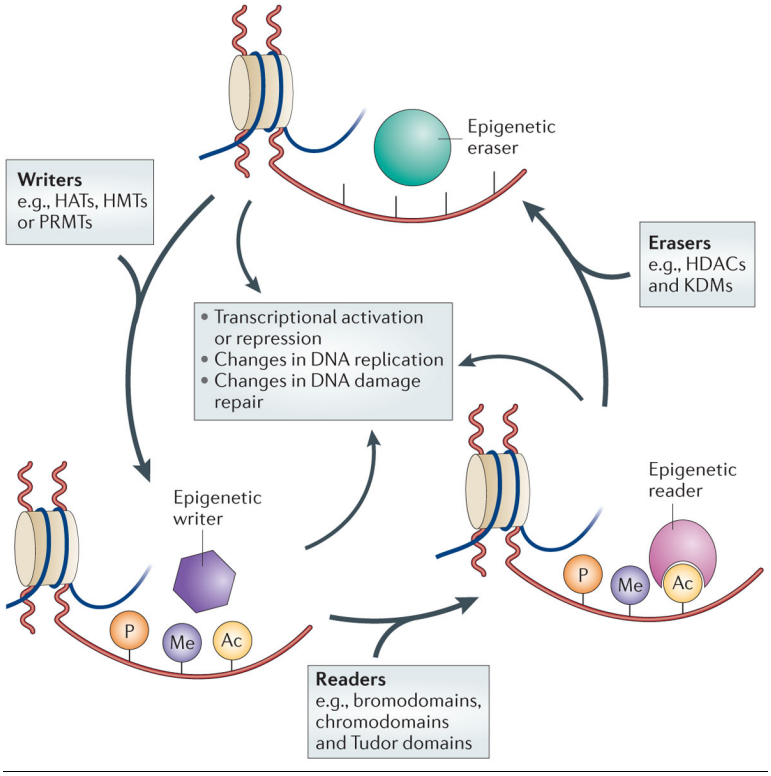



**Figure intro 1.** Epigenetic regulation is a dynamic process: epigenetic writers lay down marks on amino acid residues on the histone tails, epigenetic readers recognize and bind these marks; epigenetic erasers catalyze the removal of the histone marks. All together these addictions or removal, and the recognition of the histone marks determine a complex histone code that regulates various processes including transcription, replication and repair (Falkenberg and Johnstone, 2014).

These mechanisms act in a coordinated manner with one another and, in particular, the existence of a "histone code" or "epigenetic code" has been proposed, according to which each histone modifications are recognized by specific domains, part of the remodeling complexes or of transcription factors. (Bird A., 2001; Nakayama et al., 2001).

## DNA methylation

DNA methylation represents the most characterized epigenetic modification; it has been described in several organisms (Goldberg et al., 2007) and is involved in

many different cellular functions including gametogenesis embryogenesis, imprinting, chromosome X inactivation and transcriptional control (Bird, 2002). In mammals DNA methylation has a peculiar distribution, with CpGs dinucleotide preferentially methylated and usually clustered in CpGs rich regions defined as CpG islands. These regions are distributed generally around promoters and first exon, in the unmethylated status the genes associated are prone to be transcriptionally activated, while on the contrary the methylation of the CpG islands is strongly associate with silencing of the corresponding genes (Esteller, 2007). The enzymes responsible for the methylation of DNA strands are called DNA methyltransferases (DNMTs) and in mammals three different classes, DNMT1, DNMT3a, DNMT3b regulate the methylation levels. DNMT1 regulates the maintenance of the methylated status by recognizing hemimethylated DNA, while DNMT3a and DNMT3b are more considered as the de novo methyltransferases (Okano et al., 1999) and their activities are mostly confined to the early stages of development, where they regulate the proper DNA methylation status (Meissner et al, 2008). The differentiation process is the mainly affected by DNA methylation, which in fact is higher in differentiated cells respect stem cells. Moreover the genes mainly subjected to DNA methylation are those involved in the manteinance of pluripotency (Mohn et al., 2008).

## Histone modifications and cross-talk

Among the histone tails lysine and arginine residues are the main sites of modification that principally includes acetylation and methylation. While acetylation of the lysine abolishes the positive charge of the amino acid and eliminates in this way the electrostatic bond between histones and DNA allowing

euchromatin formation, methylation of lysine and arginine does not alter the charge (Copeland et al., 2009). There are numerous chromatin-associated factors that have been shown to recognize and interact with modified histones through many distinct domains. Lysine methylation is the histone modifications with the major number of recognizing domains, reflecting the modification's relative importance. Interestingly the histone modifications not only are recognized from specific domains, recruiting cofactors and create in this way an interaction, but they can also disrupt the interactions between histone and chromatin factors. For example, the presence of H3K4me3 can prevent the binding of NuRD complex at the H3 N-terminal tail, consistent with the fact that NuRD is a transcriptional repressor, while H3K4me3 is a marker of active transcription. (Zegerman et al., 2002)

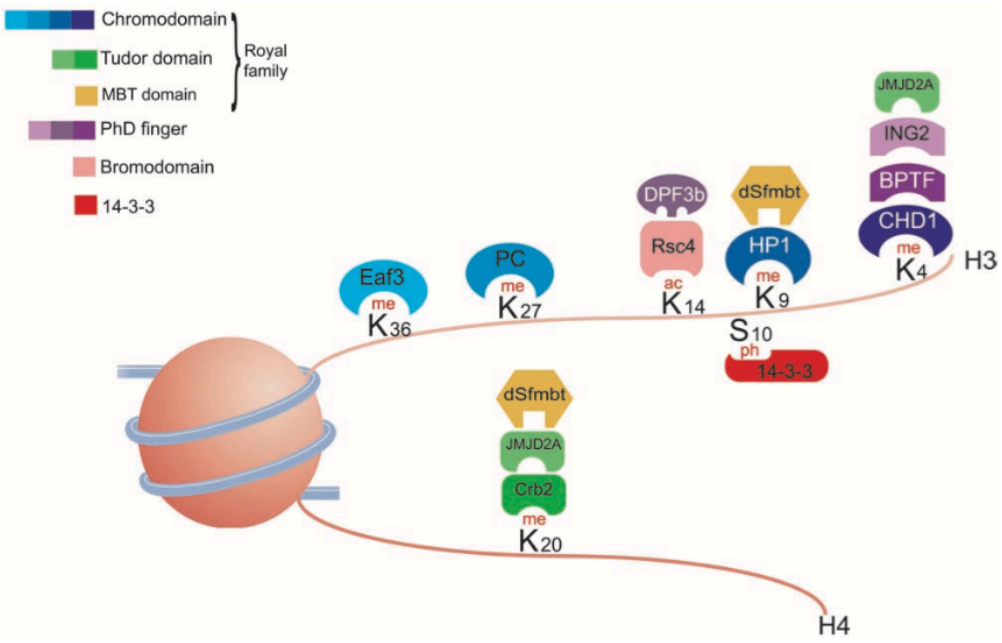



**Figure intro 2.** Examples of proteins and domain that specifically bind to modified histone (Bannister and Kouzarides, 2011)

The large number of possible chromatin modifications and their recognition provide a precise and tight control of chromatin structure. In addiction to these

regulations there is an additional level of complexity and regulation, due to the cross-talk among the different modifications. These cross-talk can occur through multiple mechanism (Kouzarides, 2007):

- Competitive antagonism could be among modifications that occur at the same site;
- one modification may be dependent upon another (Lee et al., 2007; Kim et al., 2009);
- adjacent modifications can disrupt the binding of another modification (Fischle et al., 2005);
- the activity of an enzyme may be affected by the modification of its substrate (Nelson et al., 2006);
- two modifications can cooperate in order to recruit specific factors.

Moreover, there may be also interactions among DNA methylation and histone modifications. These interactions could be cooperative but also inhibitive, as for the KDM2A which binds nucleosomes with H3K9me3 only when the DNA is not methylated (Bartke et al., 2010).

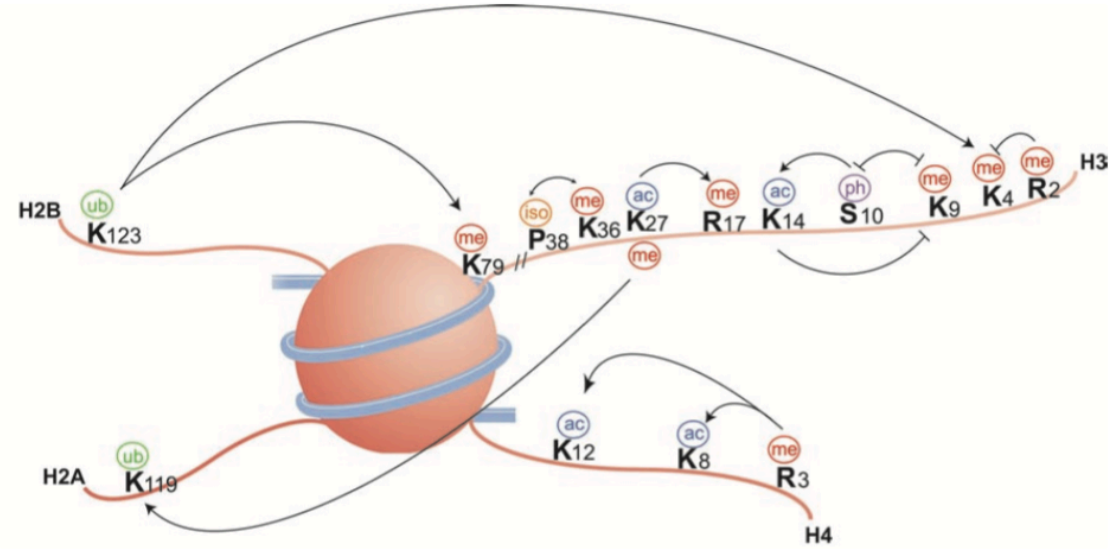



**Figure intro 3.**Histone modification cross-talk: the histone modifications can affect the other adjacent or distant modification in a positive or negative manner. (positive effects are indicated by an arrowhead while negative effects are indicated by a flat head (Bannister and Kouzarides, 2011).

## Histone acetylation: Lysine acetyltransferases and deacetylases

Histone acetylation, first reported in 1964 by Allfrey and colleagues, is highly dynamic histone modification regulated by the opposite action of two families of enzymes: histone acetylation (HATs) and histone deacetylation (HDACs) (Allfrey et al., 1964). The histone acetylases utilize acetyl CoA as cofactor and catalyze the transfer of the acetyl group to the ε-amino group of lysine side chains. This addiction neutralizes the lysine's positive charge, weakening the interactions between histones and DNA. HATs can be divided in two major classes:type-A and type-B. Type-B are predominantly cytoplasmatic and acetylates free histones not already deposited into chromatin, in particular histones H4 at K5 and K12. This acetylation is important for the correct deposition of the histones (Parthun, 2007). The type-A comprises three more classes of HDAC: GNAT, MYST and CBP/p300 (Hodawadekar and Marmorstein, 2007). The enzymes of these classes

function as coactivators and are often associated with large multiprotein complexes (Yang and Seto, 2007). The effects of the HDAC are opposite to HATs, reversing the histone acetylation and restore the positive charge of the histones, stabilizing the chromatin architecture and acting as transcriptional repressor. There are four classes of HDAC: class I proteins are homologous of the yeast Rpd3 and localize into the nucleus; class II can be further divided into two classes, class III proteins are homologous of the yeast Sirt2 and differ structurally from the other classes, requiring NAD+ as cofactors, class IV contains singe HDAC with catalytic domain shared with class I/II. All the HDAC are typically present in multiple distinct complexes, also often with other members of HDAC family.

## Histone methylation: Lysine methyltranferases

The enzymatic methylation of histone is performed by lysine methyltransferases (KMTs) and arginine methyltransferases (PRMTs) with S-adenosyl-L-methionine as the methyl donor and can involve the transfer of up to three methyl groups resulting in mono-, di- or thrimethylates lysine (Martin and Zhang, 2005) or the transfer of one or two groups to the arginine (Greer and Shi, 2012). Since the same modification could also lead to opposite activities, due to recruitment of different enzyme, the presence of a histone code has in the latest years shifted to a "language" which is more dependent on the context. Up to date more than 50 lysine human methyltransferases (KMTs) have been reported and these enzymes are characterized by the high selectivity concerning the histone lysine residue target. Several residues undergo methylation: H3K4, H3K9 and H3K27 among the others. The methyltransferases are divided in three different families: one acts on arginine (PRMTs family) and two on lysine the DOT-like proteins and the SET

domain-containing proteins. The SET proteins methylate lysines in histone as well as non-histone substrates and can be divided into four additional families: SET1, SET2, SUV39 and RIZ, that generally act in multiprotein complexes.

## LSD1 LYSINE-SPECIFIC HISTONE DEMETHYLASES 1A

LSD1 was the first experimental evidence of the existence of enzymatic demethylation and was first isolated as partner of the histone deacthylase HDAC2 in HeLa cells (Tong et al, 1998). Subsequent analysis identified LSD1 in CoREST transcription repressor complex (Humphrey et al., 2001; You et al., 2001; Hakimi et al., 2003), in association with NuRD complex (Wang et al., 2009) and with several others factors. While some of the interactions of LSD1 still remain unclear, the incorporation of LSD1 into protein complex such as CoREST and NuRD is determinant for the LSD1 ability to demethylate nucleosome; in particular, LSD1 requires association with RCOR1 of the CoREST complex or MTA2 of the NuRD complex (Shy et al., 2005; Lee et al., 2005). Moreover some interactions, like LSD1:RCOR1, also prevent LSD1 from proteasomal degradation, while association with PHF21A binds un-methylated H3K4 and leads to a stabilization of the LSD1 with chromatin, promoting the activity of LSD1 as a transcriptional repressor.

The discover of LSD1 opened the way for the isolation of an entire family of demethylases, the JMJ proteins (Kooistra and Helin, 2012). Recently, in human another homolog of LSD1, named LSD2, was characterized, which exhibits poor sequence identity with LSD1 (only 30%) but shared similar domains (Fang et al, 2010). The structure of both LSD proteins includes a N-terminal SWIRM domain and an amino oxidase domain (AO) containing two binding sites, one forms a

non-covalent FAD binding site and the other one forms the substrate binding and recognition site. These two pockets represent the catalytic domain of LSD proteins, for which the enzymatic activity was first demonstrate by Shi and collegues in 2004.

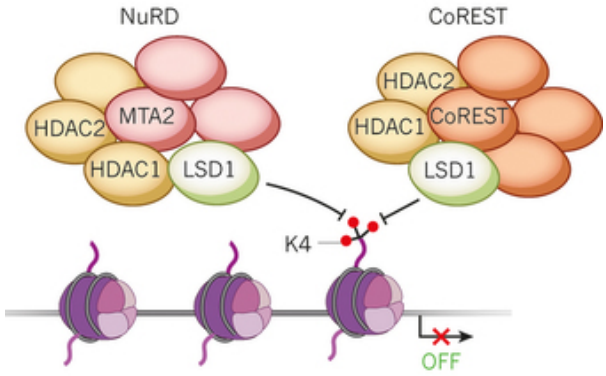



**Figure intro 4.** Histone demethylases LSD1. LSD1 is part of several chromatin complexes, including nucleosome remodelling and histone deacethylase (NuRD) and CoREST, in which it catalyse the demethylation of H3K4me2 and H3K4me1. (Helin K. and Dhanak D., 2013)

The protein acts as demethylase on mono- and di- methyl-lysine 4 of histone H3 (H3K4me1/me2) with production of formaldehyde and reduced form of FAD, while it is not able to demethylate H3K4me3 (Shi et al., 2005). Nevertheless the substrate specificity of the protein is influenced by its association with other cofactors: while generally LSD1 demethylates H3K4me1/me2, leading to transcriptional repression, when it interacts with the androgen receptor (AR), its enzymatic activity switches to H3K9me1/me2 with consequent transcriptional activation (Metzger et al., 2005). Contrary to LSD2, LSD1 contains a TOWER domain hairpin that protrudes from its globular portion and is essential for the interaction with other proteins like Co-REST. Interestingly, interaction of LSD1

with co-factors could affect the ability to demethylase and also the protein stability in vivo (Shi et al., 2005).

LSD1 is highly expressed in many cancer type, including breast prostate, acute myeloid leukemia, lung cancer and neuroblastoma (Lynch et al., 2012); is an essential gene in mammalian biology and many specific roles have been reported. Germline murine knockout shows embryonic lethality before E7.5. Orkin and collegues used in vitro experiement to uncover the role of LSD1 in the multi-lineage hematopoietic differentiation (Saleque et al., 2007); role that has been confirmed by conditional KO/KD murine system, highlightening a clear involvement of LSD1 in several steps of the physiological hematopoietic differentiation (Sprussel et al., 2012; Kerenyi et al., 2013). Conditional KD of LSD1 leads to abnormal expansion of immature progenitors and consequent impairment in most lineage terminal differentiation, a phenomenon that can be reverted by restoring LSD1 expression (Sprussel et al., 2012). Morover, and consistent with these findings, the complete deletion of the gene in hematopoietic compartments leads to alteration of HSC self renewal and impairment of terminal granulocytic and erythroid maturation. Both mRNA and LSD1 protein are highly expressed in undifferentiated human embryonic stem cells, while the expression progressively decreases during differentiation, suggesting a significant role in the differentiation through maintenance of pluripotency by the fine control of H3K4 methylation at genes exhibiting bivalent domains (marked by both H3K4me2/me3 and H3K27me3) such as FOXA2, EOMES and BMP2 (Adamo et al., 2011).

## Pharmacological inhibition of LSD1 and potential clinical use

Given its enzymatic activity and the observation of high-level of expression of LSD1 in many malignancies in the last few years there is a growing interest in developing potent and specific pharmacologic inhibition of LSD1 (Amente et al., 2013). In particular, in neuroblastoma LSD1 is often found overexpressed (Schulte et al., 2009) while the overexpression in prostate cancer serves as a marker for cancer recurrence (Kahl et al., 2006) and poor prognosis in SCLC, where it promotes invasion progression and proliferation (Amente et al., 2013). On the contrary, other studies have proposed LSD1 as tumor suppressor. Wang and colleagues demonstrated the involvement of LSD1 in breast cancer migration and the down-regulation of the protein in breast carcinoma (Wang et al., 2009). The involvement of LSD1 in cellular differentiation has been demonstrated in several studies (Wang et al., 2007; Musri et al., 2010) and its role in hematopoietic differentiation, through the repression of GFI1 targets, was proposed by Saleque and colleagues (Saleque et al., 2007).

Tranylcypromine (TCP: trans-2-phen-ylcyclopropylamine), a non-selective and irreversible monoamine oxidase inhibitor (MAOI), and approved drug for the treatment of depression, was the first drug reported to inhibit LSD1 through an inactivation mechanism involving covalent modification of the cofactor FAD. Because of the relative lack of potency and specificity, together with several side effects, many derivatives of tranylcypromine were subsequently developed. Several reversible LSD1 inhibitors have been reported in the last years, most of which remain at an early phase of development but maintain the potential to establish a novel type of LSD1 inhibitor with alleviation of some of the side effects reported for the tranylcypromine on erythropoiesis.

The tranylcypromine derivative inhibitor GSK2879552 was screened in a panel of 165 cell lines revealing that AML and SCLC lines are particularly sensitive. In vivo the drug effectively prevents growth of xenografted SCLC cell without inducing hematologic toxicities and showing prolonged survival (Mohammad et al., 2015). The inhibitor is currently undergoing evaluation in early phase of clinical trial (phase I) for patients with AML or small cell lung cancer, together with the Orizon compound ORY-1001, which is in phase I/II clinical trial for patient with relapsed or refractory acute leukemia. Also the inhibitor OG86, known as Compound B, is effective in human primary APL cells and cell line, with promotion of differentiation and in vivo blocked of accumulation of blasts in the blood and impaired the proliferative potential of AML cells but not normal hematopoietic stem and progenitor cells (Harris et al., 2012). RN-1 is another tranylcypromine derivative inhibitor which is effective across a large panel of AML and acute lymphoblastic leukemia cell lines, with increasing effects on AML cells harboring the t(8;21) and MLL-rearrangements. The inhibitor completely stops the Kasumi-1 xenograft growth at well-tolerated doses (McGrath et al., 2016).

In AML LSD1 inhibitors were tested also in combination therapies: tranylcypromine or tranylcypromine derivatives have shown to increase the effects with all-trans retinoic acid, potentiating differentiation and loss of clonogenic potential (Binda et al., 2010;Schenk et al., 2012) while the combination with HDAC inhibitor (using vorinostat and tranylcypromine) has also shown a response in glioblastoma cell line and patient-derived xenograft (Singh et al., 2015). Moreover synergistic effects have also been reported between RN-1 and cytarabine and between RN-1 and EZH2 in vivo using AML cell

lines (McGrath et al., 2016). Because of the cross talks between histone modification and the interaction with chromatin remodeling proteins, combination of epi-inhibitors could represent an interesting approach for the future therapeutic interventions. Combination of drugs that altered chromatin or DNA methylation status has already been tested with promising results and synergistic reactivation of tumor-suppressor genes and enhanced anti-cancer effect in several malignancies. The improvement of combination therapies compared to single agent drugs could be the reduction of acquired resistances and the limitation of the side effects through the use of lower dosages of one or both the drugs.

Despite the number of inhibitors and their effectiveness is greatly increased in recent years, the mechanism of action of the inhibitors has not yet been fully clarified and this is the purpose of this thesis, in which we analyze the role of LSD1 in the combined treatment of a novel LSD1 inhibitor and retinoic acid, proposing a mechanism of action that involves changes of histone modifications in regions associated to the protein and changes in gene expression profile.

## HISTONE MODIFICATION AND TRANSCRIPTIONAL REGULATION

In the last years with the increasing number of high throughput data many efforts have been done that aim to integrate different data types, from histone modification to transcriptional data, trying to develop new hypothesis regarding the regulatory functions of all these chromatin features. Integration of histone modification maps with chromatin accessibility, nucleosome positioning, transcription factors binding sites, RNA expression and genome annotation is providing increasingly unified vision of chromatin structure and function.

Although there is a continuous stream of new discoveries, today we are able to characterize certain states of chromatin and their functions by the presence or absence of chromatin remodeling complexes, histone modifications and nucleosome structure. The possibility to map histone modifications at genomic scale and to generate the expression information related to these regions had lead to the identification of chromatin features indicative of both transcribed or repressed genes and to a comprehensive picture of the epigenome (Enrst et al., 2011). The DNA methylation occurs throughout all the genome except for functional regulatory regions, which includes promoter and enhancer. Large heterochromatin domains are associated with the presence of H3K9me2 and H3K9me3 as well as HP1 binding while euchromatic domains are characterized with localized signals of H3K4me and histone variant H2A.Z and H3.3 occurring mainly at functional regions as promoters, enhancer and insulator. Promoter regions, corresponding to cis-regulatory regions, which span the transcriptional start site (TSS) and are necessary to guide the transcription activation, could be active and marked with H3K4me3 or repressed and associated with H3K27 and H3K9 trimethylated marks (Barski et al., 2007). These regions are also depleted from nucleosome as evident from the DNaseI hypersensitivity. Even if H3K27me3 and H3K4me3 seem to be alternatively present in genome regions, they coexist in the early development at genes called “bivalent”, characterized by the presence of both the marker and by the possibility to easily switch from one to another. While the presence of both the markers determines the “poised” state, the loss of H3K27me3 and the acquisition of H3K4me3, or vice versa, determines the respective activation or repression of the genes. The pattern of PTMs associated to enhancer is still not fully elucidate, although seems to include high levels of

H3K4me1 and low levels of H3K4me3, together with p300 association (Heintzman et al., 2007). Also enhancer could be sub-divided in "poised" when presenting H3K27me3, or "active" when associated to H3K27ac (Rada-Iglesias et al., 2011). All these data support the hypothesis that all the chromatin components are not independent elements but they instead can influence one another and that a complex interplay is necessary for a fine regulation needed by the cell. It is not surprising that epigenetic alteration even at single gene level could predispose to disease, including cancer.

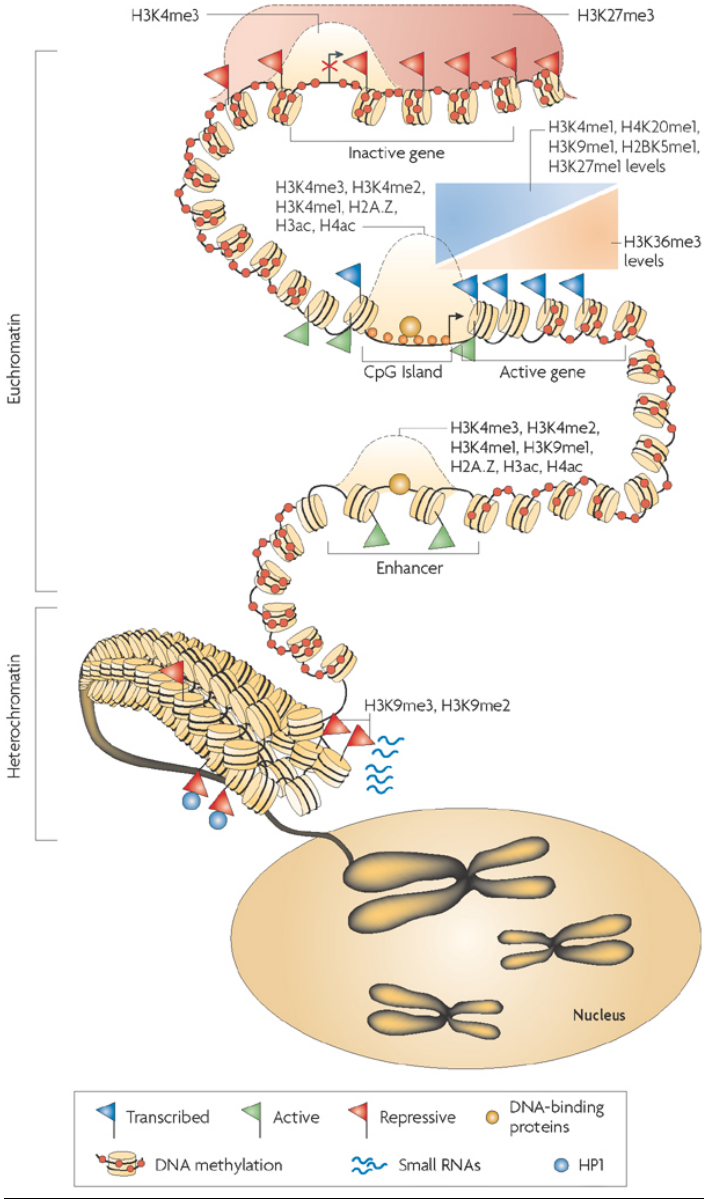


**Figure intro 5**. The interaction of DNA methylation and histone modification together with other regulative factors such as small RNAs constitute the epigenome of the cells which regulates gene expression and allows cells to remember their identity. In general, chromatin is divided in euchromatin (more accessible regions, rich in expressed or poised genes) and heterochromatin (poorly accessible regions enriched in silenced genes). Heterochromatic regions are mainly associated to H3K9 methylation, while methylated form of lysine 4,27 and 36 of the histone are enriched in eurcromatin (Schones and Zhao, 2008).

## EPIGENETIC ALTERATION IN CANCER

Since the epigenetic machinery contributes to regulate most of the cellular functions, it is not surprising that perturbation of the normal epigenome could lead to improper activation or repression of several transcriptional pathways, which could impair both initiation and progression of many diseases, first of all cancer. The mechanisms by which aberrant histone modification profiles, DNA methylation or dysregulated activity of the associate enzyme lead to cancer onset are at least two: the alteration of oncogene or oncosoppressor expression at single genes level or the impairment of the organization of more wide regions which may affect genome integrity and/or chromosome segregation.

A global reduction of DNA methylation characterizes the cancer cell, together with a specific acquisition of hypermetilation at the CpG islands of certain promoter, such as VHL, p16, BRCA and Rarβ, or genes involved in DNA repair, such as MGMT (Esteller and Herman, 2002, Kulis and Esteller, 2010).

Generally, DNA hypomethylation often occurs at repetitive element, impairing the chromosomal stability, or at specific oncogene promoters, while the hypermethylation is observed at genes with oncosuppressive functions, involved

often in DNA repair, cell cycle control and apoptosis and probably causes their silencing (Nguyen et al., 2001; Sadikovic et al., 2008).

In AML, recent genome-wide studies of methylation on 344 AML samples revealed that AML samples could be divided in 16 subclasses according to the methylation signature that are often associated to cytogenetic or molecular characteristics (Figueroa et al., 2010), suggesting that aberrant DNA methylation is very specific and related with driving genetic lesion (Akalin et al, 2012).

Histone modifications are also globally altered in cancer (Baylin and Jones, 2011; Elsheikh et al., 2009). Fraga and colleagues recently demonstrated that global loss monoacetylation at lysine 16 of histone H4 and trimethylation at lysine 20 of histone H4 is a common hallmark of human cancer cells (Fraga et al., 2005). Moreover global histone acetylation has been suggested as independent prognostic factor in several cancer types (Barlesi et al., 2007; Manuyakorn et al., 2010; Seligson et al., 2005).

These alteration in the epigenetic profile are due to different mechanism which include direct genetic alteration of enzyme responsible for the deposition or the removal of the histone modifications (e.g. the translocation of the mixed lineage leukemia MLL which is an histone methyltransferases), or alteration due to abnormal recruitment of histone modifiers (such as the aberrant recruitment of HDAC-cointaining complex and DNMTs on the target genes of RAR due to the PML-RAR alpha (Minucci and Pelicci, 2006).

The involvement of DNA methylation and histone modification in cancer lead in the last year to the approval of the first two drugs (azacitidine and decitabine) inhibiting the DNA methyltransferases enzymes DNMT1 and DNMT3 and

currently used for the treatment of patients with myelodisplastic syndrome (Kantarjian et al., 2007; Issa and Kantarjian, 2009).

Few years later also two histone deacetylases inhibitors (SAHA and romidepsin) were approved for the treatment of refractory cutaneous T-cell lymphoma and others are in various stages of development (Mercurio et al., 2010; Nebbioso et al., 2012).

Although the introduction of these epi-drugs, directly targeting epigenetic regulators, has been a huge success for the field, the clinical results are not satisfying as expected and a large numbers of scientific challenges still remain, most of them due to the fact that the exact mechanism of actions of these drugs is often unknown and to the lack of specificity of the compounds. Moreover the lack of reliable molecular biomarkers for the prediction of clinical resistance of sensitivity is a serious drawback that preclude the correct stratification in the clinical trials. Additional reasons are also the context-specific effects of the epigenetic drugs that act in a different manner on the different tumor cell subtype, like cancer stem cells versus the bulk of the tumor, (Santoro et al., 2013; Shlush et al., 2014) and the use of these epigenetic drugs as single agent when the complexity of the epigenome alteration in cancer requires simultaneous interference.

## GENOME WIDE APPROACHES FOR THE STUDY OF CHROMATIN MODIFICATIONS

Since the discovery of the role of DNA and its structure in 1953 one of the biggest challenge to scientists has been "cracking the code" and deciphering the DNA sequencing in order to assign them the functional role in the cell. The sequencing was the first method through which scientists could actually sequenced the genetic information and its mechanism (Sanger et al, 1992). The Sanger sequencing is based on the selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during in vitro DNA replication. Although this method was widely used as sequencing method for approximately 39 years it is heavily impaired by many limitations such as sequencing speed, scalability and laboratory protocols. The early years of 2000 marked the turning point of the sequencing techniques with the introduction of so-called Next Generation Sequencing (NGS) (Marziali et al., 2001) that introduced a completely new way of sequencing, able to processes millions of bases in the range of hours with high precision level. The technique still relied on the incorporation on nucleotides in a fragment of DNA re-synthetized from a template strand but, contrary to Sanger sequencing, all the NGS techniques rely on light emitted by the incorporates nucleotides, each labeled with a specific fluorescent marker. Since the first appearance in 2004 many new platforms have been released with an exponential decrease of time and cost for each experiment. The affordable price of the technique, the possibility to sequence different experiment in the same run and to adjust the coverage and average length of the reads have made NGS the most widely used technique for any type of genome wide experiment, a type of analysis which in recent years has replaced the gene-centric approach.

The high–throughput methods were then applied to any type of assay, making them genome-wide.

Although chromatin immunoprecipitation has been used since 1988, its combination with next generation sequencing has provide a precise and comprehensive views of transcription factors and histone modification landscape, highlighting roles of chromatin structure across different genomic features.

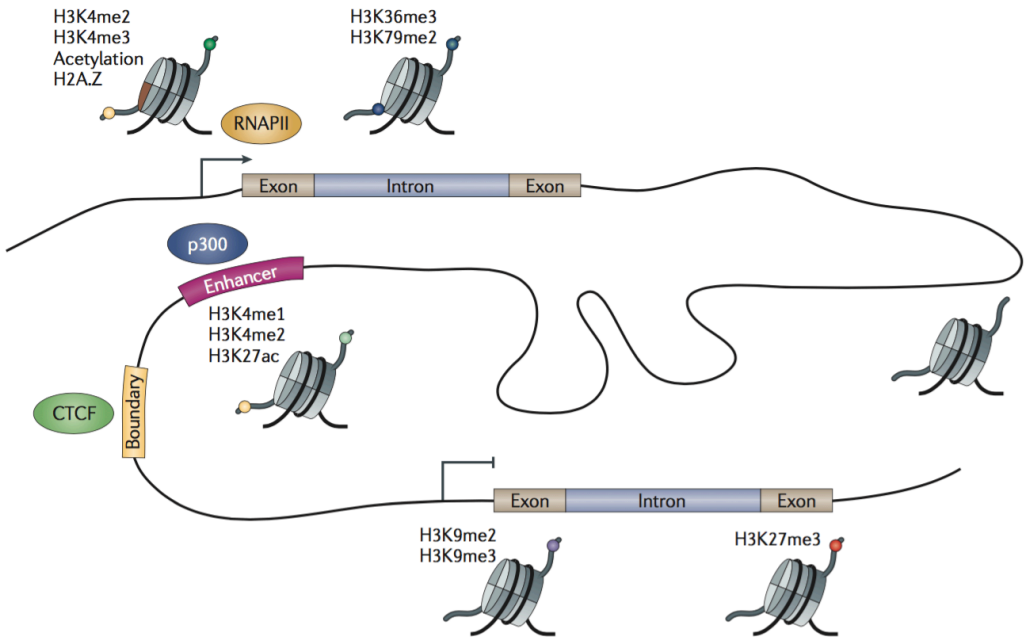



**Figure intro 6.** Histone modifications mark functional elements at the genome. Schematic representation of promoters, gene bodies and boundary element. Active promoters are commonly marked by H3K4me2, H3K4me3, acetylation and histone variant H2A.Z, while the transcribed genes body are marked by H3K36me3 and H3K79me2. Enhancers are relatively enriched for H3K4me1, H3K4me2 and H3K27ac and the histone acethyltransferase p300; repressed regions may be located in large domain characterized by the presence of H3K27me3 and H3K9 me2 and/or me3. (Zhou et al, 2010)

The development of the next generation sequencing has brought also the study of transcriptome to a whole other level, allowing the quantification of all the transcripts present in the cell, and lead to enhancing understanding of

mechanism behind several cell process and human disease. By retro-transcribing RNA to cDNA the NGS method could be used to study the transcriptome with times and cost that continue to decrease since the first usage in the 2007 (Emrich et al., 2007).

## BIOINFORMATICS WORKFLOW

The next generation sequencing and the huge shift in data collection and analysis have introduced new challenges from computational point of view and the NGS era is characterized also from the constant needing of new bioinformatic tools and workflows (Stein, 2011).

### ChIP-seq

The use of specific antibody against protein of interest or histone modification followed by high-throughput sequencing leads in the last years to genome localization of thousand of transcription factors and histone modification. The ChIP-seq process enriches specific DNA sequences cross-linked to the protein or histone marks of interest that subsequently undergo sequencing. The reads obtained from sequencing, after quality score, were aligned to the reference genome. Only parts of the total amount of the reads of the experiment will be uniquely mapped and the percentage of the these uniquely mapped reads varies among organisms; for human above 70% is quite normal, whereas less than 50% could be a problem. A low-percentage of uniquely mapped reads is often due to

excessive PCR cycles of amplification, inadequate sequences length or sequencing platform problems. Nevertheless the use of NGS provides relatively high resolution, low noise and higher genomic coverage compared to ChIP-chip assay (ChIP followed by microarray hybridization) and is nowadays the widely used approach to the study of genome-wide DNA-protein interaction, and mapping histone modification.

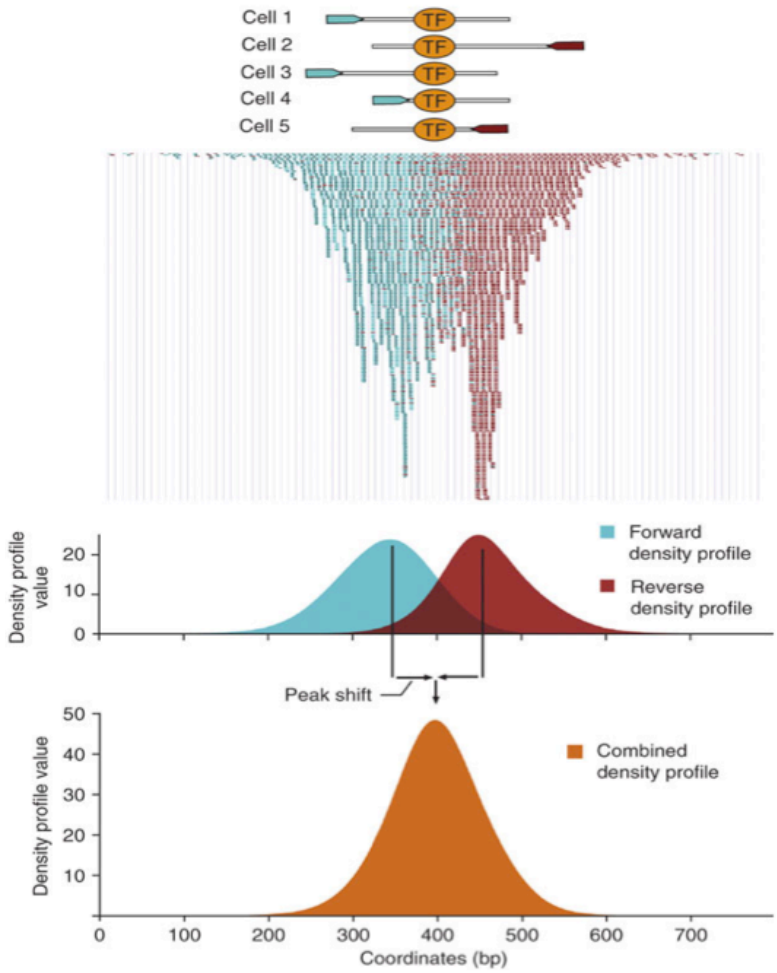



**Figure intro 7.** A ChIP-seq peaks overview (image adapted from http://biocluster.ucr.edu/~rkaundal/workshops/R_feb2016/ChIPseq/ChIPseq.html)

The most discussed part of the ChIP-seq analysis is the identification of true "peaks" in the data, where a peak is a region enriched of mapped reads that produce a pileup. Most of the time the ChIP-seq was performed as single-end and

sequenced from their 5' ends only: two distinct peaks ere generated, one for each strand where the shift between the two peaks represent the exact binding site of the protein. For the peakcalling the use of a control is strictly suggested, and the possible controls used are (i) an input DNA sample, DNA cross-linked and sonicated but non immunoprecipitated or (ii) an IgG "mock" ChIP, using an antibody that will non bind nuclear proteins. Once identified the peaks related to the protein of interest or the histone mark, the aim is to associate the ChIP-seq peaks to functionally relevant regions such as gene promoters, TSS or intergenic regions. This step is called annotation and it is only the first part of all the possible downstream analysis.

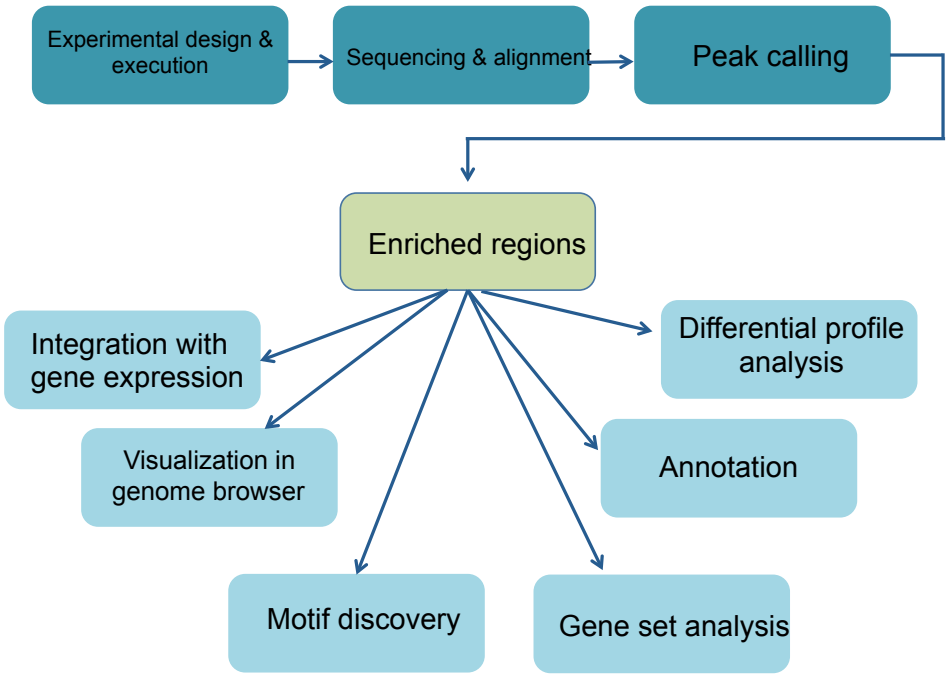



**Figure intro 8.** A typical ChIP-seq data analysis workflow

## RNA-seq

Typical workflow for RNA-seq experiments start with isolation and purification of total or part of the RNA present in the cells, followed with fragmentation and

creation of the libraries with cDNA synthesis. Once the cDNA has been synthetized and sequencing adapters added the sample is almost ready to undergo sequencing and subsequent bioinformatics analysis. The analysis starts with the initial filtering of the sequencing reads, eliminating reads with low base quality confidence, on the basis of the phred quality score (Q) present in the FASTQ file, where the Q values is logarhytmically related to the base calling error probability (P). The reads selected were then assembled into transcripts when the reference genome is not available or aligned to reference genome. When the reference genome is available the first step is the reassignment of the DNA fragment extracted to their original location in the genome, step called "reads alignment". Once aligned, the transcript levels for each genes were normalized to reflect the relative abundance of the transcripts in the starting library, also in function of the transcript length and the total number of mappable reads in the experiment, obtaining the RPKM measure (reads per kilobase of exon model per million mapped reads) which also allows comparison among different experiment (Mortazavi et al., 2008)

$$RPKM = \frac{total\ reads}{mapped\ reads\ (millions) \cdot exon\ length\ (KB)}$$

$$RPKM = \frac{10^6\ C}{NL/10^3}$$

**Figure intro 9.** RPKM formula where C= number of mappable reads into the gene's exons, N=total number of mappable reads in the experiment, L=length of the transcript in bp

Generally the final goal of an RNA-seq experiment is the identification of differentially expressed genes among conditions such as treated vs control,

disease vs healthy, treatment #1 vs treatment #2. The goal of this analysis (known as differential expression, DE) is the identification of statistically significant transcriptional variability among the samples. Since the appearance of the NGS many statistical methods based on Poisson distribution, negative binomial distribution have been proposed and compared and the debate for the correct one is still open.

## ACUTE PROMYELOCYTIC LEUKEMIA

Acute myeloid leukemia (AML) is a genetically heterogeneous disorder, characterized by the accumulation of genetic alteration in hematopoietic stem and/or progenitor cells. The main characterization of AML is a severe block in the differentiation and rapid clonal proliferation and expansion of immature myeloid cells in the bone marrow and peripheral blood. This expansion occurs at the expense of the normal production of their terminally differentiated counterparts.

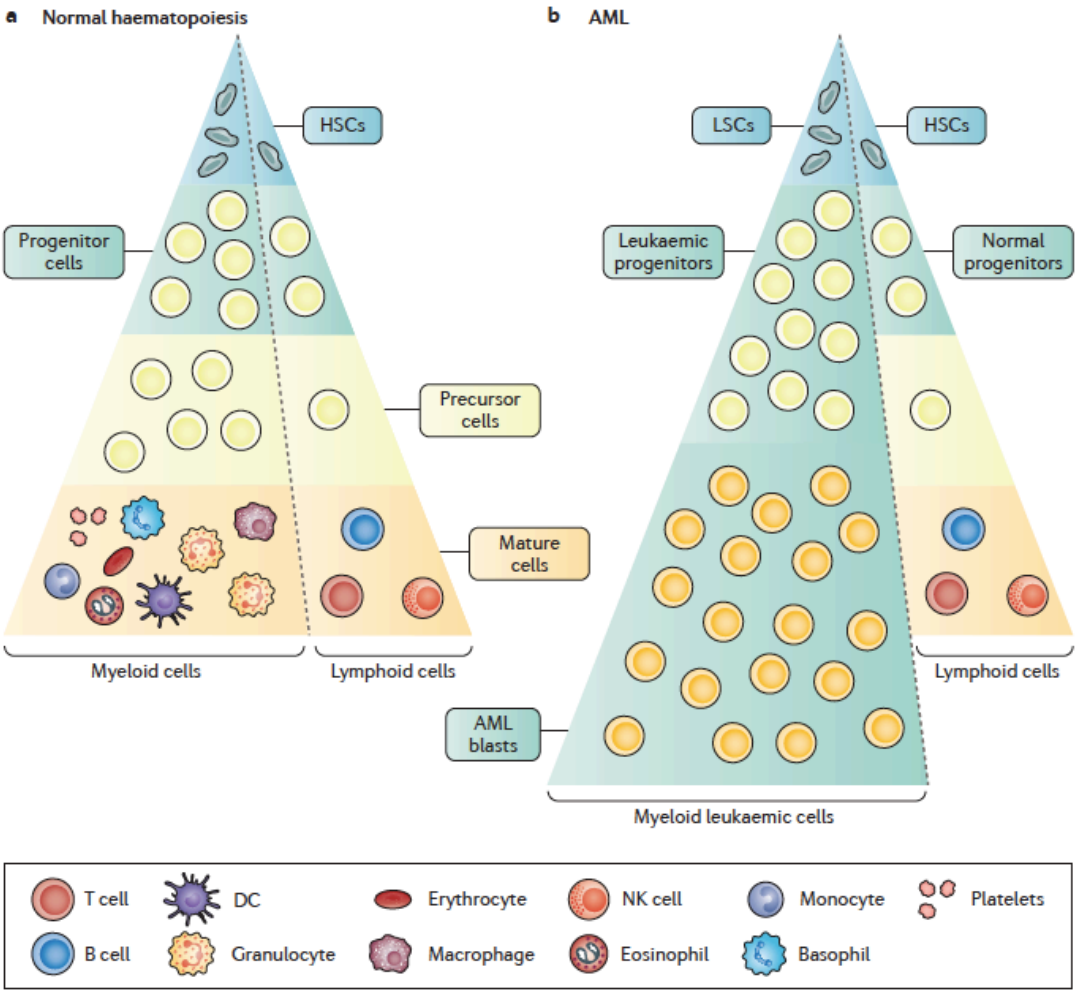



**Figure intro 10. a** The general hierarchical structure of normal haematopoiesis comprises the long-term haematopoietic stem cells (HSCs), with extensive self-renewal potential, that give rise to various haematopoietic progenitor cells. These progenitor cells maintain the proliferative ability but have lost self-renew capability. Progenitors produce various precursor cells and then mature haematopoietic cell types. **b.** Aberrant haematopoiesis in acute myeloid leukaemia (AML). Leukaemic stem cells (LCSs) reside at the top of the developmental pyramid, giving rise to AML progenitor cells and the more mature (but still morphologically primitive) myeloid blast cells that make up the bulk of the neoplasm. DC, dendritic cell; NK, natural killer. (Khwaja et al., 2016)

Diagnosis is based on accumulation of myeloblasts in the bone marrow and blood, while immunophenotyping and cytogenetic and molecular characterization of myeloblasts are used to distinguish AML from other leukemias or to define AML subtypes. The FAB classification (French-American-British) characterized the AML subtypes based mainly on cytological features of

the expanding population. The WHO classification superseded the previous FAB classification and defines seven main subtypes of AML based largely on genetic criteria (such as chromosomal translocation) along with morphological, immunological, cytochemical and clinics characteristics.

Acute promyelocitic leukemia (APL) is a subtype of acute myeloid leukemia, with specific clinical and biological features, characterized by accumulation of progenitors blocked at the promyelocitic state. The genetic hallmark of the disease is the balanced reciprocal translocation t(15;17)(q22;q12), present in about 95% of the patient, involving the promyelocytic leukemia (PML) gene and the retinoic acid receptor alpha (RARa) gene (Minucci and Pelicci, 2007). The result of these chromosomal branches causes the creation of the fusion oncoprotein PML-RAR (PR) (De The, Chomienne et al. 1990), responsible for the block of differentiation of leukemic promyelocytes. In vitro studies demonstrated that PR confers self renewal abilities to murine CMP and GMP progenitors, enhancing their colony forming capacity (Welch et al., 2011) while in mice the expression of the fusion protein leads to APL development, with a intermediate pre-leukemic phase without evident phenotype, suggesting that other mutation are required for the onset of the disease (Di Croce et al., 2002; Grignani et al., 1998).

## PML-RAR oncogenic potential

The fusion protein contains all the functional domains of both the parental proteins and, most important, retains the ability to bind DNA and to recruits a wide spectrum of chromatin modifiers proteins (such as HDAC-containing

complexes, DNMTs, histone methyltransferases, polycomb group proteins) which allow the formation of a condensed chromatin structure, not permissive for the transcription of target genes. The repression of the RAR target genes becomes constitutive in the presence of the fusion protein due to its non-sensitivity to physiological doses of retinoic acid. Pharmacological doses of RA are needed to release the co-repressor and recruitment of co-activators that lead to transcriptional reactivation of the target genes and proteosomal-dependent degradation of the fusion protein. Two different studies demonstrated that the x-RAR/RXR heterooligomerization is required for the X-RAR fusion's oncogenic activity (Zhu et al., 2007; Zeisig et al., 2007). PML-RAR shows more relaxed DNA binding properties than wild type RAR. ChIP sequencing for PR and RXR in primary APL cells provides the evidence that PML-RAR/RXR oligomers interact on a large portion of genomic regions (Martens et al., 2010). Moreover another genome wide distribution analysis reveals that PR is recruited to non-canonical RAR binding motif, highlighting an increased pattern of target genes respect to RAR (Mikesh et al., 2010).

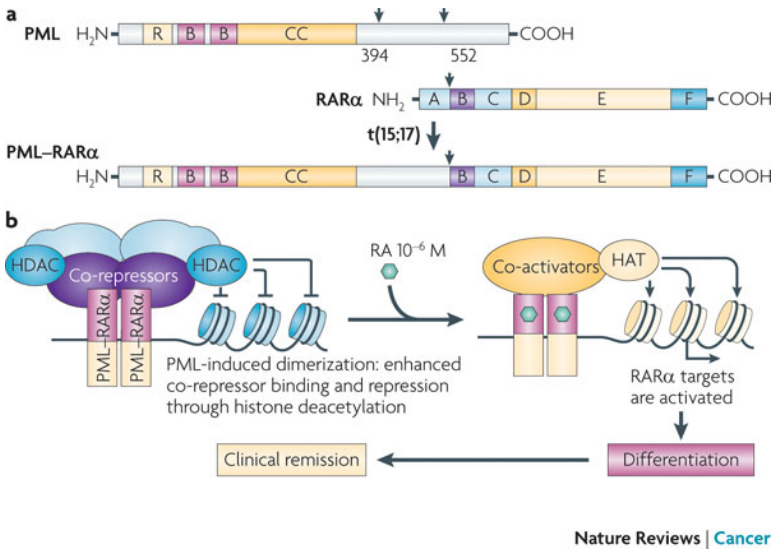


**Figure intro 11. a.** Structure of the promyelocytic leukemia (PML) and the retinoic acid receptor-a (RARa) proteins. The PML-RARa fusion protein comprises the RING (R) domain, the B boxes (B) and coiled-coil domain (CC), RARa DNA-binding domain (C), hormone-binding domain (E) and other regulatory domain. The fusion protein retains the functional domains of both proteins allowing dominant-negative activities on both PML and RARa. **b.** The PML-RARa homodimers bind and repress RARa target genes also through recruitment of co-repressor. Retinoic acid (RA) at pharmacological doses allows fusion protein degradation, recruitment of co-activators and restores differentiation (De Thé and Chen, 2010)

## APL treatment

First studies for the APL treatment with all-trans retinoic acid (ATRA) began in the 1980s and demonstrated that ATRA induces terminal differentiation towards granulocytes and mature granulocyte-like cells, with subsequently apoptosis in vitro e in vivo (Wang ZY and Chen Z, 2008). Thus ATRA treatment for APL was the first successful differentiation therapy in human cancer, although a high number of patients show ATRA resistanse relapsed APL after the treatment. The treatments with RA alone generally resulted in transient disease clearance but only few definitive cures, while the combination of RA with anthracyclines was later able to cure more than 70% of the patients. Then, after the introduction of arsenic trioxide as a second-line or consolidation agent in 1994, the rates of patient that reached complete remission increased to 95% (Sanz and Lo-Coco, 2013). Both ATRA and arsenic target the stability of PML-RAR through two different mechanisms: ATRA induce PML-RAR degradation via its RARA moiety, while arsenic acts on the PML part of the fusion protein (Zhang et al., 2010). In the last years, thanks to transplantation experiment it has become evident that all the leukemia cells are not equal and that there is a hierarchy in myeloid leukemia as in normal hematopoiesis, opening the way to the new concept of LICs, leukemia initiating cells, a subpopulation of cells able to reconstitute the disease

in a second recipient mice. Thus, the eradication of LICs has been proposed now as the new goal for the eradication of the disease. Recent study demonstrate that treatment with low doses of retinoic acid leads to differentiation of the cells, while treatment with arsenic leads to loss of LICs but only partial differentiation. Instead high doses of retinoic acid or low doses combined with arsenic lead to loss of the leukemia initiating cell, differentiation and eradication of APL (Ablain and De Thé, 2011). Moreover new study suggested that the transcriptional activation of PR target genes and degradation of the protein are not associated: the transcriptional activation seems to be associated with differentiation activity while the degradation of the protein seems to be responsible for the LIC clearance (Nasr et al., 2008; Ablain et al., 2013). The evaluation of the combined therapy RA/arsenic is underway while alternative therapies are under investigation. Since the importance in leukemogenesis of epigenetic enzyme and histone modifications, new strategies point to epigenetic target as a the strategy for the development of new compound that alone or in combination can be added to the standard therapy of APL (Mercurio et al., 2010).

# AIMS

Taking advantage of a new LSD1 inhibitor showing high specificity and acting at relatively low concentrations (Binda et al., 2010) we've been able to investigate the role of LSD1 in APL. We use as model system NB4 cells, a cell line derived from APL patient, which recapitulate numerous characteristics of APL blasts (Lanotte et al, 1991). The expression of the PML-RAR fusion protein makes the cells insensitive to physiological doses of retinoic acid (RA low doses) while higher concentration (RA high) causes the growth arrest and terminal differentiation.

The thesis started from the initial observation that treatment with LSD1 inhibition sensitizes NB4 cells to physiological doses of retinoic acid which impairs cell growth and triggers cell to differentiation with an effect event stronger compared to RA high alone (Binda et al., 2010).

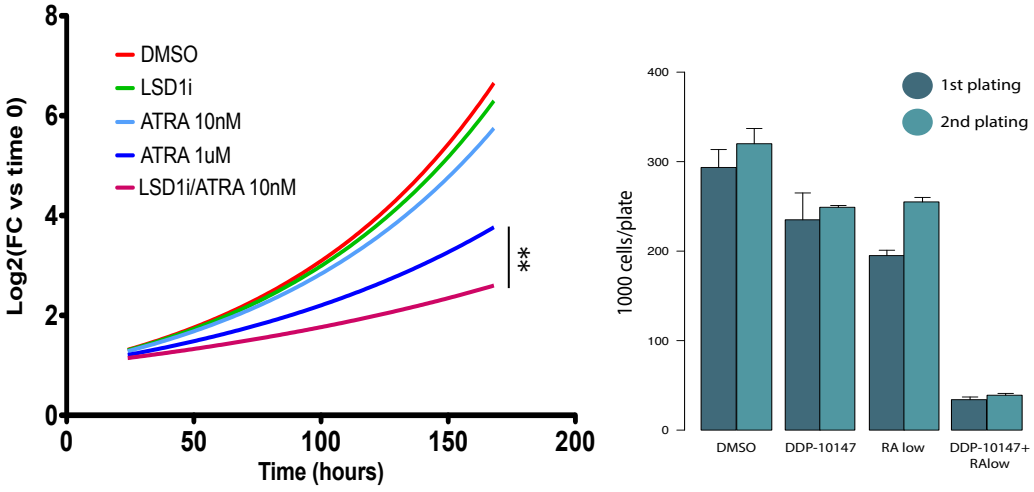
To investigate the role of LSD1 and its molecular mechanism in this phenomenon of sensitization we performed genome wide profiling of the protein and of both transcriptional and epigenetic changes accompanying the LSD1 inhibition and the other treatment (DMSO as control, retinoic acid at low and high concentration, and combination of LSD1 inhibition and RA low). Moreover we investigate the interaction between LSD1 and the fusion protein PML-RARalpha.

The thesis aims to validate the role of LSD1 in the differentiation of APL, identify the transcriptional pathways involved, characterized the interplay between LSD1 and the fusion protein and define the molecular mechanism and the chromatin changes caused by LSD1 that mediate the differentiation.

# RESULTS

## LSD1 INHIBITION SENSITIZES APL CELLS TO PHYSIOLOGICAL CONCENTRATIONS OF RETINOIC ACID

We previously demonstrated that inhibition of LSD1 sensitizes APL-derived cell line NB4 to retinoic acid treatment and induces cell growth arrest and differentiation when combined with physiological concentration of retinoic acid (Binda et al. 2010). We found that, while NB4 cells are not sensitive to LSD1 inhibition or to physiological doses of retinoic acid, the combination of LSD1 inhibition and retinoic acid at low concentration (from now on: cotreatment) significantly reduces the cell proliferation in liquid culture and colony forming ability in semi-solid cloture, with an effect even stronger than the retinoic acid at high concentration (Figure 1 a-b).



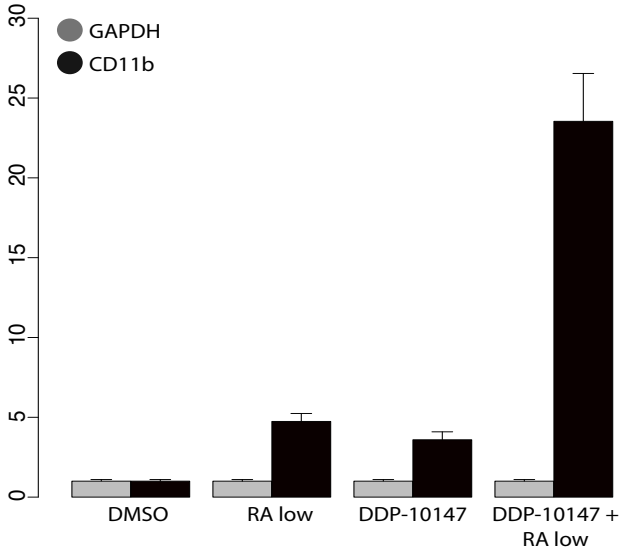



**Figure 1: a.** Growth curve of NB4 cells treated as indicated with LSD1 inhibitor DDP_10147 and/or retinoic acid 10nM and 1uM. Slopes were found significantly different: the comparison between LSD1+RAlow and RAhigh was significant at p-value of 0.0028 **b.** Colony forming ability, scored after 7 days, of 1000 NB4 cells plated in methylcellulose medium and treated with LSD1 inhibitor DDP_10147 and/or retinoic acid 10nM and 1uM. Mean and standard deviation of three independent experiments are shown **c.** Analysis of CD11b mRNA levels in NB4 cells treated as described for 24 hours. Fold changes were normalized against DMSO and GAPDH (used as housekeeping). Graph represent the mean and standard deviation of three independent experiments.

To assess whether the observed phenotype was due to induction of myeloid differentiation, we checked the CD11b marker level by qPCR (Figure 1c). The histogram in figure 1b clearly shows an increase in the CD11b level after 24 hours of treatment with LSD1 inhibition plus retinoic acid at low concentration. These data together with the strong differentiation-associated morphological changes observed, demonstrate a clear synergistic effect between the LSD1 inhibition and the physiological concentration of retinoic acid in the induction of differentiation and growth arrest of NB4 cells, even stronger than the effect induced by pharmacological doses of retinoic acid.

## A TRANSIENT WAVE OF LSD1 INHIBITION IS SUFFICIENT TO INDUCE APL CELL SENSITIVITY TO PHYSIOLOGICAL RETINOIC ACID CONCENTRATION

We perform washout experiment to define the temporal window after LSD1 inhibition sufficient to commit NB4 cells to differentiation in the presence of physiological concentration of retinoic acid. NB4 cells were treated with LSD1 inhibition and retinoic acid low, then the inhibitor was removed at different time point while retinoic acid was kept continuously in the medium for 96h. Short periods of LSD1 inhibition (6h and 12h) were not sufficient to trigger differentiation of NB4 cells while treatment with LSD1 inhibitor for 24 hours in presence of retinoic acid at low concentration leads to a response comparable to 96 hours of continuous co-treatment (Figure 2b). For these reason we selected 24h time point as the treatment condition for all the following experiments.

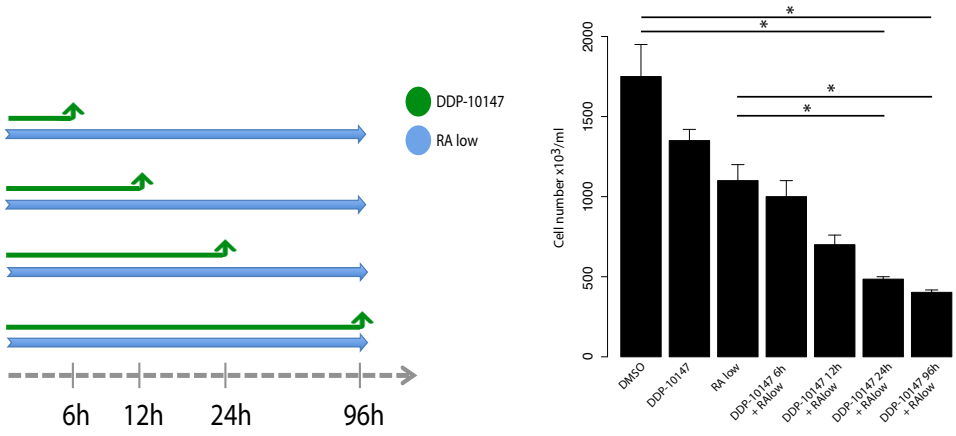



**Figure 2: a.** Schematic representation of wash out experiments. NB4 cells were treated with RA low for 96h and DDP-10147 at 2 µM for 6h, 12h, 24h and 96h. The inhibitor has been washed out after the indicated time-points and was used at the concentration of 2 µM, in which has been shown to reach the maximum target modulation (Binda et al, 2010). **b.** NB4 Cells were counted at 96h. 24h of LSD1 inhibition is sufficient to induce similar growth arrest to the longer

treatment where the inhibitor is kept in the medium for 96h. Error bars represent SD of four independent experiments.

### COTREATMENT WITH LSD1 INHIBITION AND PHYSIOLOGICAL CONCENTRATION OF RA ACTIVATES THE TRANSCRIPTIONAL DIFFERENTIATION PROGRAM OF NB4 CELLS

Given the results of the washout experiment, we determine the gene expression profile of NB4 cells after 24 hours of treatment with DDP-10147 and/or RAlow by RNA-Seq.

Since LSD1 has been generally associated with transcriptional repression, we first assessed if LSD1 target genes were relatively less expressed compared to the LSD1 unbounded ones. Unexpectedly we observe the same transcriptional level of the LSD1-target genes (identify by LSD1 ChIP-seq, see next paragraph) compared to the non-target ones (Figure 3) both considering up-regulated genes and down-regulated genes.

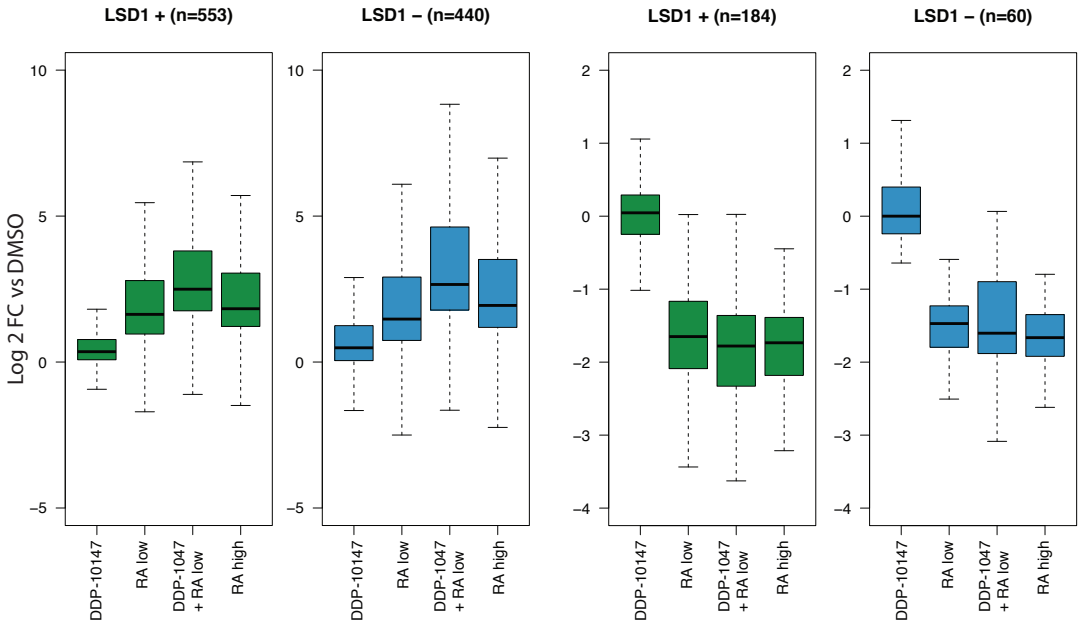


**Figure 3**: Boxplots indicate expression levels in Log2 fold-change of FPKM respect DMSO of genes up-regulated (**a**) or down-regulated (**b**) in at least one treatment. Genes are divided in LSD1 positive (green) or negative (blue) based on peaks annotation (peak within 22kb upstream or in the gene body)

Since there were no differences in the expression of LSD1 target genes and non-target genes, the subsequent analysis has been made without further subdivision.

Global analysis of the differentially expressed genes (Figure 4a) revealed that overall the majority of the regulated genes are up-regulated respect to DMSO and are differentially modulated among the different treatment (Figure 4b), while down-regulated genes show a constant trend both in the number and in the intensity of modulation after the different treatments, except for the treatment with inhibitor that shows only one down-regulated gene. For this reason we decide to focus our analysis on the up-regulated genes.

In order to better understand the global modulation in the genes expression we perform a heatmap showing all the genes regulated in at least one of the four treatments (Figure 4c, n= 1235). The heatmap highlight the association of LSD1 with almost all the regulated genes (green flag on top of the heatmap) and a clear clusterization based on transcription level among the treatments: while LSD1 inhibition has a modest impact on the global gene regulation and is almost comparable to DMSO, treatments with retinoic acid at low and high concentration and cotreatment with LSD1 and RAlow show a similar pattern of expression. However cotreatment shows a clear distinct pattern of expression compared to all the other treatments and includes almost all the regulated genes. It is noteworthy that the two treatments (cotreatment and RA high) leading to cell differentiation, share many of the upregulated genes. However it is also clear that

the cotreatment leads to an upregulation of a subset of genes, which are not activated after treatment with retinoic acid at high concentrations.

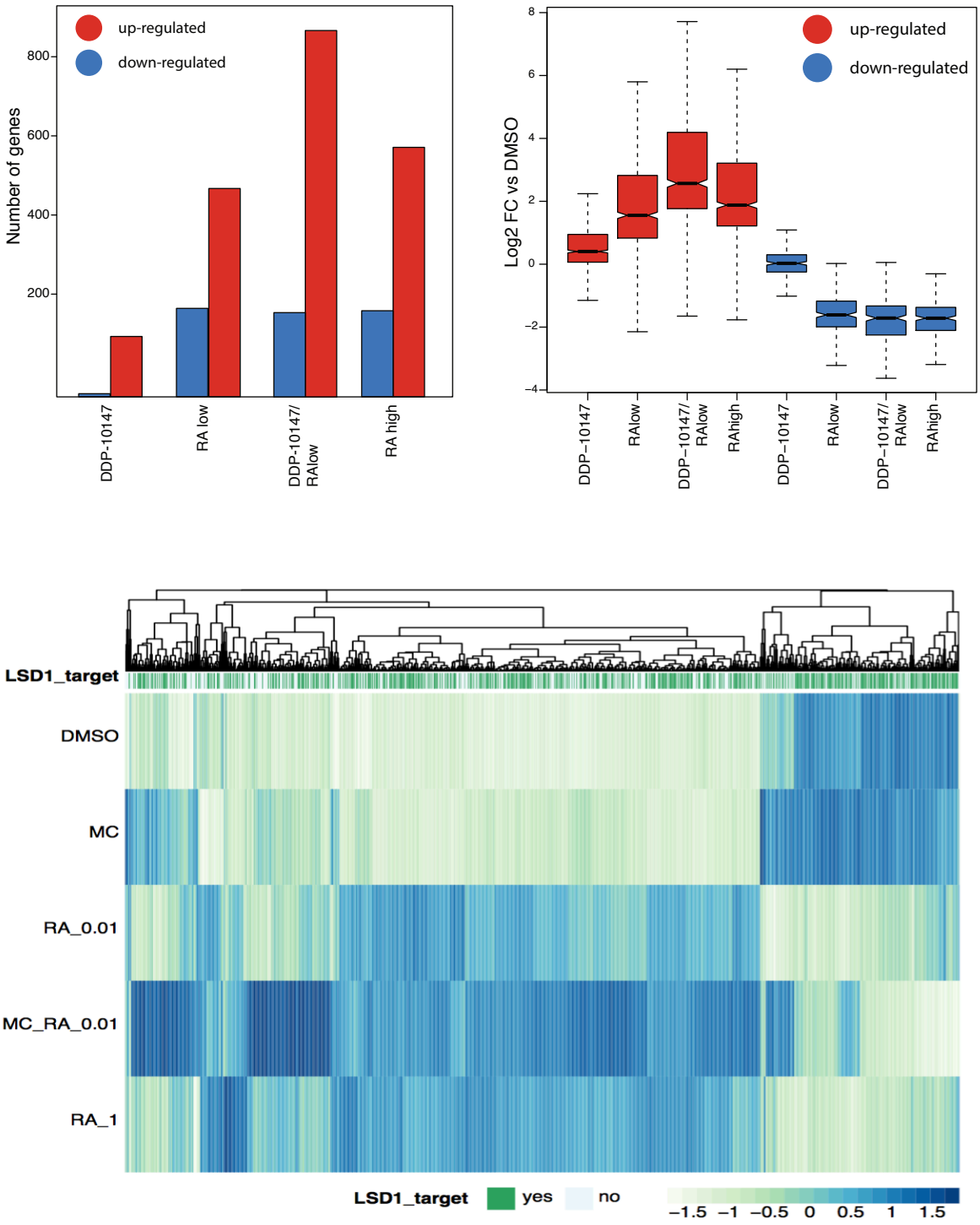



**Figure 4: a.** RNA sequencing was performed in NB4 cells treated with DDP-10147, Ralow, Rahigh and DDP-10147+Ralow for 24h. Barplot represents number of genes regulated (up or down regulated respect to control DMSO,

RPKM>0.5, absolute (FC)>1.5, FDR<0.5) upon the indicated treatments. **b.** Boxplot shows degree of induction by the indicated treatment vs control (DMSO). **c.** Heatmap shows clusterization of gene regulated in at least one of the four treatments against the DMSO. The expression levels range from light blue (less expressed) to dark blue (more expressed) while the green flag represent the LSD1 binding.

Contrary to what happens with LSD1 inhibition alone, that has very modest effect on transcription both considering number of up-regulated genes and magnitude of their regulation, treatment with the inhibitor greatly potentiates the effect of physiological concentration of retinoic acid, increasing the number and the magnitude of the genes regulated by RA alone. Consistently with biological data that indicate a more efficient effect of the combination in arresting cell proliferation, the transcriptional changes induced by cotreatment are comparable and even stronger than those induced by RA high.

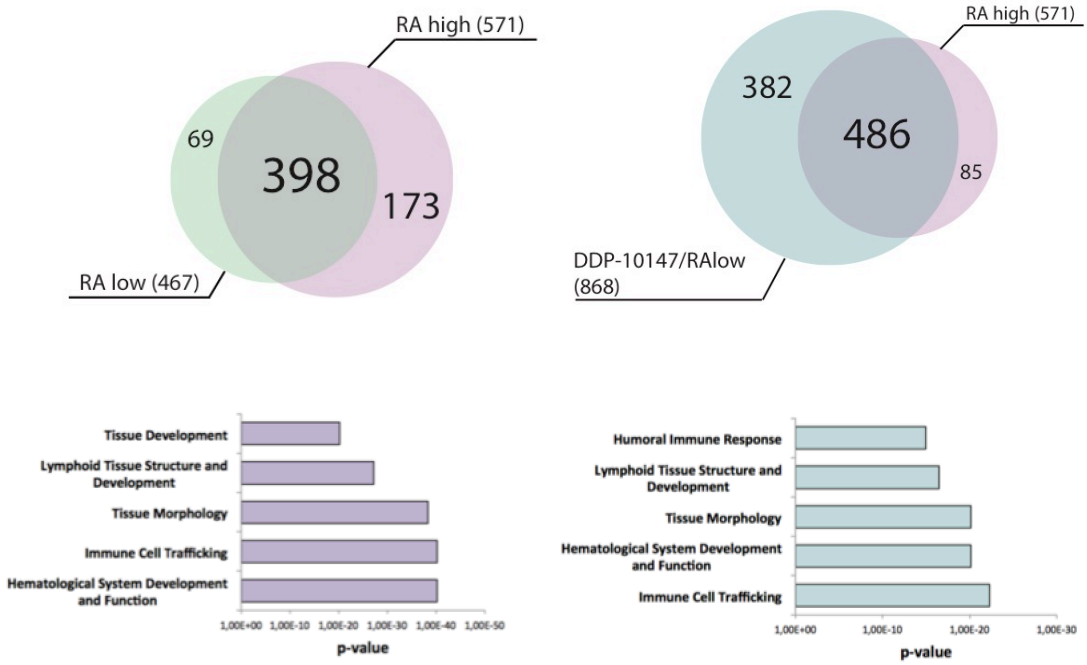



**Figure 5:** Venn diagrams illustrate the similarity of genes expression changes caused by RAhigh and RAlow **(a)** and cotreatment and Rahigh **(b)**. Genes were considered upregulated if RPKM>0.5, absolute (FC)>1.5 respect to DMSO, FDR<0.5. **c.** Ingenuity Pathway Analysis of genes regulated by both RA high and cotreatment (486 genes, violet) and cotreatment-only (382, lightblue)

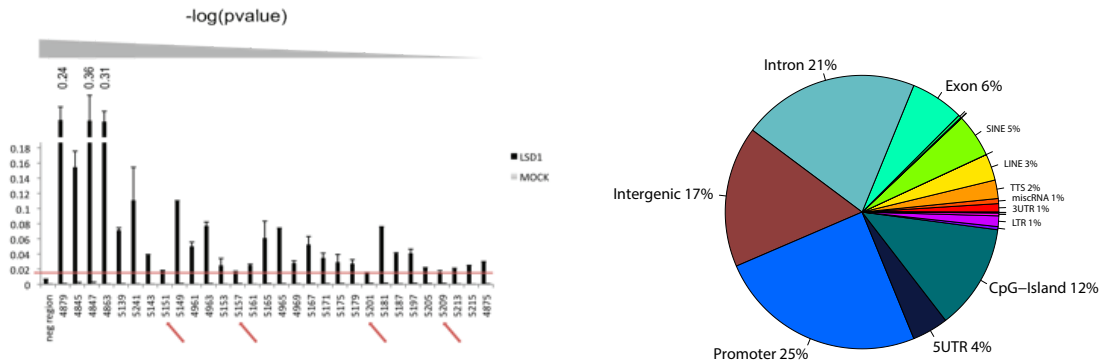
To further explore the transcriptional induction effect of the different treatment we compared the genes upregulated among the different treatment. Almost all the genes upregulated by retinoic acid at low doses are also upregulated by RAhigh (Figure 5a), showing that RAlow starts the induction of genes involved in cell differentiation but not sufficiently to activate the differentiation program. Importantly, in the same manner almost all the genes upregulated by RAhigh are upregulated also by cotreatment (Figure 5b), and this overlap includes also 107 out of 173 of the genes induced by RAhigh but not by RAlow and probably responsible for the differentiation. In addition to the overlapped genes with RAhigh the cotreatment regulated also a specific subset of genes (n=382). This observation suggests that the differentiation program induced by pharmacological doses of retinoic acid is the same one induced by cotreatment that in addition activates another subset of genes cotreatment-specific that could explains the stronger biological effect observed.

Gene ontology performed on overlapping genes (n=486) and combination treatment-specific genes (n=382) revealed among the top scoring network enriched hematological system and function, tissue morphology and immune cell trafficking for both the subset of genes (Figure 5c).

### **LSD1** GENOMIC DISTRIBUTION

In order to explore the role of LSD1 we decided to assess the genomic distribution of LSD1 and to correlate the obtained results with transcriptomic profiling and with the histone modifications associated with its enzymatic activity. Thus we performed ChIP-seq for LSD1 in untreated (DMSO) NB4 cells.

After the peakcalling we perform q-PCR on several regions in order to set a threshold for the identification of high confidence peaks (Figure 6a). Using the threshold of -log10(p)=16.9 we validated 15.187 regions of LSD1 binding. Peaks annotation performed with GIN (Cesaroni et al., 2008) revealed that 45% of the peaks are proximal to the TSS (considered as +/- 2.5 Kb) of annotated genes and about 35% lay in intergenic and intronic regions (Figure 6b), possibly representing regulative regions and consistent with other finding in mESC (Whyte et al, 2012). This proportion in the genomic annotation is not affected by the peaks score and remain constant among highly enriched peaks and the rest of the peaks, highlighting that LSD1 more enriched peaks are not preferentially bind to specific genomic location (Figure 6c). Moreover we could identify two different types of LSD1 peaks, sharp and multiple or broad, with majority of broad peaks located on TSS proximity (Figure 6d). Nevertheless both the broadness of the peaks and the number of sharp peaks associated to genes are not related to change in genes expression (data not shown).



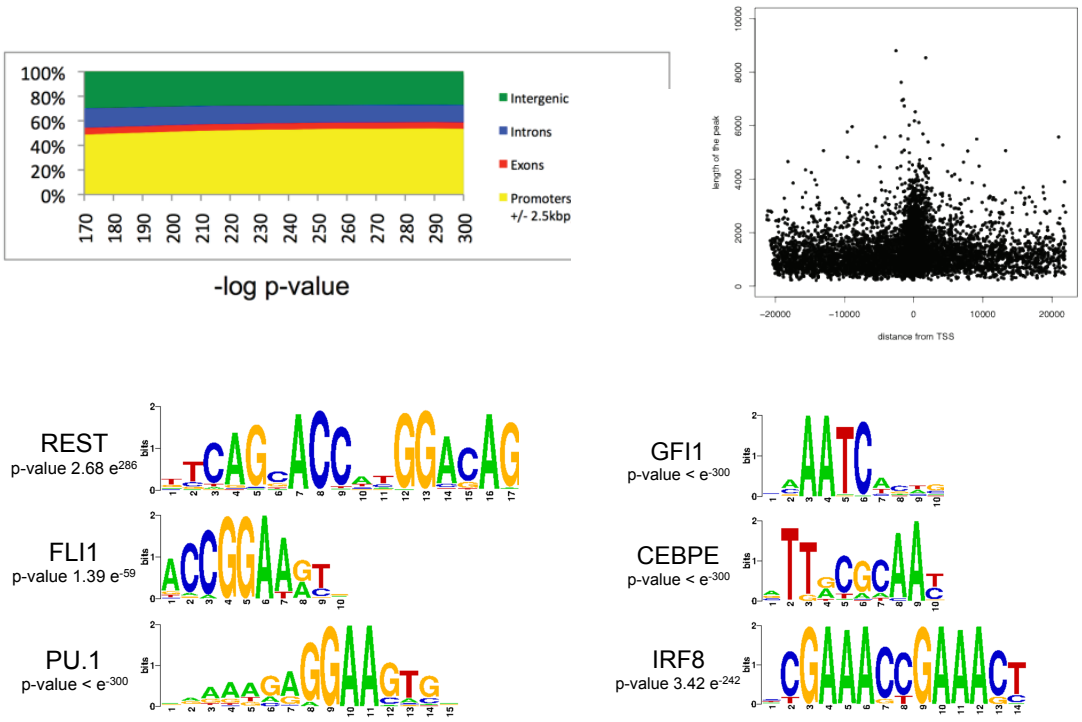



**Figure 6: a.** Validation by ChIP-qPCR assay of LSD1 positive regions at several p-values (obtained with MACS peakcalling) in three independent experiments. First region represent an intergenic region negative for LSD1 used as control, anti-IgG antibody was used as mock control. The red line was set as the enrichment of negative control plus three times the relative SD. 26/30 regions were considered as validated. (The four arrows point the 4 regions not validated). **b.** Genomic distribution of LSD1 validated peaks. Regions spanning around the TSS (+/- 2.5 kb) were considered as promoters, while intergenic regions correspond to regions more than 22kb distant from the nearest gene **c.** Genomic distribution of LSD1 peaks across the different p-value score. **d.** Plot indicates variation of the peaks length among the different distance from TSS. **e.** PSCAN motif discovery on genomic regions bounded by LSD1.

Since LSD1 has been shown to cooperate with several transcription factors (TFs) in hematopoiesis (Saleque et al., 2007) we performed transcription factor binding sites enrichment analysis on the LSD1 bound regions. The analysis was performed using PSCAN-ChIP algorithm (Zambelli et al., 2013) in order to find over-represented transcription factor binding site. The algorithm scans the given LSD1 peaks coordinates and looks for over represented sequence motif, based on motif descriptors of the TRANSFAC (Matys et al., 2003) and JASPAR database

(Portales-Casamar et al., 2010). Results of the scanning on regions bound by LSD1 revealed a large number of transcription factor involved in granulocytic and monocytic differentiation as PU.1, FLI1, GFI1, CEBPE. Moreover among the significantly enriched matrices we found also REST, a known LSD1 recruiter (Mosammaparast and Shi 2010), IRF8 and the canonical PML-RAR binding sequences (RAR:RXR DR5) suggesting an interaction between PML-RAR and LSD1, at least in certain genes and confirm the key role of LSD1 in hematopoietic regulation.

### LSD1 INHIBITION CAUSES A GLOBAL INCREASE IN THE HISTONE H3K4 DIMETHYLATION

Since from RNA-seq analysis we observed that LSD1 inhibition alone does not dramatically alter transcription of NB4 cells but has a great impact in combination with retinoic acid, we reasoned that the contribution of LSD1 inhibition should be investigated in the chromatin alteration occurring after the treatment. For this reason we performed ChIP-seq for H3K4me1/me2/me3 and H3K27 acetylation after LSD1 inhibition, RAlow, combination of the two treatments and RAhigh. Peakcalling analysis respect to DMSO revealed a significant increase of the number of the H3K4me2 peaks in regions bounded by LSD1 after inhibition while the other histone modification are only little affected by the treatment. Similarly, we observed a significantly increase of the number of H3K27ac regions after RA low treatments. Only after treatment with combination of LSD1 inhibition and physiological doses of retinoic acid we observed an increase of both the histone marks H3K4me2 and H3K27ac (Figure 7). These observations and the biological results that individuate cotreatment as the only

one with cell differentiation, led us to hypothesize that the individual contribution of the two treatments, increase of H3K4me2 by the inhibitor and increase of H3K27 by retinoic acid, which individually has no effect on transcription nor on differentiation, once combined, leads to the activation of the differentiation program in NB4 cells. Thus the presence of both H3K4me2 and H3K27ac could be a prerequisite for the transcriptional activation and cells differentiation.

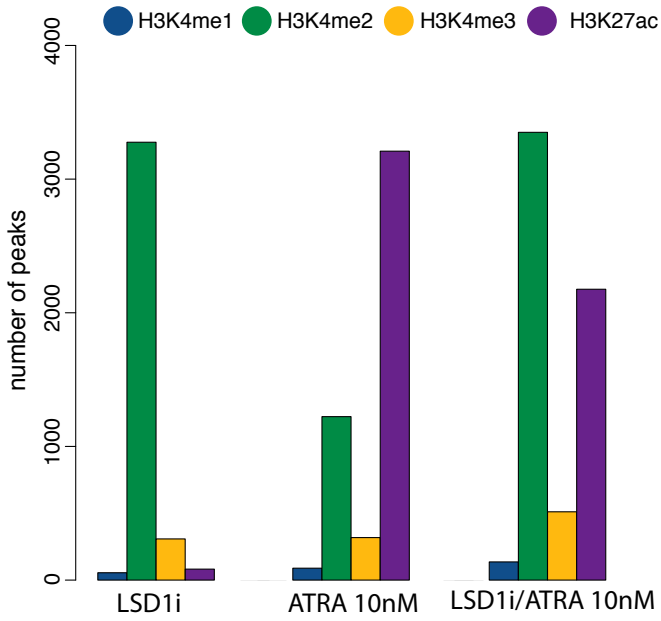



**Figure 7:** ChIP-seq of indicated histone modifications were performed in NB4 cells treated with DDP-10147, Ralow, Rahigh and DDP-10147+Ralow for 24h. Barplot represents number of regions (peaks) increasing in mono-, di-, tri-methyl K4 of histone H3 and acethylation of K27 of histone H3 peaks in LSD1 positive regions.

To further explore the role of LSD1 on the chromatin changes and to correlate the obtained results with the transcriptomic profiling and cells differentiation we decided to individuate the subset of genes cotreatment-specific, and thus differentiation-related genes, and then explore the chromatin associated to this particular set of genes. We divided the up-regulated genes into classes based on the transcriptional induction after the different treatments, obtaining 10 classes of possible combinations of the three treatments (Table 1).

| | DDP-10147 | RA low | DDP-10147/RA low | # genes | % RA high overlap |
|---|---|---|---|---|---|
| **CLASS 1** | ns | ns | ↑ | **254** | 16.9 |
| **CLASS 2** | ns | ↑ | ↑ | **335** | 91.3 |
| **CLASS 3** | ns | ↑ | ↑↑ | **45** | 95.6 |
| **CLASS 4** | ns | ↑ | ns | **33** | 27.3 |
| **CLASS 5** | ↑ | ns | ns | **27** | 11.1 |
| **CLASS 6** | ↑ | ns | ↑ | **30** | 6.7 |
| **CLASS 7** | ↑ | ns | ↑↑ | **13** | 38.5 |
| **CLASS 8** | ↑ | ↑ | ns | **0** | 0.0 |
| **CLASS 9** | ↑ | ↑ | ↑ | **5** | 100.0 |
| **CLASS 10** | ↑ | ↑ | ↑↑ | **12** | 100.0 |

**Table 1**. Number of genes regulated inside each class. One row indicates genes upregulated respect DMSO (RPKM>0.5, Log2(FC vs DMSO)>1.5), two row indicate further increase respect one row to Log2(one row)>1.5. the last column indicates the percentage of genes in each class upregulated also after RAhigh treatment.

Most of the classes are poorly represented while the majority of the gene is included in the classes from 1 to 3. We individuated a subset of 254 genes named class 1 as cotreatment-specific genes, up-regulated only after treatment with both

LSD1 inhibition and RAlow. To investigate the chromatin changes related to LSD1 in this class each gene was associated to LSD1 peak (or peaks) and then reads coverage for all the histone marks were calculated inside the windows delimitated by the LSD1 peaks (Peaks were considered associated to gene if there are no more than 22kb of distance between them).

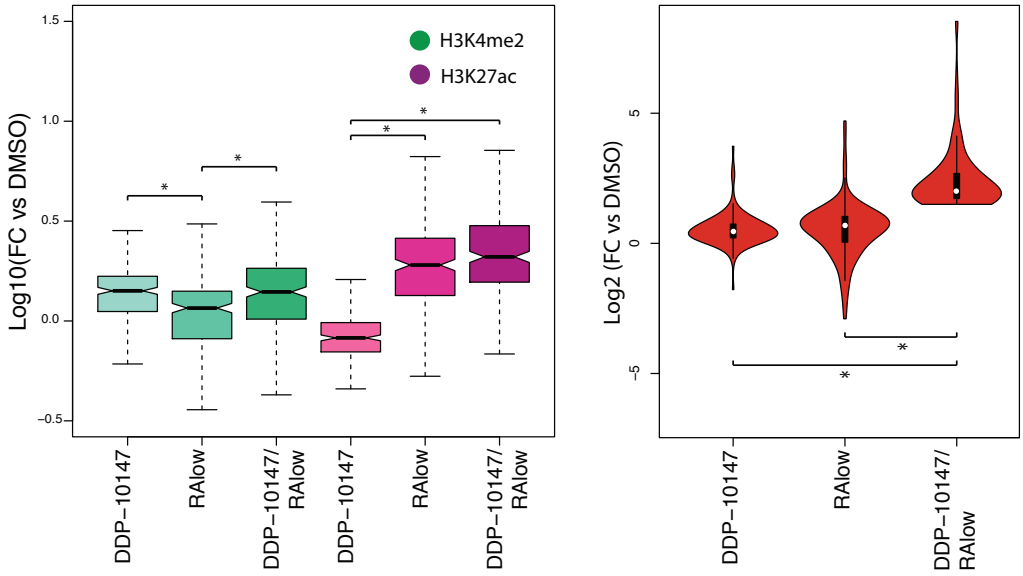



**Figure 8: a.** Boxplot shows level of K4 dimethylation (green) and K27 acetylation (purple) of histone H3 in regions LSD1 positive and associated to genes cotreatment-specific (class 1). Statistics was performed with repeated misures ANOVA comparing log10 FC of RPKM treatment values respect to DMSO followed by Tukey's multiple comparison tests. **b.** Violin plot showing the degree of induction by the indicated treatment vs control (DMSO) in cotreatment-specific subset of genes (class 1, n= 256). Statistics were performed as Repeated Measures ANOVA comparing log2 FC of RPKM treatment values respect DMSO values of each gene included in the class, followed by Tukey's Multiple Comparison Test. LSD1+RAlow was significanly higher respect DDP-10147 and RAlow with * p value <0,0001.

Chromatin coverage in the LSD1 binding site associated to class 1 revealed a statistical significant increase in the H3K4me2 after LSD1 inhibition, comparable

to the cotreatment level, while both the values significantly higher respect to RAlow. Conversely, there is an increase in the level of H3K27ac after RAlow that is comparable to cotreatment values but not present after LSD1 inhibition (Figure 8a). Only in presence of both these chromatin modifications after cotreatment we observed increase in the level of transcription, significantly higher that the two other treatments, in which only one modification at a time increases (Figure 8b). Differently to the genes comprised in the other two classes, regions associated to class 1 are the only ones with this statistical increase in H3K4me2 and H3K27ac mark, suggesting that this is a cotreatment-specific phenomenon.

These data prompted us to propone a model in which LSD1 inhibition poises chromatin for differentiation in NB4 cells by physiological doses of retinoic acid.

### KO RECAPITULATES THE RESULTS OBTAINED WITH LSD1 INHIBITION

To assess whether the effect of the compound was due to its action on LSD1, LSD1 was depleted by CRISPR-Cas9 (Figure 9a). LSD1 depletion does not affect cell growth and colony forming activity (Figure 9 b,c), while sensitivity of NB4 to physiological concentration of RA was clearly enhanced (Figure 9 d,e). We also performed ChIP-seq of H3K4me2 and H3K27ac and RNA-seq of NB4 cells after LSD1 KO. As expected we observed a very small number of genes affected by the LSD1 KO, comparable with the level obtained after treatment with LSD1 inhibitor for 24h (Figure 9 c). Moreover, we observed number of regions with enrichment in H3K4me2 comparable with the level obtained after 24h of LSD1 inhibition (Figure 9d). Taken together these data demonstrate that LSD1 depletion mimics

the effect of LSD1 inhibition, confirming a direct role of LSD1 in RA sensitization of APL cells.

FFf

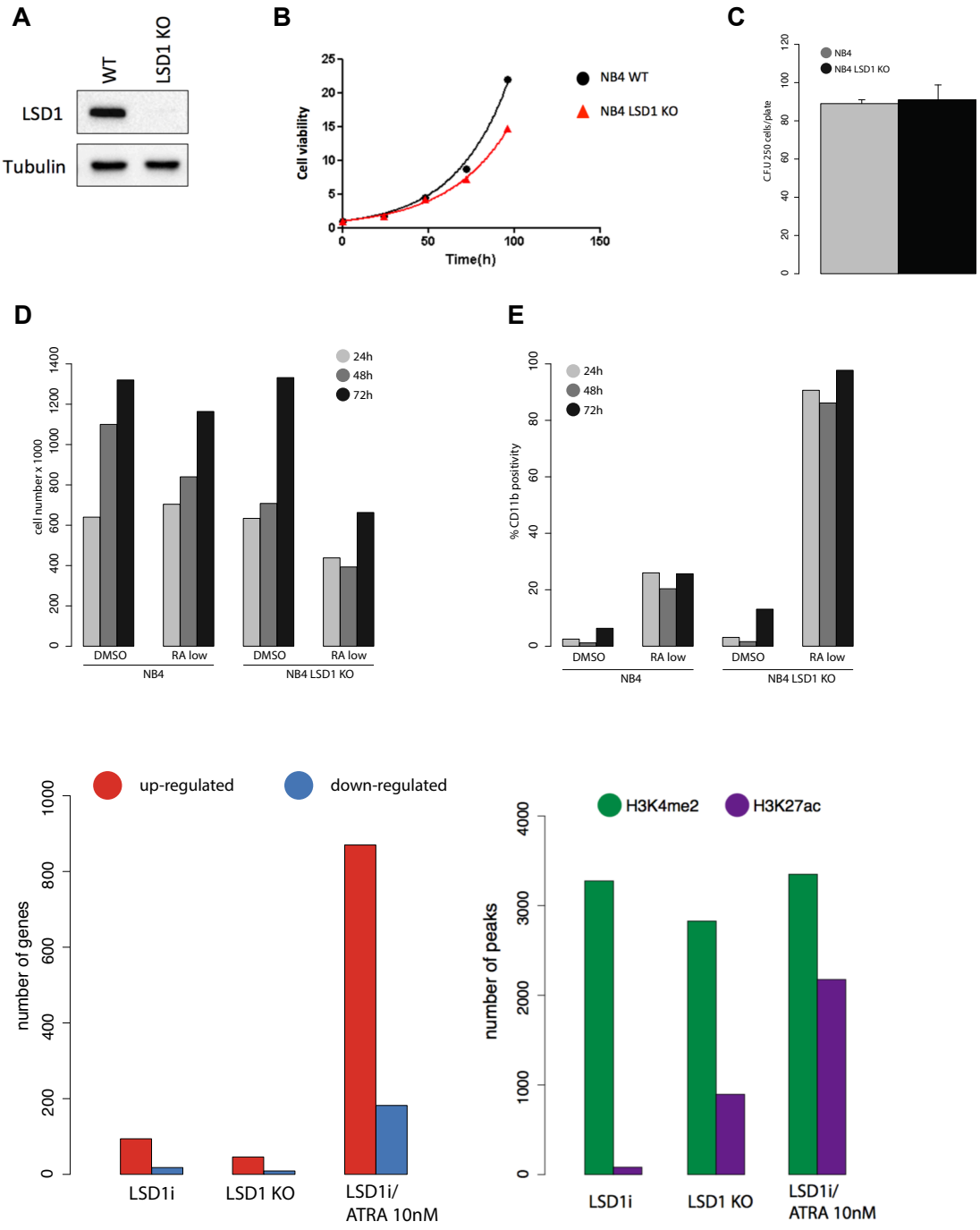



**Figure 9:** RNA and ChIP sequencing was performed in NB4 with LSD1 depleted by CRISPR-Cas9 (genes shared in two different replicates are shown).
**a.** Western blot showing extracts deriving from NB4 wild type and knock out cells. Tubulin has been used as loading control. **b.** Cell viability assay (CellTiterGlo) of NB4 wild type and knock out cells **c.** Colony forming ability, scored after 7 days, of 250 NB4 and NB4 KO cells plated in methylcellulose

medium. Mean and standard deviation of three independent experiments are shown **d.** Proliferation of NB4 wild type and LSD1 knock out cells treated with 10nM retinoic acid and DMSO as control in liquid culture. **e.** Percentage of CD11b positive cells assessed by FACS after 24,48 and 72 hours of treatment in liquid culture. **f.** Barplot represents number of genes regulated (up or down regulated respect to control, RPKM>0.5, absolute (FC)>1.5) in the indicated treatments. **g.** Barplot represents number of regions (peaks) increasing in H3K4me2 or in LSD1 positive regions.

### LSD1 INHIBITION DRIVES DIFFERENTIATION THROUGH A SPECIFIC TRANSCRIPTIONAL PROGRAM, PML-RAR INDEPENDENT

Acute promyelocytic leukemia is characterized in 95% (Sanz and Lo Coco, 2011) of the cases by the translocation t(15;17) that give rise to the oncogenic fusion protein PML-RAR. The expression of the fusion protein was proposed as the initiating event in APL and determines a differentiation block that confers major self-renewal and growth properties to the leukemia clone leading to aberrant proliferation of myeloid progenitors. PML-RAR acts as potent transcriptional repressor that can multimerize and form large protein complexes recruiting several histone-modifiers, such as HDAC1 (Minucci and Pelicci, 2007), which all cooperate to enforce the transcriptional and epigenetic repression. In presence of pharmacological doses of retinoic acid a conformational change occurs with subsequent degradation of the fusion protein, release of corepressor and recruitment of coactivators that allow active transcription of the genes previously repressed by PML-RAR and rapid cell differentiation in vitro and in vivo. (Chen et al., 1991, Chomienne et al, 1990). Contrary to what happens with pharmacological doses of RA (RA high) that triggers PML-RAR to degradation, physiological doses of RA (RA low) do not (Nasr and Guillemin, 2008). For this reason, we wanted to understand the effect of the different treatments on the

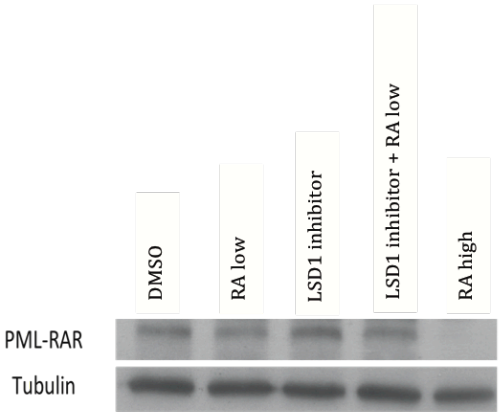
degradation of the protein and, in particular, if the cotreatment determines the degradation of the fusion protein as occurs after treatment with retinoic acid at high concentrations, since they both the treatments drives APL cells differentiation. By western blot (Figure 10a) we observe that the PML-RAR protein remained stable after LSD1 inhibition and RA low, presenting similar levels to DMSO treated cells, while treatment with high doses of retinoic acid leads to a complete degradation of the protein as expected and already demonstrated. Interestingly, the protein level of PML-RAR is not affected by the cotreatment, suggesting that the LSD1 inhibition is able to drive NB4 cells to differentiation and growth arrest in presence of the fusion protein.

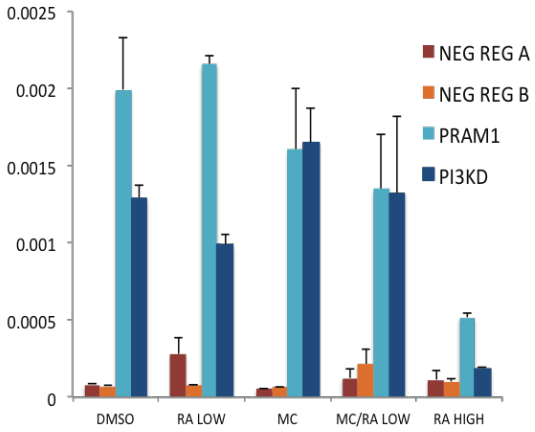
Moreover, since PML-RAR exploits its oncogenic potential ability as an aberrant transcription factor (Saeed et al., 2011), we performed PML ChIP-qPCR on two PML-RAR targets established by ChIP-seq, PRAM1 and PI3KD (Martens et al., 2010), in order to control whether the protein still bind its target genes upon cotreatement.

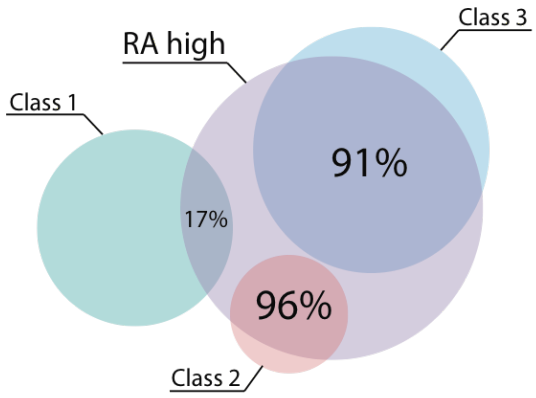
We observed that LSD1 inhibition, RA low and cotreatment did not alter PML-RAR recruitment, while RA high displaced PML-RAR from chromatin (Figure 10 b). These results clearly demonstrated that LSD1 inhibition triggers differentiation and growth arrest of APL cells without affecting PML-RAR stability and recruitment on chromatin.

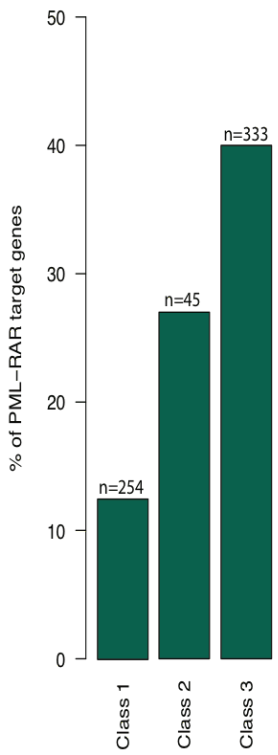
To further confirm the PML-RAR independency of cell differentiation driven by LSD1 inhibition, we compared genes of class 1, with transcriptional induction only after cotreatment (class 1), and genes activated by pharmacological doses of RA (RA high); we observed a very little overlap (17%) between the two sets of genes (Figure 10c), much lower respect the other two other major classes (class 2

and class 3), suggesting that the activation of class 1 genes is probably governed by different mechanisms, PML-RAR independent. In agreement with this observation, it is interesting to note that class 1 genes show a significant lower PML-RAR occupancy (PML-RAR binding site derived from previously NB4 by ChIP-Seq, Martens et al., 2010) compared to genes activated by high concentrations of RA (Figure 10 d,e).


DMSO
RA low
LSD1 inhibitor
LSD1 inhibitor + RA low
RA high
PML-RAR
Tubulin


0.0025
0.002
0.0015
0.001
0.0005
0
NEG REG A
NEG REG B
PRAM1
PI3KD
DMSO
RA LOW
MC
MC/RA LOW
RA HIGH


RA high
Class 3
Class 1
91%
17%
96%
Class 2


% of PML–RAR target genes
50
40
30
20
10
0
n=254
n=45
n=333
Class 1
Class 2
Class 3

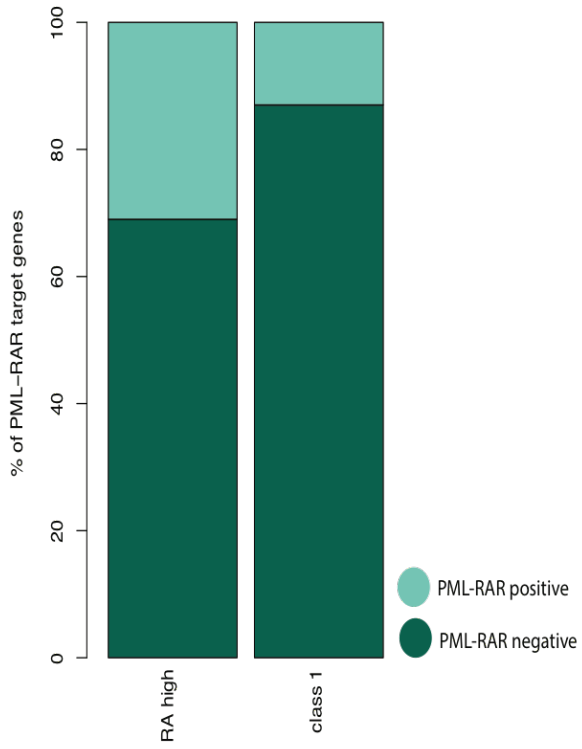

% of PML–RAR target genes
100
80
60
40
20
0
PML-RAR positive
PML-RAR negative
RA high
class 1

**Figure 10: a.** Western blot showing PML-RAR level in NB4 cells treated as indicated for 24h. Tubulin served as loading control. **b.** Histogram represents the ChIP-qPCR for PML in NB4 cells. Enrichment signal was normalized versus input. Two negative regions were used as controls to distinguish aspecific signal and correspond to intergenic (NEG A) or intronic (NEG B) PR-negative region. **c.** Venn diagram represent the overlap of genes contained in the different 3 classes and genes upregulated after RAhigh treatment. **d.** Barplot showing percentage of PML-RAR occupancy in the different 3 classes. **e.** Stacked barplot represent the percentage of genes PML-RAR target in class 1, cotreatment-specific and in genes upregulated after RAhigh treatment, the two percentages are significantly different (Fischer exact text $p<0.0001$).

### THE PRIMING MODEL OF COTREATMENT-SPECIFIC GENES

All the results obtained by the analysis and the comparison of transcriptional and chromatin profiles among the different treatments allowed us hypothesize a mechanistic model of action cotreatment.

In regions associated to genes belonging to class 1, and then upregulated only after cotreatment, the two treatments alone lead to the increase of two different histone modifications H3K4me2 and H3K27ac; the same two treatments combined in a single combination drive to activation of transcription and cell differentiation. The effects of the two treatments alone, LSD1 inhibitor and retinoic acid at physiological concentrations, although leading to an increase in histone markers are not sufficient to activate the genes responsible for cell differentiation. The synergistic action of the two treatments leads to a transcriptional activation and cell differentiation. In addition to this model, that takes into account the histone markers and the presence of LSD1, further consideration must be made regarding the presence of the fusion protein PML-RAR. As already demonstrated, treatment with low doses of retinoic acid is not

sufficient for the degradation of the fusion protein, which is degraded instead with pharmacological doses of retinoic acid, leading to the removal of transcriptional blocking and activation of genetic pathways responsible for cell differentiation.

Treatment with physiological doses of retinoic acid and the inhibitor of LSD1 determines cell differentiation while maintaining the fusion protein, which is not degraded but also retains the binding with DNA; the contrary happens with pharmacological doses of retinoic acid, that determines the detachment and degradation of the protein. Cell differentiation as a result of the cotreatment is caused by the activation of a different pathway genes compared to those activated by retinoic high, and this probably derives from the need to bypass the transcriptional block caused by the fusion protein. This is also consistent with the low number of PML-RAR target genes among the genes activated by cotreatment.

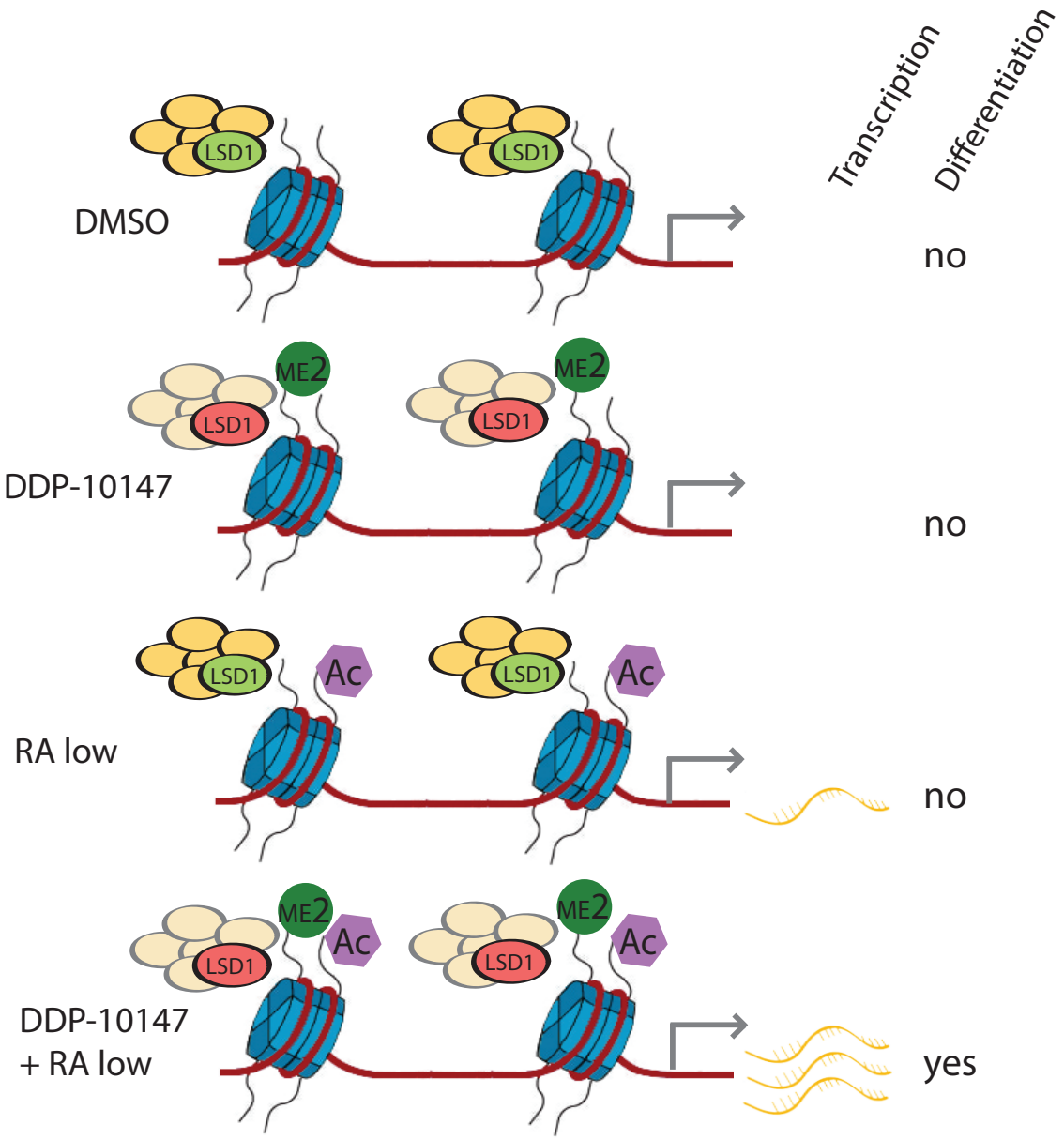



**Figure 11**: Model summarizing chromatin dynamics in class 1 genes associated regions upon the indicated treatments. At specific regions, bounded by LSD1, treatment with LSD1 inhibitor leads to H3K4me2 increase while treatment with low doses of retinoic acid determined increase in H3K27ac. In both the treatment these increase are not sufficient to activate the associated genes and cells remain undifferentiated. Combination of the two treatments caused an increase in both H3K4me2 and H3K27ac that drive the activation of the associated genes and cell differentiation.

### LSD1 BINDING IN PRIMING MODEL-SELECTED REGIONS ARE CONSERVED AFTER TREATMENTS

The priming model proposed considers as positive LSD1 genes those genes that were identified as protein binding site according to the ChIP-seq made on untreated NB4 cells (DMSO). The next question was whether this model could be considered valid even taking into account a possible repositioning of the protein as a result of the treatments. Recent studies demonstrate that inhibition of LSD1 in Kasumi-1 and SKNO-1 cells evict LSD1 from chromatin (McGrath et al., 2016). For this reason we performed LSD1 ChIP-seq in NB4 cells after 24h of treatment with inhibitor, cotreatment and retinoic acid at high concentration. Similarly to the previous LSD1 ChIP-seq we validate the peaks obtained with the peakscalling for each treatments by q-PCR and set the threshold to 234, 357,239,196 MACS score respectively for DMSO, LSD1 inhibitor, cotreatment and RAhigh. The analysis on the peaks, called on the NB4 input and filtered according to the thresholds obtained by q-PCR, revealed that almost all the binding regions present in the DMSO are maintained after the three treatment analyzed.

This result allowed us to confirm that the genes and their respective regions of the proposed priming model were retained as protein binding sites even after the treatments. Thus the model was valid also considering the possible repositioning of protein.

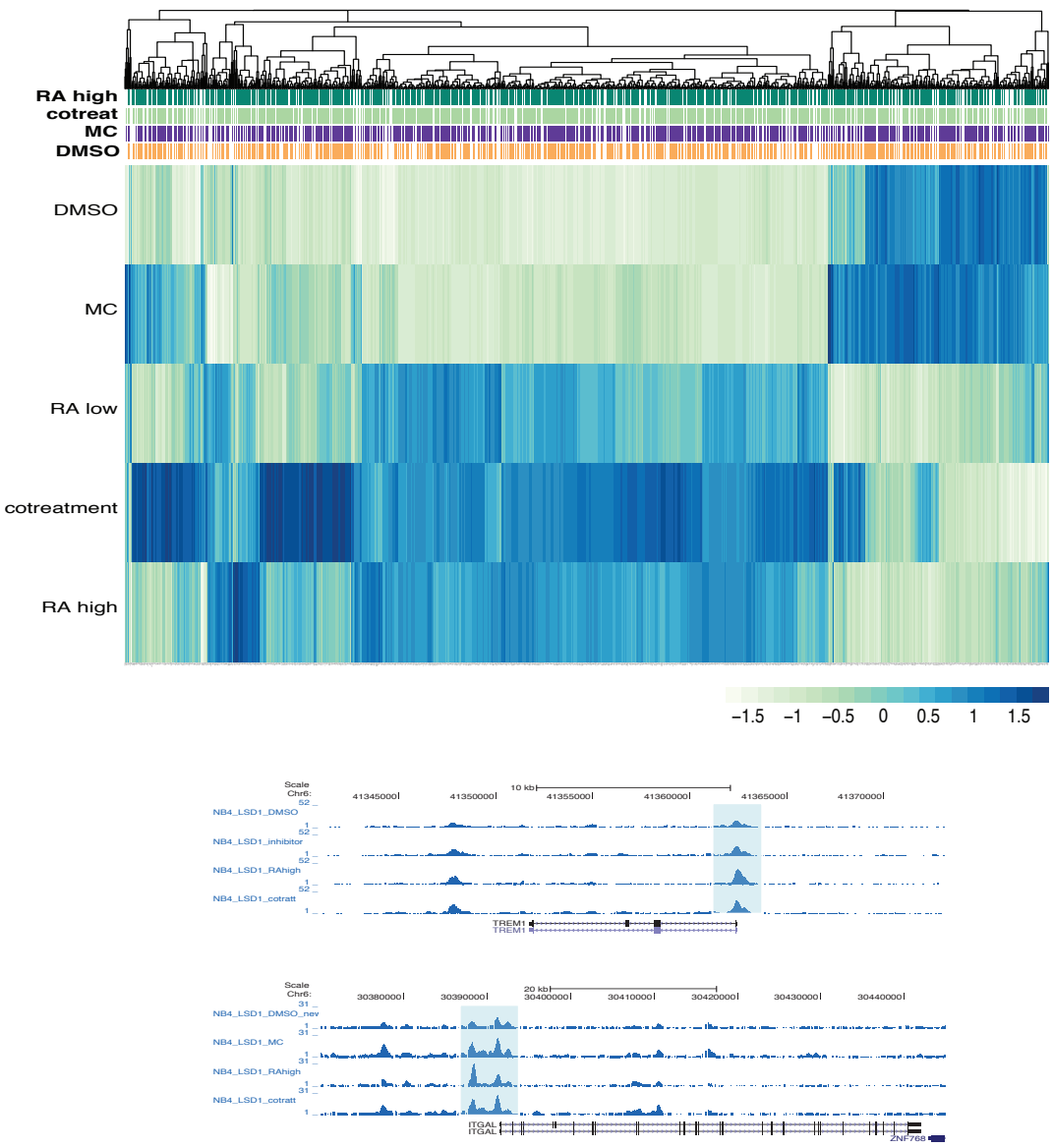



**Figure 12: a.** Heatmap shows expression level of genes regulated in at least one of the four treatments against the DMSO (same gene as figure 4a, n=1235) The expression levels range from light blue (less expressed) to dark blue (more expressed). The colored flags on the left represent the LSD1 binding in the indicated treatment. **b.** UCSC genome tracks of two class 1 genes show the maintenance of LSD1 peaks among the different treatments.

Moreover, to explore the correlation between LSD1 and global modulation of gene expression we extended the previous analysis by adding to the heatmap of modulated genes in figure 4a also the binding sites of the protein after the different treatments (Figure 12a). As expected there is no clustering of protein binding compared to the various subsets of genes expressed, the binding does not appear to change significantly among treatments and affects almost all genes

considered (genes significantly expressed in at least one of the treatments compared to DMSO, n=1235).

# DISCUSSION

### LSD1 INHIBITION SENSITIZE NB4 CELLS TO RA-INDUCED DIFFERENTIATION

Our lab takes part in the development of a new compound working as a LSD1 specific inhibitor (Binda et al., 2010). Taking advantage of the high specificity of the inhibitor we aimed to characterize the role of LSD1 in APL cells. We found, by morphological characterization and analysis of differentiation-associated marker, that both LSD1 KO and pharmacological inhibition of the protein sensitizes the NB4 cells, a PML-RAR expressing APL cell line, to retinoic acid-induced differentiation. Results in cell proliferation and colony forming ability revealed an even stronger effect of the combination of LSD1 inhibitor and RA acid at concentration 100-fold lower respect to pharmacological doses of RA, capable of induce differentiation. We demonstrate that a pulse of inhibition was sufficient to prime APL cells to differentiate, suggesting that LSD1 dependent activities play a role in the initial phases of RA induced differentiation.

### LSD1 GENOMIC DISTRIBUTION

To investigate the function of LSD1 in NB4 cells, we first identified the binding sites of the protein by using chromatin immunoprecipitation coupled with massively parallel DNA sequencing (ChIP-seq). We found that LSD1 occupancy in APL cells is mainly promoter-associated and reads density are higher around the TSS of the annotated genes. Nevertheless we also found a significant proportion of LSD1 peaks in promoter distal regions, in accordance with previous results

obtained in mESCs and immature murine granulocytic cell line (Whyte et al., 2012, Kereniy et al., 2013). Motif discovery analysis revealed that promoters bounded by LSD1 were enriched for binding matrix of SPI1, EGR1 and E2F1-3, transcription factors involved in the granulocytic/monocytic differentiation. Moreover we also found REST, a known LSD1 recruiter (Mosammaparast and Shi 2010) and the canonical PML-RAR binding sequences (RAR:RXR DR5), suggesting that LSD1 can interact also with the oncogenic fusion protein characteristic of APL, at least in the promoters of certain genes.

### CO-TREATMENT ACTIVATES THE TRANSCRIPTIONAL DIFFERENTIATION PROGRAM OF NB4 CELLS AND DRIVE CELL TO DIFFERENTIATION

Despite its well-supported activity as transcriptional repressor, LSD1 binds are all across the genome also at actively transcribed promoters and enhancer (Whyte et al, 2012). We found indeed a large overlap between LSD1 and PolII occupancy, suggesting that LSD1 may cooperating to keep transcription of target genes under a certain threshold rather than repressing it completely or that its activity is dynamically counterbalanced by co-located transcription activator. Our RNA-seq on NB4 cells treated for 24 with inhibitor, retinoic acid or combination of the two revealed that the binding of the protein doesn't affect the level of transcription in none treatment both in up-regulated and in down-regulated genes. We also observed that inhibition of LSD1 by itself does not have a big impact on the transcription but in accordance with its repressor role the few genes modulated respect to the DMSO are up-regulated while only one gene is down-regulated. Retinoic acid at physiological concentration has only a mild

effect while we observe a significant increase in transcription after co-treatment and retinoic acid at pharmacological concentrations, consistent with the phenomenon of cell differentiation only visible as result of these two treatments.

Interestingly the cell differentiation observed after co-treatment is activated by two different pathways of genes, one subset of genes shared with retinoic acid at high concentration and the other subset induced only after co-treatment but always related, as for the first subset, with hematological system development and function, tissue morphology and immune cell trafficking.

These results suggest that the activity of LSD1 per se is not directly involved in transcription but has a poising function that contributes to the activation by retinoic acid low of an alternative pathway of genes specific for the co-treatment and different respect retinoic acid at higher concentration.

### LSD1 INHIBITOR-DEPENDENT EPIGENETIC MODULATION PRIMED NB4 DIFFERENTIATION

The poising hypothesis suggested by transcription data is corroborated by the analysis of chromatin alteration occurring after the different treatments. ChIP-seq analysis revealed a global significant increase in H3K4me2 after treatment with LSD1 inhibitor alone or in combination, co-localized with LSD1 peaks within promoters and TSS distal regions.

Interesting we can not observe this increase in NB4 cells treated with retinoic acid at low concentration, where we can instead observe increase in the H3K27ac level, comparable to the one observed in co-treatment. It is important to note that this increase of H3K27ac, generally associated to an active transcription, in this

case is not correlated with the expression levels of the genes, which remain comparable to the control.

Since only the combination of the two treatments induces cell differentiation, we reasoned that the presence of both H3K4me2 and H3K27ac is a prerequisite for the activation of the differentiation program in NB4 cells. Moreover this effect is observed only in regions bounded by LSD1 and associated with a subset of cotreatment-specific genes and not with RAhigh induced genes, indicating that the H3K4me2 increase after LSD1 inhibitor poises the chromatin for the activation by retinoic acid of an alternative pathway of genes responsible for differentiation.

### CHARACTERIZATION OF LSD1 - PML-RAR INTERPLAY

APL patients present in the 95% of the cases the expression of the fusion protein PML-RAR (PR) due to the translocation t(15;17). The oncogenic potential of the fusion protein is expressed by the influence on both PML pathway and retinoic acid receptor functions. Treatment with retinoic acid reactivates the downstream RARa pathway-inducing differentiation of APL blasts through degradation of the fusion protein, a step that has been proposed as crucial for differentiation and eradication of APL. At molecular level only high (pharmacological) concentration of retinoic acid triggers PML-RAR to proteosomal degradation while physiological doses does not (Nasr et al., 2008). For this reason, we wanted to understand if LSD1 inhibition combined with low doses of retinoic drives APL cells differentiation through the degradation of PML-RAR, as RA high does. Interestingly, PML-RAR protein remained stable after LSD1 inhibition, presenting

similar levels to DMSO treated cells, while it was only very little affected by RA low, and complete degradated by RA high. Moreover since PML-RAR exploits its oncogenic potential ability as aberrant transcription factor (Saed et al., 2011) we investigated PML-RAR binding of its target genes after the different treatments.
PML ChIP-qPCR on PRAM1 and PI3KD, two PML-RAR targets established by ChIP-Seq (Martens et al., 2010), revealed that LSD1 inhibition and RA low did not alter PML-RAR recruitment, while RA high displaced PML-RAR from chromatin. Further confirmation of the independency of cell differentiation driven by the cotreatment from the PML-RAR fusion protein is the little overlap of PML-RAR target genes and cotreatment-specific genes observed. The almost total absence of genes target of the fusion protein and the little overlap between cotreatment genes and genes induced by RA high, suggest that the action of LSD1 leads to an activation governed by different mechanism respect to RA high. Contrary to RA high the cotreatment doesn't altered the levels and the binding of the protein, thus LSD1 inhibition acts on non-target PML-RAR genes and allows the gene activation through an alternate mechanism which bypasses the block determined by the fusion protein.

Collectively, our experiments characterized the role of LSD1 in APL, suggesting a mechanistic interpretation of its action. LSD1 is bound across almost all the genes regulated in this system upon differentiating condition but its binding doesn't influence directly the transcription. Its actions seem to be selective mediated by its H3K4me2 demethylase activity in specific regions related to genes overexpressed only after cotreatment. LSD1 acts on these non-PML-RAR target

genes overcoming the block of transcription due to the fusion protein and poises chromatin for a later phases induction of differentiation by RA low.

# APPENDIX

We extended in the last few months the analysis of the chromatin markers also to treatment with retinoic acid at high concentration (RA high), which shows similar phenotype to co-treatment.

The same analysis performed on the class 1 region was also performed on the ChIP-seq of NB4 cells after RA high. Although the expression of those genes does not increase after RA high treatment, the level of H3K4 di-methylation increase and is comparable to the inhibitor and the co-treatment levels, suggesting that this change in H3K4me2 can occur even without pharmacological inhibition of LSD1 (Figure 13).

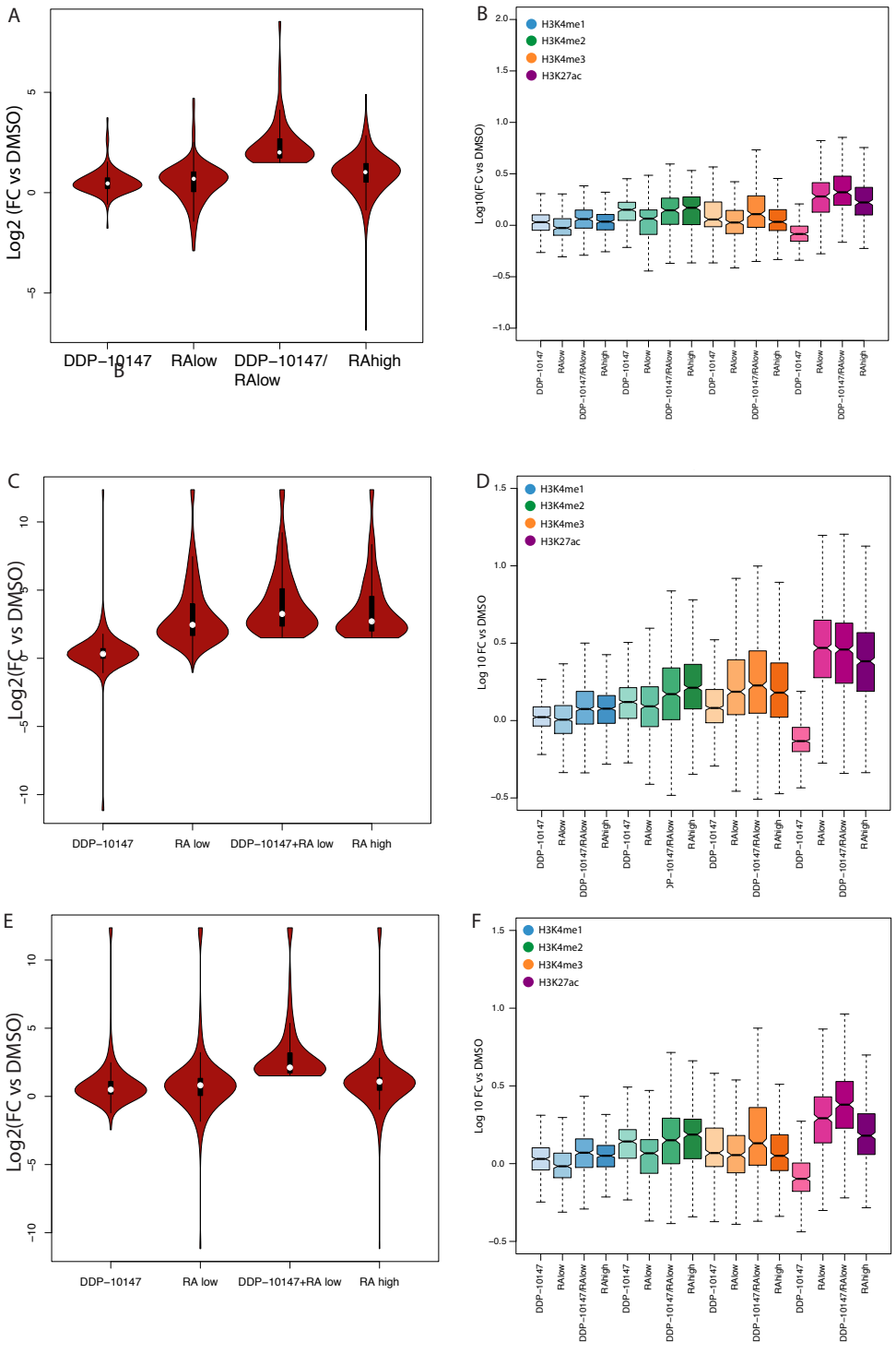



**Figure 13.** Violin plot showing the degree of induction by the indicated treatment respect to control (DMSO) in **a.** class 1(n= 256), **c.** genes induced by RA high and co-treatment (n=486) **e.** cotreatment-specific genes (n=382). Boxplot shows level of H3K4me1 (blue), H3K4me2 (green), H3K4me3 (orange) and H3K27ac (purple) in regions LSD1 positive and associated to genes in **b.** class 1(n= 256), **d.** common genes between RA high and co-treatment (n=486) **f.** cotreatment-specific genes (n=382)

We determined the chromatin alteration occurring in the subset of genes specifically induced by the two treatments that trigger the cells to differentiation: co-treatment specific genes (n=382) and genes induced by both RA high and co-treatment (n=486). Not surprisingly active transcription is positively correlated with the presence of both H3K4me3 and H3K27ac; LSD1 inhibition as expected induces an increase in the levels of H3K4me2 both alone and to a greater extent in combination with RA low. Remarkably the H3K4me2 marker increases not only after LSD1 inhibition but also after high doses of retinoic acid in genes regulated by RA high as well as those not induced by RA high treatment.

The results of these new analyzes show hat the increase in H3K4me2 is not directly related to the activity of the inhibitor neither to transcriptional activation, and therefore our initial hypothesis cannot be entirely correct. While the lack of association with LSD1 inhibition by drugs could be explained by models where the enzymatic activity of LSD1 can be reverted by recruitment of other factors (with an unknown effect on chromatin), it is more difficult to reconcile the lack of transcriptional activation in the condition “RA high” with a positive role for H3K4Me2 in gene activation, though we cannot exclude that other unknown histone marks/other mechanisms may “block” the positive effect and only upon LSD1 inhibition we can achieve full transcriptional activity.

One alternative and somewhat unexpected explanation for what we observed is however that the enhancement of H3K4Me2 (even upon LSD1 inhibition) is not strictly necessary for gene activation. To assess if LSD1 enzymatic activity

is required for the sensitization that we observe in APL cells to RA low, we expressed either wild type or catalytic inactive LSD1 (K661A) in LSD1 knock out NB4 cells (figure 14a) . While KO cells are very sensitive to low doses of retinoic acid in terms of cell proliferation, LSD1 expressing cells as well as K661A expressing cells are insensitive to RA low, recapitulating the phenotype of NB4 wild-type cells and clearly demonstrate that the catalytic activity of LSD1 is dispensable for the sensitization of APL cells to low doses of retinoic acid (figure 14b).

In addiction by ChIP-qPCR we demonstrate that both re-expressed wild-type LSD1 and mutant K661A are able to re-localize in the chromatin regions bounded by LSD1 and previously identified. And also that only wild-type LSD1, but not the catalytically inactive K661A, is able to reduce the level of H3K4me2 in that target regions (Figure 14c).

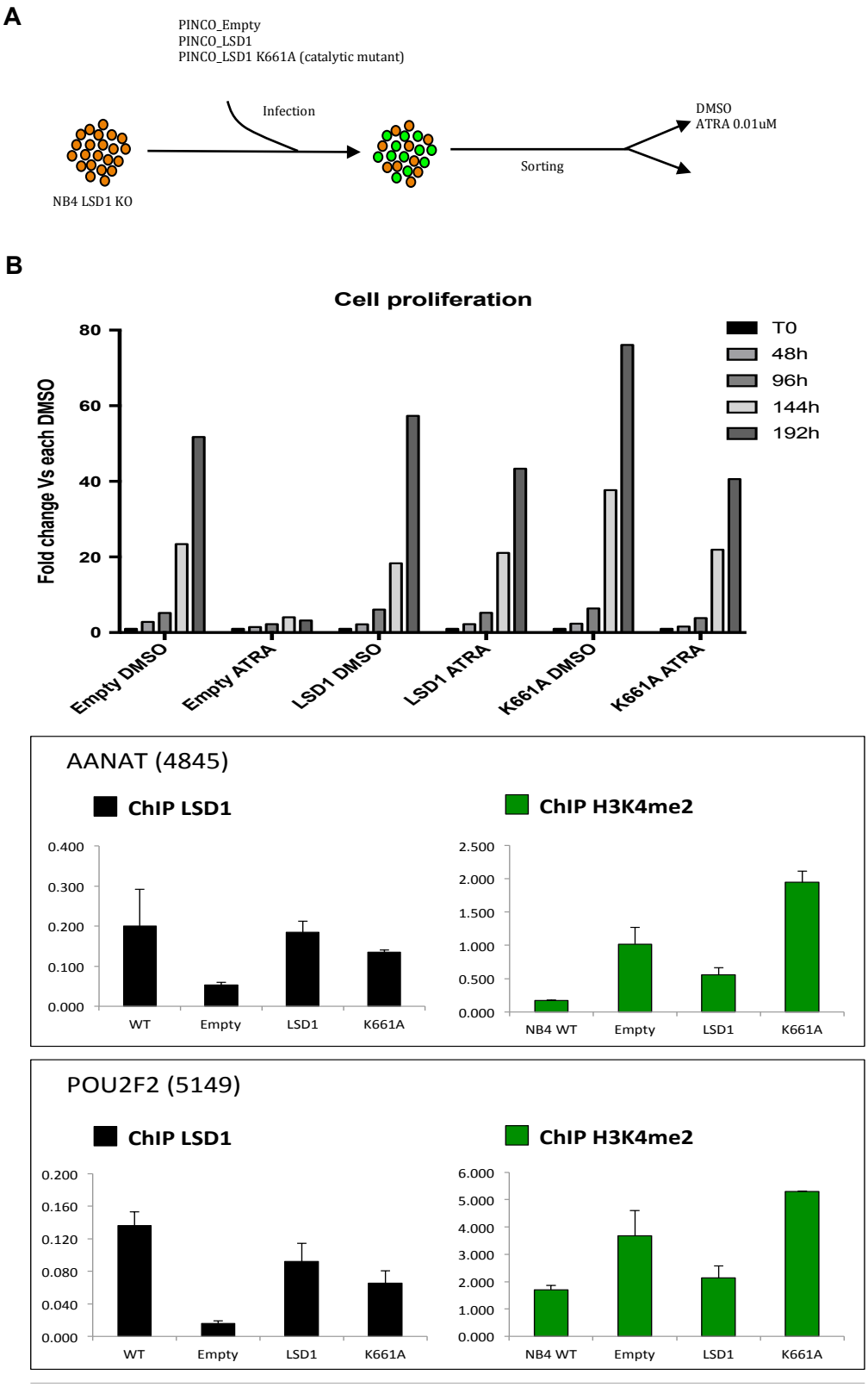



**Figure 14. LSD1 catalytic activity is dispensable for the sensitization of NB4 cells to RA. a.** Design of the experiment. NB4 LSD1 knock out cells are infected with LSD1 wild type or catalytic mutant (K661A), sorted and treated with RA. **b.** Proliferation of cells treated as indicated. **c.** ChIP-qPCR of LSD1 (left) and H3K4me2 at selected regions in NB4 wild type and knock out cells infected with empty vector (control), LSD1 wild type and K661A.

These unexpected observations opens a plenty of questions on the mechanism of action of LSD1 and on his altered function after inhibition.

The well known role of LSD1 as a member of many protein complexes containing histone deacetylases and methyltransferases such as NURD and COREST suggests that LSD1 should exerts its actions, at least those related to the observed phenotype in APL, more as a member and scaffold of these complexes than as a single enzyme.

This hypothesis is supported by the increase in the binding of the protein that we observe after differentiation-related treatment (cotreatment and RAhigh), which is to a lesser extent observable also after inhibition of the protein alone (figure 15).

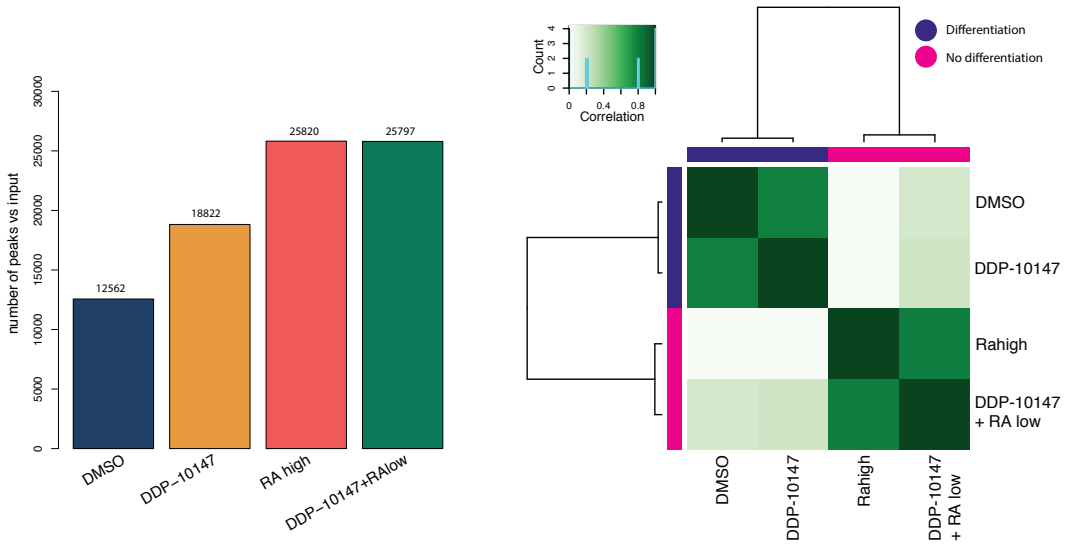



**Figure 15. LSD1 chromatin occupancy increase after treatment with inhibitor, RA high and cotreatment. a.** Number of enriched regions (peaks) respect to input present at indicated treatment. **b.** Correlation heatmap using occupancy data.

# APPENDIX DISCUSSION AND PERSPECTIVES

In conclusion these data support the therapeutic implications of LSD1 in AMLs, APL in particular, as LSD1 fulfill a pivotal role in normal and malignant hematopoiesis. We demonstrate that the enzymatic acitivity of LSD1 is dispensable for the retinoic acid sensitization of APL. This unexpected result suggests a novel role of LSD1 as a key interactor and scaffold of multiprotein complexes. Since it is well known that LSD1 is a member of many protein complexes containing histone deacetylases and methyltransferases and it's been described the interaction of LSD1 with many different partner (e.g PRC2, MLL) further investigation should be performed for the identification of LSD1 interactors in these APL model, and the characterization of their function related to the observed phenotype.

# LIST OF ABBREVIATIONS

| | |
|---|---|
| aa | amino acid |
| ac | (es. H3K27ac) acetylated residues |
| AML | Acute Myeloid Leukemia |
| APL | Acute Promyelocytic Leukemia |
| ATP | Adenosine Triphosphate |
| ATRA/RA | All trans retinoic acid |
| bp | base pairs |
| ChIP | Chromatin Immuno Precipitation |
| DDP-10147 | LSD1 inhibitor |
| DNMT | DNA methyltransferase |
| HAT | Histone/lysine acetyltransferases |
| HDAC | Histone deactylases |
| HSC | Hematopoietic stem cell |
| Kb | kilobase |
| LSD1 | lysine specific demethylases |
| MC | LSD1 inhibitor DDP-10147 |
| me | (es. H3K4me3) methylated residues |
| mRNA | messanger RNA |
| PCR | Polymerase chain reaction |
| PML | Promyelocitic leukemia protein |
| PML-RAR | Acute promyelocitic leukemia fusion protein |
| PMTs | post translational modifications |
| PR | PML-RAR protein |
| RA low | Physiological concentration of RA (0.01 μM) |
| RA high | Pharmacological concentration of RA (1 μM) |
| RARα | Retinoic acid receptor |

# MATERIALS AND METHODS

Pierluigi Rossi and Roberto Ravasio performed the wet-lab experiments of the project.

**Cell culture**

NB4 cells, isolated from an APL patient by Lanotte and colleagues (Lanotte et al., 1999), have characteristics similar to APL blasts and were grown in RPMI medium plus 10% of fetal calf serum, 2mM glutammine and 1% Penicillin/Streptomycin.

**Treatments**

NB4 cells, plated at 100.000/ml, were treated with: dimethyl sulfoxide (DMSO) 1/1000; inhibitor ,2 μM, for 6,12,24 or 96h; RAlow, 0.01 μM, for 24 or 96h; RAhigh 1 μM for 24 or 96h.

**NB4 LSD1 KO**

Single-guide sequence specific to LSD1 (exon 1) was designed using the CRISPR design tool (http://tools.genome-engineering.org, Ran et al., 2013) and cloned into lentiCRISPR (Sanjane et al.,2014). The sequence selected (based on the lowest number of predicted off-target in exon and the highest predicted efficiency) was the following one: 5'-CACCGCGCGGAGGCTCTTTCTTGCG-3'. After infection cells were selected with puromycine for 3 days and then seeded in 96-

well plated by dilution and expanded. Clones were screening using Suveyor Assay (Ran et al., 2013) and western blot analysis. Positive clones were subjected to Sanger Sequencing.

Figure 9a: NB4 cells were plated at 3500 cells per well in 96 well tissue culture dishes. Relative cells numbers were assessed by Cell Titer-Glo luminescent cell viability assay (Promega, Madison, WI USA). GraphPad 6 (GraphPad Software, Inc., La Jolla, CA USA) was used for curve fitting.

Figure 9b: Colony forming ability, scored after 7 days, of 250 NB4 and NB4 KO cells plated in methylcellulose medium. Mean and standard deviation of three independent experiments are show.

**Gene expression analysis**

mRNA-seq was performed according to the True-seq Low sample protocol selecting only polyadenylated transcript. RNA-seq analysis was performed with the TopHat and Cufflinks algorithm (Trapnell et al., 2010). We adopt a 36 bp paired end sequencing strategy. The number of reads obtained was comparable among the samples. The values considered to quantify the relative expression of a given gene correspond to the number of reads aligned per kilobases of the transcript per million mappable fragment detected (FPKM, fragment per kilobase of exon per million fragments mapped). These values were used for all the comparative analysis.

The threshold set to consider a gene as being regulated was FDR≤ 0.05, FPKM ≥ 0.5 and Fold change respect DMSO greater than absolute Log2(1.5) .

Class in table 1 are mutually exclusive, genes were considered up-regulated (one arrow) if they fulfill the threshold set above, while the further-upregulated genes

(two arrow) were genes up-regulated in both the treatments but with a log2FC greater than 1.5 between them.

Gene ontology was performed through the use of QIAGEN's Ingenuity Pathway Analysis (IPA, QIAGEN Redwood City, www.qiagen.com/ingenuity)

**ChIP-seq analysis**

Raw data corresponding to reads coming from Illumina Genome Analyzer II were analyzed according to the Fish the ChIPs pipeline (Barozzi et al., 2011). Reads were mapped to Human NCBI36/hg18. Only sequences showing unique alignment were used for peak detection, allowing for a maximum of two mismatches.

Peaks calling was performed with MACS with the threshold of -log10(p)=5 for all the ChIP-seq except LSD1, for which we perform qPCR-ChIP and set a more stringent threshold. For the validation of LSD1 positive regions we selected regions with several p-values (obtained with MACS peak calling) and perform ChIP-qPCR assay in three independent experiments. Using intergenic region negative for LSD1 as control and anti-IgG antibody as mock control we set the minimum value for positive true enrichment as the enrichment of negative control plus three times the relative SD. 26 out of 30 regions tested were considered as validated and thus we set as MACS score threshold the minimum score among these 26 validated regions. With this new threshold of -log10(p)=16.9 we obtained 15187 LSD1 peaks. Reads from each sample were normalized to the input of NB4 cell line. Peaks were associated to Refseq annotated genes according to GIN (Cesaroni et al., 2008) while intergenic regions were considered as having more than 22kb of distance from the nearest gene.

For LSD1 ChIP-seq after treatment we used the same validation strategy and set the thresholds to −10*log10(pvalue)= 357,239,196 MACS score respectively for LSD1 inhibitor, cotreatment and RAhigh.

UCSC Genome tracks were generated normalizing each sample to the same sequencing depth.

The intersection among the peaks datasets was performed with bedtools intersect tool, peaks are considered overlapping if they share at least 1bp.

Reads coverage for the comparison of chromatin changes in the first three classes of gene expression (Table 1) was performed with bedtools suite inside the window delimitated by the LSD1 peaks associated to each gene; differences in methylation, acetylation or gene expression values, as FPKM, among the different treatments were estimated by means of Repeated Measures ANOVA followed by Tukey's Multiple Comparison Test.

All plots were performed with R (R Core Team, 2016).

Vioplot: Daniel Adler (2005). vioplot: Violin plot. R package version 0.2. http://wsopuppenkiste.wiso.uni-goettingen.de/~dadler

Pheatmap: Raivo Kolde (2015). pheatmap: Pretty Heatmaps. R package version 1.0.8. https://CRAN.R-project.org/package=pheatmap

DiffBind: Stark R and Brown G (2011). *DiffBind: differential binding analysis of ChIP-Seq peak data*.

# REFERENCES


Ablain J & de The H (2011) “Revisiting the differentiation paradigm in acute promyelocytic leukemia”. *Blood* 117, 5795–5802.

Ablain J, Leiva M, Peres L, Fonsart J, Anthony E, de Thé H. (2013). "Uncoupling RARA transcriptional activation and degradation clarifies the bases for APL response to therapies." *J Exp Med* 210(4): 647-53.

Adamo A, Sesé B, Boue S, Castano J, Paramonov I, Barrero MJ, Izpisua Belmonte JC. (2011). “LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells”. *Nat. Cell Biol*. 13(6), 652–659

Akalin A, Garrett-Bakelman FE, Kormaksson M, Busuttil J, Zhang L, Khrebtukova I, Milne TA, Huang Y, Biswas D, Hess JL Allis CD, Roeder RG, Valk PJM, Lowenberg B, Delwel R, Fernandez HF, Paietta E, Tallman MS, Schroth GP, Mason CE, Menick A, Figueroa ME. (2012) “Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia”. *PLoS Genet* 8, e1002781.

Allfrey VG, Faulkner R, Mirsky AE. (1964) “Acetylation and methy- lation of histones and their possible role in the regulation of RNA synthesis”. *Proc Natl Acad Sci USA,* 51:786-794.

Amente S, Lania L, Majello B. (2013). "The histone LSD1 demethylase in stemness and cancer transcription programs." *Biochim Biophys Acta*, 1829(10): 981-6.

Bannister AJ, Kouzarides T, (2011) “Regulation of chromatin by histone modifications”. *Cell Res*, 2001, 21: 381-395.

Barlesi F, Giaccone G, Gallegos-Ruiz MI, Loundou A, Span SW, Lefesvre P, Kruyt FA & Rodriguez JA. (2007). “Global histone modifications predict prognosis of resected non small-cell lung cancer”. J *Clin Oncol* 25, 4358–4364.

Barozzi I, Termanini A, Minucci S, Natoli G. (2011) “Fish the ChIPs: a pipeline for automated genomic annotation of ChIP-Seq data”. *Biol Direct*, 6:51

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

Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. (2010). “Nucleosome-interacting proteins regulated by DNA and histone methylation”. *Cell,* 143(3):470-484.

Baylin SB & Jones PA (2011) A decade of exploring the cancer epigenome -

biological and translational implications. *Nat Rev Cancer* 11, 726–734.

Binda C, Valente S, Romanenghi M, Pilotto S, Cirilli R, Karytinos A, Ciossani G, Botrugno OA, Forneris F, Tardugno M, Edmondson DE, Minucci S, Mattevi A, Mai A.(2010). "Biochemical, structural, and biological evaluation of tranylcypromine derivatives as inhibitors of histone demethylases LSD1 and LSD2." *J Am Chem Soc* 132(19): 6827-33.

Bird, A. "Molecular biology: methylation talk between histones and DNA". Science, 2001, 294:2113-2115.

Bird A. "DNA methylation patterns and epigenetic memory." *Genes Dev*, 2002, 16(1): 6121.

Cesaroni M, Cittaro D, Brozzi A, Pelicci PG, Luzi L (2008). "CARPET: a web-based package for the analysis of ChIP-chip and expression tiling data". *Bioinformatics* 24: 2918–2920

Chomienne C, Ballerini P, Balitrand N, Daniel MT, Fenaux P, Castaigne S, Degos L. (1990) "All-trans retinoic acid in acute promyelocytic leukemias. II. In vitro studies: structure-function relationship". *Blood*, 76(9):1710–1717.

Copeland RA, Solomon ME, Richon VM. (2009) "Protein methyltransferases as a target class for drug discovery". *Nat Rev Drug Discov*, 8:724-32

De The, H., C. Chomienne, Lanotte M, Degos L, Dejan A. (1990). "The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor alpha gene to a novel transcribed locus." *Nature*, 347(6293): 558-61.

De Thé H & Chen Z, (2010) "Acute promyelocytic leukaemia: novel insights into the mechanisms of cure" *Nature Reviews Cancer,* 10:775-783.

Di Croce L, Raker VA, Corsaro M, Fazi F, Fanelli M, Faretta M, Fuks F, Lo Coco F, Kouzarider T, Nervi C, Minucci S, Pelicci PG. (2002) "Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor". *Science* 295, 1079–1082

Elsheikh SE, Green AR, Rakha EA, Powe DG, Ahmed RA, Collins HM, Soria D, Garibaldi JM, Paish CE, Ammar AA, Grainge MJ, Ball GR, Abdelghany MK, Matrinez-Pomares L, Heery DM, Ellis IO. (2009) "Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome". *Cancer Res* 69(9):3802–9.

Emrich, SJ, Barbazuk WB, Li L, Schnable PS. (2007). "Gene discovery and annotation using LCM-454 transcriptome sequencing". *Genome Res*, 17(1), 69-73.

Ernst, J., P. Kheradpour, et al. (2011). "Mapping and analysis of chromatin state dynamics in nine human cell types." Nature 473(7345): 43149.

Esteller, M. (2007). "Cancer epigenomics: DNA methylomes and histone-modification maps." *Nat Rev Genet* 8(4): 286-298.

Esteller M and Herman JG (2002). "Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours." *J Pathol* 196(1): 117.

Falkenberg KJ and Johnstone RW (2014) " Histone deacetylases and their inhibitors in cancer, neurological disease and immune disorders" *Nature Reviews Drug Discovery* 13, 673-691

Fang, R, Barbera AJ, Xu Y, Ruthenberg M, Leonor T, Bi Q, Lan F, Mei P, Yuan GC, Lian C, Peng J, Cheng D, Sui G, Kaiser UB, Shi Y, Shi YG. (2010). "Human LSD2/KDM1b/AOF1 regulates gene transcription by modulating intragenic H3K4me2 methylation." *Mol Cell* 39(2): 222-33.

Figueroa ME, Lugthart S, Li Y, Erpelinck- Verschueren C, Deng X, Christos PJ, Schifano E, Booth J, van Putten W, Skrabanek L, Campagne F, Mazumdar M, Greally JM, Valk PJ, Lowenberg B, Delwel R, Melnick A. (2010) "DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia". *Cancer Cell* 17(1):13–27.

Fischle W, Tseng BS, Dormann HL,Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H, Allis CD. (2005) "Regulation of HP1-chromatin binding by histone H3 methylation and phos- phorylation". *Nature* 2005; 438:1116-1122.

Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Pérez-Rosado A, Calvo E, Lopez JA, Cano A, Calasanz MJ, Colomer D, Piris MA, Ahn N, Imhof A, Caldas C, Jenuwein T, Esteller M. (2005) "Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer". *Nat Genet* 37(4)391–400.

Goldberg AD, Allis CD, Bernstein E. "Epigenetics: a landscape takes shape" *Cell*, 2007, 128(4): 635-8.

Greer, E. L. and Y. Shi (2012). "Histone methylation: a dynamic mark in health, disease and inheritance" *Nat Rev Genet* 13(5): 343-57.

Grignani F, De Matteis S, Nervi C, Tomassoni L, Gelmetti V, Cioce M, Fanelli M, Ruthardt M, Ferrara FF, Zamir I, Seiser C, Grignani F, Lazar MA, Minucci S, Pelicci PG. (1998) "Fusion proteins of the retinoic acid receptor-alpha recruit histone deacetylase in promyelocytic leukaemia". *Nature* 391(6669):815–818.

Hakimi MA, Dong Y, Lane WS, Speicher DW, Shiekhattar R. (2003) "A candidate X-linked mental retardation gene is a component of a new family of histone deacetylasecontaining complexes". *J. Biol. Chem*. 278(9):7234–7239

Harris WJ, Huang X, Lynch JT, Spencer GJ, Hitchin JR, Li Y, Ciceri F, Blaser JG,

Greystoke BF, Jordan AM, Miller CJ, Ogilvie DJ, Sommervaille TC. (2012) "The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells". *Cancer Cell* 21(4): 473–87

Heintzman ND, Stuart RK, Hon G, Fu Ym Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. (2007). "Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome" *Nat Genet* 39(3): 311-8.

Helin K and Dhanak D. (2013) "Chromatin proteins and modifications as drug targets" *Nature* 502(7472):480–8

Hodawadekar SC, Marmorstein R. (2007) "Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design". *Oncogene* 26(37):5528- 40.

Humphrey GW, Wang Y, Russanova VR, Hirai T, Qin J, Nakatani Y, Howard BH. (2001) "Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1". *J Biol Chem.* 276(9):6817–6824

Issa JP. and Kantarjian HM (2009). "Targeting DNA methylation." *Clin Cancer Res* 15(12): 393813946.

Kantarjian H, Oki Y, Garcia-Manero G, Huang X, O'Brien S, Cortes J, Faderl S, Bueso-Ramos C, Ravandi F, Estrov Z, Ferrajoli A, Wierda W, Shan J, Davis J, Giles F, Saba HI, Issa JP, (2007) "Results of a randomized study of 3 schedules of low-dose decitabine in higher-risk myelodysplastic syndrome and chronic myelomonocytic leukemia". *Blood,* 109:52–57

Kerenyi MA, Shao Z, Hsu YJ, Guo G, Luc S, O'Brien K, Fujiwara Y, Peng C, Nguyen M, Orkin SH (2013) "Histone demethylase Lsd1 represses hematopoietic stem and progenitor cell signatures during blood cell maturation" *eLife* 2:e00633.

Kahl P, Gullotti L, Heukamp LC, Wolf S, Friedrichs N, Vorreuther R, Solleder G, Bastian PJ, Ellinger J, Metzger E, Schule R, Buettner R. (2006). "Androgen receptor coactivators lysine1specific histone demethylase 1 and four and a half LIM domain protein 2 predict risk of prostate cancer recurrence." *Cancer Res* 66(23): 11341-7.

Khwaja A, Bjorkholm M, Gale RE, Levine RL, Jordan CT, Ehninger G, Bloomfield CD, Estey E, Burnett A, Cornelissen JJ, Scheinberg DA, Bouscary D, Linch DC.(2016) "Acute myeloid leukaemia", *Nat Rev Dis Primers* 10(2):16010

Kim J, Guermah M, McGinty RK, *et al.* RAD6-mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* 2009; 137:459-471.

Kooistra SM and Helin K (2012). "Molecular mechanisms and potential functions

of histone demethylases." *Nat Rev Mol Cell Biol* 13(5):297-311.

Kouzarides T. (2007) "Chromatin modifications and their function". *Cell* 128:693-705.

Kulis M. and Esteller M (2010). "DNA methylation and cancer." *Adv Genet* 70: 27156.

Lanotte M, Martin-Thouvenin V, Najman S, Balerini P, Valensi F, Berger R (1991). "NB4, a maturation inducible cell line with t(15;17) marker isolated from a human acute promyelocytic leukemia (M3)." *Blood* 77(5): 1080-6.

Lee JS, Shukla A, Schneider J, Swanson SK, Washburn MP, Florens L, Bhaumik SR, Shilatifard A. (2007) "Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS". *Cell* 131(6):1084-96.

Lee MG, Wynder C, Cooch N, Shiekhattar R. (2005) "An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation". *Nature* 437(7057), 432–435

Lynch JT, Harris WJ, Sommervaille TC (2012). "LSD1 inhibition: a therapeutic strategy in cancer?" *Expert Opin Ther Targets* 16(12):1239-49

Manuyakorn A, Paulus R, Farrell J, Dawson NA, Tze S, Cheung-Lau G, Hines OJ, Reber H, Seligson DB, Horvath S, Kurdistani SK, Guha C, Dawson DW (2010) "Cellular histone modification patterns predict prognosis and treatment response in resectable pancreatic adenocarcinoma: results from RTOG 9704". *J Clin Oncol* 28, 1358–1365.

Martens JH, Brinkman AB, Simmer F, Francojis KJ, Nebbioso A, Altucci L, Stunnenberg HG (2010). "PML-RARalpha/RXR Alters the Epigenetic Landscape in Acute Promyelocytic Leukemia." *Cancer Cell* 17(2): 173-85.

Martin C and Zhang Y (2005). "The diverse functions of histone lysine methylation." *Nat Rev Mol Cell Biol* 6(11): 8381849.

Marziali A and Akeson M. (2001). "New DNA sequencing methods*". Annu Rev Biomed Eng*, 3, 195-223

Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, Kloos DU, Land S, Lewicki-Potapov B, Michael H, Munch R, Reuter I, Totert S, Saxel H, Scheer M, Thiele S, Wingender E. (2003) "TRANSFAC: transcriptional regulation, from patterns to profiles". *Nucleic Acids Res* 31(1):374–8

McGrath JP, Williamson KE, Balasubramanian S, Odate S, Arora S, Hatton C, Edwards TM, O'Brien T, Magnuson S, Stokoe D, Daniels DL, Bryant BM, Trojer P (2016) "Pharmacological inhibition of the histone lysine

demethylase KDM1A suppresses the growth of multiple acute myeloid leukemia subtypes". *Cancer Res.* 76(7):1975–88

Meissner A, Mikkelsen TS, Gu H, Werning M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES (2008). "Genome-scale DNA methylation maps of pluripotent and differentiated cells." *Nature* 454(7205): 766-70.

Mercurio C, Minucci S & Pelicci PG (2010) "Histone deacetylases and epigenetic therapies of hematological malignancies". *Pharmacol Res* 62(1):18–34.

Metzger E, Wissman M, Yin N, Muller JM, Schneider R, Peters AHFM, Gunter T, Buettner R & Schule R. (2005) "LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription" *Nature*, 436-39

Mikesch JH, Gronemeyer H, So CWE (2010). "Discovery of novel transcriptional and epigenetic targets in APL by global ChIP analyses: Emerging opportunity and challenge." *Cancer Cell* 17(2): 112-114.

Minucci S and Pelicci PG (2006) "Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer." *Nat Rev Cancer* 6(1): 38151.

Minucci S and Pelicci PG (2007). "Determinants of oncogenic transformation in acute promyelocytic leukemia: the hetero1union makes the force." *Cancer Cell* 12(1): 113.

Mohammad HP, Smitheman KN, Kamat CD, Soong D, Federowicz KE, Van Aller GS, Schneck JL, Carson JD, Liu Y, Butticello M, Bonnette WG, Gorman SA, Degenhardt Y, Bai Y, McCabe MT, Pappalardi MB, Kasparec J, Tian X, McNulty KC, Rouse M, McDevitt P, Ho T, Crouthamel M, Hart TK, Concha NO, McHugh CF, Miller WH, Dhanak D, Tummino PJ, Carpenter CL, Johnson NW, Hann CL, Kruger RG (2015). "A DNA hypomethylation signature predicts antitumor activity of LSD1 inhibitors in SCLC". *Cancer Cell* 28(1):57–69

Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schubeler D (2008). "Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors." *Mol Cell* 30(6): 755-66.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods, 5(7), 621-628. doi: 10.1038/nmeth.1226

Mosammaparast N and Shi Y. (2010) " Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases" *Annu Rev Biochem*, 79:155-79

Musri MM, Carmona MC, Hanzu FA, Kaliman P, Gomis R, Pàrrizas M (2010). "Histone demethylase LSD1 regulates adipogenesis." *J Biol Chem* 285(39):30034-41.

Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SI (2001) ”Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly”. *Science*,292: 110–113.

Nasr, R, Guillemin MC, Ferhi O, Soilihi H, Peres L, Berthier C, Rousselot P, Robledo-Sarmiento M, Lallemand-Breitenbach V, Gourmel B, Vitoux D, Pandolfi PP, Rochette-Egly C, Zhu J, de Thé H. (2008). "Eradication of acute promyelocytic leukemia initiating cells through PML-RARA degradation." Nat Med 14(12): 1333-42.

Nebbioso A, Carafa V, Benedetti R & Altucci L (2012) Trials with ‘epigenetic’ drugs: an update. Mol Oncol 6, 657–682.

Nelson CJ, Santos-Rosa H, Kouzarides T. “Proline isomer- ization of histone H3 regulates lysine methylation and gene expression”. *Cell* 2006; 126:905-916.

Nguyen C, Liang G, Nguyen TT, Tsao-Wei D, Groshen S, Lubbert M, Zhou JH, Benedict WF, Jones PA (2001) “Susceptibility of nonpromoter CpG islands to de novo methylation in normal and neoplastic cells”. *J Natl Cancer Inst* 93(19):1465–72.

Okano M, Bell DW, Haber DA, Li E(1999). "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." *Cell* 99(3): 247-57.

Parthun MR.(2007) “Hat1: the emerging cellular roles of a type B histone acetyltransferase” *Oncogene* 2007; 26:5319-28.

Portales-Casamar E, Thongjuea S, Kwon AT, Arenillas D, Zhao X, Valen E, Yusuf D, Lenhard B, Wasserman WW, Sandelin A. (2010) “JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles”. *Nucleic Acids Res* ;38(Database issue):D105-10

Rada-Iglesias A, Bajpai Swigut T, Brugmann SA, Flynn RA, Wysocka J (2011). "A unique chromatin signature uncovers early developmental enhancers in humans." *Nature* 470(7333): 279-283.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013) “Genome engineering using the CRISPR-Cas9 system” *Nature protocols* 8:2281-2308

Sadikovic B, Al-Romaih K, Squire JA & Zielenska M (2008)” Cause and consequences of genetic and epigenetic alterations in human cancer. Curr Genomics 9, 394–408.

Saeed S, Logie C, Stunnenberg HG and Martens JHA. (2011) “Genome-wide functions of PML–RAR*α* in acute promyelocytic leukaemia” *British Journal of Cancer* 104, 554–558

Saleque S, Kim J, Rooke HM, Orkin SH (2007). "Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1." *Mol Cell* 27(4):562-72.

Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B. (2004) “JASPAR: an open-access database for eukaryotic transcription factor binding profiles”. *Nucleic Acids Res*. 32(Database issue):D91-4.

Sanjana NE, Shalem O, Zhang F (2014) “Improved vectors and genome-wide libraries for CRISPR screening” *Nat Methods* 11(8):783-4

Sanger, F, Nicklen S, Coulson AR (1992). “ sequencing with chain-terminating inhibitors”. 1977. *Biotechnology*, 24, 104-108.

Sanz, M. A. and F. Lo Coco (2011). "Modern approaches to treating acute promyelocytic leukemia." *J Clin Oncol* 29(5): 495-503.

Santoro F, Botrugno OA, Dal Zuffo R, Pallavicini I, Matthews GM, Cluse L, Barozzi I, Senese S, Fornasari L, Moretti S et al. (2013) “A dual role for Hdac1: oncosuppressor in tumorigenesis, oncogene in tumor maintenance”. *Blood* 121, 3459–3468.

Schenk T, Chen WC, Gollner S Howell L, Jin L, Hebestreit K, Klein HU, Popescu AC, Burnett A, Mills K, Casero RA jr, Marton L, Woster P, Minden MD, Dugas, Wang JCY, Dick JE, Muller-Tidow C, Petrie K & Zelent A. (2012) “Inhibition of the LSD1 (KDM1A) demethylase reactivates the all-trans-retinoic acid differentiation pathway in acute myeloid leukemia”. *Nat Med* 18(4), 605–611

Schones DE and Zhao K. (2008) “Genome-wide approaches to studying chromatin modifications” Nature reviews genetics, 9(3):179-91

Seligson DB, Horvath S, Shi T, Yu H, Tze S, Grunstein M & Kurdistani SK (2005) “Global histone modification patterns predict risk of prostate cancer recurrence”. *Nature* 435, 1262–1266.

Shi YJY, Matson C, Lan F, Iwase S, Baba T, Shi YJY. (2005) “Regulation of LSD1 histone demethylase activity by its associated factors.” *Mol Cell* 19(6):857–864

Shlush LI, Zandi S, Mitchell A, Chen WC, Brandwein JM, Gupta V, Kennedy JA, Schimmer AD, Schuh AC, Yee KW McLeod JL, Doedens M, Medeiros JFJ, Marke R, Kim HJ, Lee K, McPherson JD, Hudson TJ, The HALT Pan-Leukemia Gene Panel Consortium, Brown AMK, Youfis F, Trinh QM,

Stein LD, Minden MD, Wang JCY, Dick JE, (2014) “Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia”. *Nature* 506, 328–333.

Schulte JH, Lim S, Schramm A, Friedrichs N, Koster J, Versteeg R, Ora I, Pajter K, Klein-Hitpass L, Kuhfittig-Kulle S, Metzger E, Schule R, Eggert A, Buettner R, Kirfel J. (2009). "Lysine-specific demethylase 1 is strongly expressed in poorly differentiated neuroblastoma: implications for therapy." *Cancer Res* 69(5): 2065-71.

Singh MM, Johnson B, Venkatarayan A, Flores ER, Zhang J, Su X, Barton M, Lang F, Chandra J (2015)*”* Preclinical activity of combined HDAC and KDM1A inhibition in glioblastoma”. *Neuro. Oncol.* 17(11), 1463–73

Sprussel A, Schulte JH, Weber S, Necke M, Handschke K, Thor T, Pajtler KW, Schramm A, Konig K, Diehl L, Mestdagh P, Vandesompele J, Speleman F, Jastrow H, Heukamp LC, Schule R, Duhrsen U, Buettner R, Eggert A, Gothert JR (2012) “Lysine-specific demethylase 1 restricts hematopoietic progenitor proliferation and is essential for terminal differentiation”. *Leukemia* 26(9), 2039–2051

Stein, L. D. (2011). “An introduction to the informatics of "next-generation" sequencing”. *Curr Protoc Bioinformatics*, Chapter 11, Unit 11 11.

Tong JK, Hassig CA, Schnitzler GR, Kingston RE, Schreiber SL. (1998). “Chromatin deacetylation by an ATPdependent nucleosome remodelling complex*”. Nature* 395(6705): 917–921

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Baren MJV, Salzberg SL, Wold BJ Pachter L (2010) “transcript assembly and quantification by RNA.seq reveals unannotated transcripts and isoform switching during cell differentiation” *Nature Biotechnology* 28, 511-515

Wang Y, Zhang H, Chen Y, Sun Y, Yang F, Yu W, Liang J, Sun L, Yang X, Shi L, Li R, Li Y, Zhang Y, Li Q, Yi X, Shang Y (2009). “LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer”. *Cell* 138(4): 660–672

Wang ZY, Chen Z, (2008) “Acute promyelocytic leukemia: from highly fatal to highly curable”. *Blood*, 111(5):2505-15

Wang J, Scully K, Zhu X, Cai L, Zhang J, Prefontaine GG, Krones A, Ohgi KA, Zhu P, Garcia-Bassets I, Liu F, Taylor H, Lozach J, Jayes FL, Korach KS, Glass CK, Fu XD, Rosenfeld MG, (2007). "Opposing LSD1 complexes function in developmental gene activation and repression programmes." *Nature* 446(7138): 8821887.

Welch JS, Yuan W, Ley TJ, (2011). "PML-RARA can increase hematopoietic self-renewal without causing a myeloproliferative disease in mice." *J Clin Invest* 121(4): 1636-45.

Whyte WA, Bilodeau S, Orlando DA (2012) "Enhancer decommissioning by LSD1 during embryonic stem cell differentiation". *Nature* 482(7384), 221–225.

Yang XJ, Seto E. (2007) "HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention". *Oncogene* 26:5310-5318.

You A, Tong JK, Grozinger CM, Schreiber SL. (2001). "CoREST is an integral component of the CoREST–human histone deacetylase complex". *Proc. Natl Acad. Sci. USA* 98(4), 1454–8

Zambelli, F., G. Pesole, et al. (2009). "Pscan: finding over1represented transcription factor binding site motifs in sequences from co1regulated or co1expressed genes." *Nucleic Acids Res* 37(Web Server issue): W2471252.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) "Model-based Analysis of ChIP-seq (MACS)" Genome Biol 8(9)R137

Zhang XW, Yan XJ, Zhou ZR, Yang FF, Wu ZY, Sun HB, Liang WX, Song AX, Lallemand-Breitenbach V, Jeanne M, Zhang QY, Yang HY, Huang QH, Zhou GB. Tong JH, Zhang Y, Wu JH, Hu HY, de Thé H, Chen SJ, Chen Z (2010). "Arsenic trioxide controls the fate of the PML-RARalpha oncoprotein by directly binding PML." *Science* 328(5975): 2401 243.

Zeisig BB, Kwok C, Zelent A, Shankaranarayanan P, Gronemeyer H, Dong S, So CW (2007). "Recruitment of RXR by homotetrameric RARalpha fusion proteins is essential for transformation." *Cancer Cell* 12(1): 36151.

Zegerman P, Canas B, Pappin D, Kouzarides T.(2002) "Histone H3 lysine 4 methylation disrupts binding of nucleosome remod- eling and deacetylase (NuRD) repressor complex". *J Biol Chem* 277:11621-11624.

Zentner GE and Henikoff S. (2013) "Regulation of nucleosome dynamics by hitone modifications". *Nat Struct Mol Biol*, 20(3):259-66

Zhou VW, Goren A and Bernstein BE. (2010)"Charting histone modification and the functional organization of mammalian genomes" *Nature reviews genetics,* 12:7-18

Zhu JR, Nasr R, Pérès L, Riaucoux-Lormière F, Honoré N, Berthier C, Kamashev D, Zhou J, Vitoux D, Lavau C, de Thé H (2007). "RXR is an essential component of the oncogenic PML-RARA complex in vivo." *Cancer Cell* 12(1): 23135.

MINIREVIEW

BJC

British Journal of Cancer (2016), 1–7 | doi: 10.1038/bjc.2016.36

Keywords: leukaemia; histone deacetylases; HDACi; cancer stem cells; APL; epigenetics

# Inhibition of histone deacetylases in cancer therapy: lessons from leukaemia

Elena Ceccacci[1,2] and Saverio Minucci*[,1,2]

[1]Department of Experimental Oncology and Drug Development Program, European Institute of Oncology, Via Adamello 16, 20139 Milan, Italy and [2]Department of Biosciences, University of Milan, Via Celoria 26, 20133 Milan, Italy

Histone deacetylases (HDACs) are a key component of the epigenetic machinery regulating gene expression, and behave as oncogenes in several cancer types, spurring the development of HDAC inhibitors (HDACi) as anticancer drugs. This review discusses new results regarding the role of HDACs in cancer and the effect of HDACi on tumour cells, focusing on haematological malignancies, particularly acute myeloid leukaemia. Histone deacetylases may have opposite roles at different stages of tumour progression and in different tumour cell sub-populations (cancer stem cells), highlighting the importance of investigating these aspects for further improving the clinical use of HDACi in treating cancer.

Epigenetic mechanisms have a key role in the control of biological processes, and chromatin alterations may lead to the onset and progression of many diseases, first of all cancer.

Unlike genetic alterations, epigenetic alterations are generally reversible. For this reason, drugs acting against epigenetic targets (epidrugs) have been developed and some of them have been approved for selected cancer indications, thus validating the concept of epigenetic therapy.

Histones present a great number of modifications, including acetylation, methylation, phosphorylation, ubiquitination and many others, more recently identified. This pattern is sometimes referred as 'histone code', but the parallelism with a code represents an oversimplification, as similar combinations of histone marks may result in different functional outcomes depending on the context, and different combinations may lead to a similar functional result.

Among histone modifications, lysine acetylation depends on the antagonistic activity of two enzyme classes: histone acetylases (HATs) and histone deacetylases (HDACs, subject of this review).

The human HDAC family comprises 18 proteins that can be grouped into four classes on the basis of sequence homology with yeast proteins:

- *Class I* (HDAC 1-2-3-8) are homologous to the yeast Rpd3, localise in the nucleus and contain a single deacetylase domain at the N terminus.
- *Class II HDACs* can be further divided into two classes: IIa (HDAC 4-5-7-9, localised in the nucleus and cytoplasm) and IIb (HDAC 6 and 10); HDAC6 is mainly located in the cytoplasm and contains two catalytic domains, whereas HDAC10 contains a functional N-terminal domain and a C-terminal incomplete domain.
- Class III HDACs are also termed sirtuins (SIRT1–SIRT7), which are homologs of yeast Sirt2 and differ structurally from the other classes, requiring NAD+ as a cofactor.
- Class IV contains a single HDAC (HDAC11) with a catalytic domain shared with classes I/II HDACs.

Here, we will focus on Classes I/II and IV HDACs. These HDACs mediate their function as part of large macromolecular complexes in association with other factors: HDAC1 and HDAC2 are found in the mSin3A, NURD and Co-REST complexes, and HDAC3 is found associated with N-CoR and SMRT, whereas several proteins involved in the ubiquitin pathway are found associated with HDAC6 (Minucci and Pelicci, 2006).

## HDACS AND THE CONTROL OF HISTONE (AND NON-HISTONE) ACETYLATION

The addition of acetyl groups to lysine residues in the histone tails by HATs is responsible for a relaxed and accessible chromatin structure, and is associated with transcriptional activation; conversely, HDACs remove acetyl groups and lead to a more closed chromatin structure, generally associated with transcriptional repression.

*Correspondence: Professor S Minucci; E-mail: Saverio.minucci@ieo.eu

Received 24 July 2015; Received 7 January 2016; accepted 12 January 2016

Genome-wide studies of HDACs by chromatin immunoprecipitation followed by next-generation sequencing (Chip-Seq) revealed, however, a strong association between HDACs and active genes, and suggested a role for HDACs also in active transcription: when associated with active genes, HDACs act to remove acetyl groups added by HATs during transcriptional initiation and elongation, providing a reset of the chromatin structure that is required for a second round of transcription. Indeed, excessive histone acetylation at transcribed regions could 'destabilise' chromatin, thus leading to an increase of transcription at erroneous starting sites (Wang *et al*, 2009). In *Drosophila*, the UpSET complex—that includes HDACs—is required to reduce spreading of histone acetylation from active promoter regions and therefore limiting transcriptional noise (Rincon-Arano *et al*, 2012). The canonical relationship between HDACs, low levels of histone acetylation and transcriptional repression is therefore not always valid.

Histone acetylases and HDACs are responsible for the reversible acetylation not only of histones but also of a large number of additional substrates such as transcription factors, DNA repair enzymes and nuclear and cytoplasmic proteins. Not all protein acetylation, however, necessarily derives from the enzymatic action of HATs/HDACs: protein acetylation in mitochondria may be a chemical event facilitated by the alkaline pH and high concentrations of reactive acetyl-CoAs present in the mitochondrial matrix (Wagner and Payne, 2013).

Acetylation of non-histone proteins and regulation of their function adds another layer of complexity to the action of HATs/HDACs: importantly, this also precludes from considering HDACs as purely 'epigenetic factors' (Minucci and Pelicci, 2006).

Proteomic studies have very recently led to the discovery of a large number of novel histone post-translational modifications (PTMs) that show additional acyl moieties beside acetylation (Kebede *et al*, 2015); among those PTMs, propionylation, butyrylation, crotonylation, succinylation, malonylation, glutarylation and lysine 2-hydroxyisobutyrylation share the use of short-chain acyl-coAs derived from energy metabolism as cosubstrate(s). These histone PTMs contribute to transcriptional regulation by promoting DNA unwrapping and nucleosome disassembly, reducing nucleosome stability and influencing the action of chromatin-associated factors: intriguingly, HDACs are able to remove at least a subset of these newly discovered acyl-histone marks (Kebede *et al*, 2015).

Currently, it remains unclear whether concentration of coAs (depending on the metabolic status of the cell) could be a critical factor determining the type of histone acylation, and if HDACs could have a role in transmitting metabolic signals by modulating the chromatin structure (a function already known for class III HDACs, regulated by NAD + /NADH ratio) and which impact metabolism can have on histone modifications. Importantly, tumour cells show constant alterations in metabolism that may lead, therefore, to alterations in HDAC function, as recent studies begin to suggest (Chiaradonna *et al*, 2015).

## HDACS IN CANCER AND A (WEAK) RATIONALE FOR THE DEVELOPMENT OF HDAC INHIBITORS

Given their pleiotropic roles and their involvement in essentially all cell functions, HDACs may not be considered at a first glance as attractive targets for therapy, owing to the likely interference of HDAC inhibitors with several processes occurring in normal cells and therefore high risk of side effects. Indeed, the first HDAC inhibitors were initially characterised for their antitumour activity *in vitro* before the discovery that they were known to inhibit HDACs: their use in preclinical models (*in vitro* and *in vivo*) showed a significant therapeutic window, with reduced effects on normal cells.

The discovery of the potent antitumoral effects following HDAC inhibition led to the hypothesis that HDACs themselves may act as oncogenes, and, in fact, distinct HDACs are found overexpressed in various solid tumours, in some cases showing a differential expression in tumour subtypes: as an example, HDAC1 is highly expressed in hormone receptor-positive breast cancer, whereas HDAC2 and HDAC3 are more expressed in breast cancers with a more aggressive phenotype including hormone receptor-negative cancers (Müller *et al*, 2013). It is difficult, however, to correlate the degree of expression of individual HDACs with particular functional consequences and biological phenotypes: loss of acetylation of lysine 16 and 20 of histone H4 has been observed in various cancer cell lines and primary tumours, and described as a hallmark of human cancer: this could be because of overexpression of the class III HDAC SIRT1, which is capable of deacetylating histone H4K16 (Fraga *et al*, 2005). Reduced acetylation because of enhanced expression of HDACs may lead to transcriptional repression of tumour-suppressive pathways, including cell cycle regulators and DNA repair pathways. Overexpression of HDACs has been linked (together with other histone modifications) to the epigenetic repression of the locus encoding for the tumour suppressor *CDKN1A*, and of DNA damage repair genes such as *BRCA1* and *ATR* (Eot-Houllier *et al*, 2009), but the consequence of HDAC overexpression may also impact on non-histone substrates: HDAC1 and HDAC2 regulate acetylation of the oncosuppressor p53, thus inhibiting its function (Insinga *et al*, 2004).

Overexpression of HDACs has been proposed in several cases as a negative prognostic marker, independently of tumour type and disease progression; however, this is not always the case as high levels of HDAC6 predict better prognosis in ER-positive breast cancer (Saji *et al*, 2005), or in CTCL.

Although HAT encoding genes are found frequently mutated or amplified in cancer, with consequent loss or alteration of function, there are reports of only rare, inactivating mutations of HDACs in cancer, which may suggest oncosuppressive roles for HDACs (Ropero *et al*, 2006).

We performed a preliminary analysis of the mutational landscape of HDACs using available data sets, focusing on missense mutations in the coding sequence, without further investigating the mutational impact of each mutation (Figure 1 and Supplementary Table S1).

The frequency of missense mutations of HDACs varies greatly in different human cancers: although in some cases (melanoma, lung cancer) the high frequency of mutations (close to 30% considering all HDACs) is somewhat expected because of the general high mutation rate caused by exposure to carcinogens (sunlight, smoke), other cancer types show a rate of HDAC mutations that could only in part be explained by their general mutational trend. In general, class II HDACs show higher percentages of mutations, and the distribution of missense mutations in the two most mutated HDACs, HDAC4 and HDAC9, show that although most mutations are distributed along the entire coding sequence, and therefore they are unlikely to be functionally relevant, a few sites – located in the catalytic domain – mutated in a higher number of patients exist, suggesting a potential alteration of the enzyme function (Figure 1B).

Taken together, these observations (that need to be extended) hint that just looking at mutations and/or altered expression pattern will not provide frequently a conclusive answer, and that more in-depth mechanistical insights are required to understand the altered function of HDACs in cancer cells.

Indeed, HDAC inhibition affects markedly cancer cells, inducing (depending on drug, dosage and tumour cell type) cell cycle arrest, differentiation, induction of cell death, reduction of

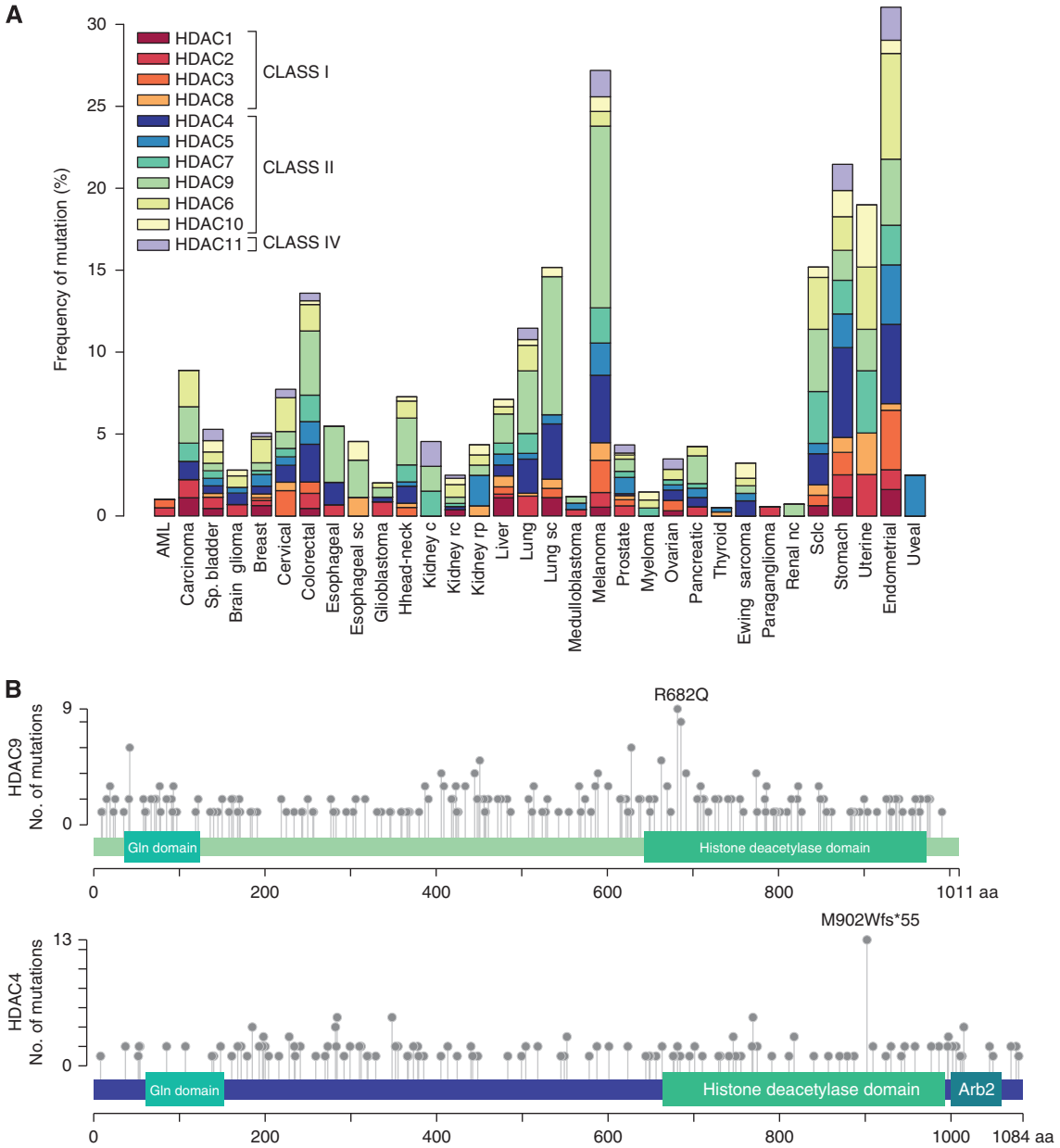



**Figure 1. HDACs class I, II, IV mutations in human cancer.** (**A**) Histone deacetylases class I, II and IV mutations across different human cancers. The histogram shows an overview of the frequency of mutations (missense) of each HDAC (classes I, II and IV) across different human cancers (analysis was performed on data downloaded from cBioPortal, see Supplementary Table S1 for a guide to the abbreviations, and the description for each cancer subtype of sample size). (**B**) Distribution of mutations across HDAC9 and HDAC4 coding sequences. Lolliplot graph of missense mutations found across all human cancers for HDAC9 (upper panel) and HDAC4 (lower panel). Note that though mutations are equally distributed along the entire coding sequence, in both cases the most frequent mutation is localised within the histone deacetylase catalytic domain. For HDAC4, in 13 patients the same mutation introduces a frameshift in the middle of the deacetylase domain, leading most likely to a functional inactivation of the domain. The two graphs were obtained by using the cBioPortal tool 'MutationMapper' (Cerami *et al*, 2012).

angiogenesis and modulation of the immune system. An 'epigenetic vulnerability' of tumour cells has been proposed, where – in contrast to normal cells that show redundancy in epigenetic regulatory mechanisms – HDACs may be essential in tumour cells for the maintenance of a set of key genes required for survival and growth (Dawson and Kouzarides, 2012).

A large number of HDAC inhibitors has been synthesised and tested in clinical trials, resulting in the approval of four inhibitors (Vorinostat, Romidepsin, Bellinostat, Panobinostat: a list of HDAC inhibitors, their classification and clinical status is provided in Table 1). This could be interpreted as a successful history of drug development, with validation of HDACs as important targets in cancer, but the situation is far more complex, and the clinical results do not reflect those expected from the preclinical work, both in terms of efficacy (observed only in selected cancer subtypes, mainly in haematology) and safety (several side effects were observed, among which the most common are fatigue, diarrhoea, bone marrow toxicity, thrombocytopenia; Subramanian *et al*, 2010). The reasons for this, at least in part, disappointing set of clinical results are not clear; one explanation could be the lack of selectivity of most of the HDAC inhibitors tested clinically and approved to date, acting as paninhibitors on all HDAC classes (not including sirtuins): the global inhibition of several non-redundant HDACs, with

**Table 1. HDAC inhibitors classified according to: (a) status of clinical advancement; (b) HDAC(s) targeted and (c) chemical class**

| Compounds | Target | Class | Highest phase trial |
|---|---|---|---|
| Panobinostat (LBH-589) | Pan-HDAC inhibitor | Hydroxamic acids | Approved in 2015 for multiple myeloma |
| Belinostat (PXD101) | Pan-HDAC inhibitor | Hydroxamic acids | Approved in 2014 for PTCL |
| Romidepsin (desipeptide-FK228) | Pan-HDAC inhibitor | Cyclic tetrapeptides | Approved in 2009 for CTCL |
| SAHA(Vorinistat, Zolinza) | Pan-HDAC inhibitor | Hydroxamic acids | Approved in 2006 for CTCL |
| Valproic acid | Pan-HDAC inhibitor | Short-chain fatty acids | Phase III |
| Tacedinaline (CI994) | Subclass I-selective inhibitor (HDACs 1, 2 and 3) | Benzamides | Phase III |
| Givinostat (ITF2357) | Pan-HDAC inhibitor | Hydroxamic acids | Phase II |
| Resminostat (4SC201) | Pan-HDAC inhibitor | Hydroxamic acids | Phase II |
| Abexinostat (PCI24781) | Pan-HDAC inhibitor | Hydroxamic acids | Phase II |
| Rocilinostat (ACY1215) | Selective class II HDAC inhibitor | Hydroxamic acids | Phase II |
| Quisinostat (JNJ-26481585) | Pan-HDAC inhibitor | Hydroxamic acids | Phase II |
| Practinostat (SB939) | Inhibit class I, II and IV HDACs | Hydroxamic acids | Phase II |
| Mocetinostat (MGCD0103) | Specific against class I and IV HDACs | Benzamides | Phase II |
| Entinostat (MS275-SNDX-275) | Class I HDAC inhibitor | Benzamides | Phase II |
| Sodium phenylbutyrate | Inhibit class I and II HDACs | Short-chain fatty acids | Phase II |
| AR42 | Pan-HDAC inhibitor | Hydroxamic acids | Phase I |
| 4SC202 | Selective class I HDAC inhibitor | Benzamides | Phase I |
| Pyroxamide (NSC696085) | Inhibitor of affinity-purified HDAC1 | Hydroxamic acids | Phase I |
| CHR-3996 | Selective class I HDAC inhibitor | Hydroxamic acids | Phase I |
| CHR-2845 | | Hydroxamic acids | Phase I |

Abbreviations: CTCL = cutaneous T-cell lymphoma; HDAC = histone deacetylase; PTCL = peripheral T-cell lymphoma. Adapted from Valente and Mai (2014), http://www.fda.gov/default.htm.

partially overlapping but with clearly distinct and sometimes contrasting functions, may result in a difficult to predict phenotype, and subtle differences among species may explain the different results observed in murine models.

Indeed, knockout studies in mice on all members of class I HDACs demonstrated the unique roles of each HDAC in the control of specific gene expression programmes: HDAC1-null mice die at day E10.5 and display severe proliferation defects and general growth retardation; HDAC2-null mice die 24 h after birth for cardiac malformations; HDAC3-null mice die before E9.5 for defects in gastrulation probably because of defective DNA repair (Haberland *et al*, 2009; Yang and Seto, 2008).

The hypothesis that inhibition of specific HDACs may have a better therapeutic outcome will be put to test once we will evaluate more selective HDAC inhibitors in clinical trials (that is supposed to happen soon).

## HDACS IN APL: DISTINCT ROLES IN SPACE AND IN TIME

In contrast to solid tumours, we have a better understanding of the altered function of HDACs in haematological malignancies.

Here, a revisitation of recent results, mainly focusing on acute myeloid leukaemia (AML) models, lead us to additional reasons to explain the complexity of the use of HDAC inhibitors in the clinical setting, linked to distinct effects of HDACs at different stages of tumourigenesis and different action of HDAC inhibitors in distinct tumour cell subtypes.

Acute promyelocytic leukaemia (APL) was one of the first diseases in which the involvement of HDACs was demonstrated mechanistically. Acute promyelocytic leukaemia is characterised by the block of myeloid differentiation at the promyelocytic stage, and is associated with the chromosomal translocation t(15;17), to generate the PML-RAR fusion protein of retinoic acid receptor-α (RAR) with the promyelocytic leukaemia protein (PML). Retinoic acid (RA) showed clinical efficacy in APL patients before the demonstration that it acted by direct targeting of the oncogenic fusion protein (Minucci and Pelicci, 2006).

Subsequent molecular studies placed RA treatment of APL among the best characterised examples of both 'transcription therapy' (whereby the drug targets specifically the oncogenic transcription factor and its aberrant action) and 'differentiation therapy', which reprograms leukaemic cells for terminal differentiation (Tallman *et al*, 1997; Huang *et al*, 1988). For all these reasons, although it is a rare disease, APL has been for several years an important model system for learning lessons that can be potentially expanded to other forms of cancer.

In normal cells, RAR acts as a transcription factor, regulating myeloid differentiation and binding in a heterodimeric form with the retinoid X receptor-specific DNA sequences (called RA-responsive element) found at RAR target genes. In the absence of RA, RAR is found in association with HDAC-containing complexes and represses transcription: RA leads to a conformational switch that causes the release of the corepressor complexes and binding of transcriptional coactivators, with consequent transcription of RAR target genes.

In APL cells, physiological concentrations of RA do not result in the release of HDAC–corepressor complexes from PML-RAR, leading to altered regulation of RAR target genes and of additional PML-RAR-specific targets, and subsequent differentiation block (Minucci and Pelicci, 2006), whereas pharmacological doses of RA (10- to 100-fold higher than physiological concentrations) reverse the action of the fusion protein, owing to induction of its degradation, and lead to reactivation of the differentiation programme of APL cells. In patients, however, treatment with RA leads only to a transient remission of the disease but cannot entail a definitive cure unless RA is combined to other drugs such as chemotherapy or arsenic.

Studies conducted in transgenic mice have demonstrated that the presence of PML-RAR alone is not sufficient by itself to confer full leukaemic potential to haematopoietic cells, and *second hits* are necessary for leukaemia development; this thus defines a preleukaemic phase where PML-RAR is acting initially, molecularly distinct from the clonal leukaemia that develops at a later stage.

**HDACs in time.** The model depicted before suggests a critical role for HDACs in APL, as they are required for the fusion protein to arrest leukaemic differentiation, but new results however have shown that the role of HDACs in APL is more complex than previously thought and demonstrated a dual role for HDACs dependent on the stage of disease progression (so changing 'in time').

Indeed, mice transplanted with haematopoietic progenitors derived from PML-RAR transgenic mice, and carrying knockdown of either HDAC1 or HDAC2, showed a strongly reduced preleukaemic phase, with accelerated leukaemia development

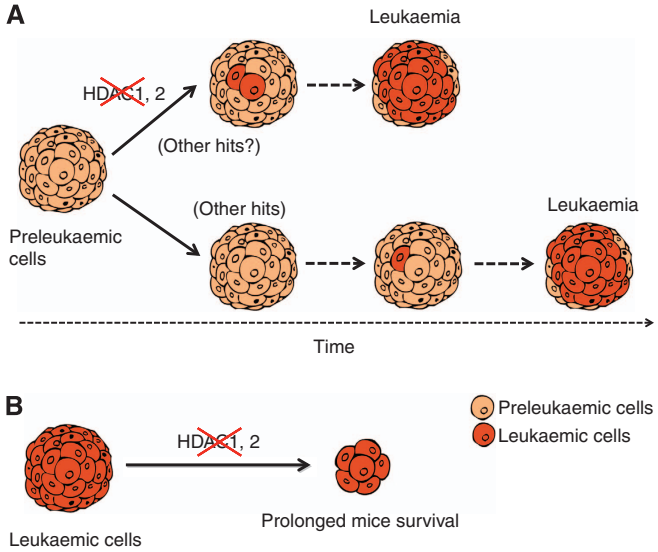



Figure 2. A time-dependent role for HDACs in leukaemia development. (**A**) During the preleukaemic phase of APL, HDAC1/2 act as tumour suppressors, and their knockdown results in accelerated leukaemia development. This can be because of higher frequency of additional hits, or to direct transformation of PML-RAR preleukaemic cells. (**B**) During the leukaemic phase, knockdown of HDAC1/2 cause differentiation and then apoptosis of APL cells, with an extended lifespan of the leukaemic mice.

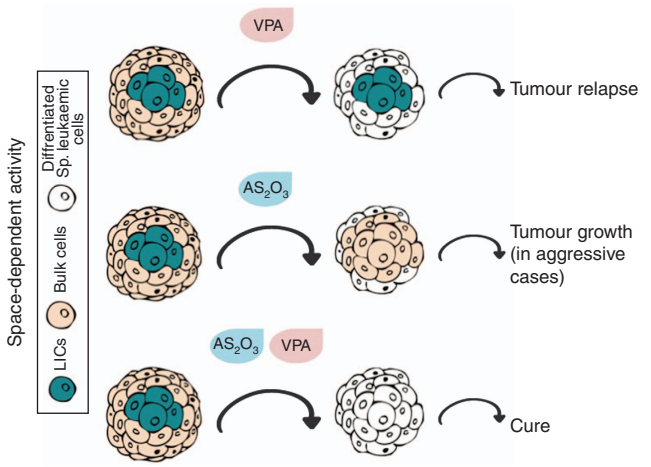



Figure 3. Histone deacetylase inhibitors have distinct effects on tumour cell sub-populations. (Upper panel) Histone deacetylase inhibition by VPA treatment of APL mice results in differentiation of bulk leukaemic cells and prolonged survival, but as LICs are not targeted disease recurs. (Middle panel) Treatment with arsenic leads to tumour regression because of significant reduction in the number of LICs, and progressive tumour exhaustion, but tumour growth in the short term may continue because of bulk of leukaemic cells and – in high-risk aggressive disease forms – lead to patient death. (Bottom panel) Combining the two treatments, by acting on both tumour cell sub-populations, may offer the best perspective in terms of disease control and LIC eradication.

(Figure 2A). Interestingly, HDAC1/2 knockdown caused right after transplantation a marked increase in cells (G + K + ) characterised by coexpression of differentiation markers (GR1) and more immature markers (C-kit). This cell sub-population in the absence of HDAC knockdown is found markedly expanded in the leukaemic stage, and is enriched in leukaemia-initiating cells (LICs: see below). Treating mice in the preleukaemic phase with the HDAC inhibitor (HDACi) valproic acid (VPA) mimicked the effect of HDAC1/2 knockdowns.

These results are in striking contrast with those obtained in the leukaemic phase, where knockdown of the same HDACs caused differentiation and apoptosis of APL cells, leading to prolonged mice survival (Figure 2B). Valproic acid treatment induces selectively in leukaemic cell differentiation followed by apoptosis because of the activation of the death receptor pathway. Taken together, these results imply a dual role of HDAC1 and HDAC2 in APL initiation and maintenance, and suggest that they may act as oncosuppressors in the preleukaemic phase, and as oncogenes in leukaemia (Santoro *et al*, 2013) (Figure 2). Interestingly, HDAC1/HDAC2 knockdown or knockout accelerated development of other tumour types (lymphomas and skin tumours), suggesting that this oncosuppressive role may be more general (Winter *et al*, 2013).

Not all HDACs share this dual, time-dependent function: HDAC3 acts as an oncogene also during the preleukaemic phase, and its knockdown or inhibition by selective drugs leads to cell differentiation and enhanced apoptosis, and lack of leukaemia development (Matthews *et al*, 2015).

Several questions remain unanswered: (I) How do HDAC1/HDAC2 contribute to decrease in the oncogenic potential of PML-RAR-expressing cells in preleukaemia? (II) Is HDAC1/2 loss a second hit sufficient to transform cells, or further hits are required? (III) Most importantly, as HDACs may have an oncosuppressive function, does this imply that clinical treatment with HDACi (such as selective HDAC1i/HDAC2i from those studies discussed above) may favour secondary cancers?

**HDACs in space.** Most cancers are heterogeneous, and the continuous expansion of the tumour mass is sustained by the self-renewing properties of a sub-population termed 'cancer stem cells' (CSCs) (in leukaemia CSCs are also called LICs) (Kreso and Dick, 2014). The inability of existing therapies (such as chemotherapy) to eradicate CSCs, and to act mainly on the bulk of the tumour mass (that does not proliferate indefinitely) is thought to be one of the most relevant causes for recurrence. Experimental protocols have been set up to measure LICs, based on limiting transplantation experiments in recipient mice of leukaemic cells and recent studies have started to explore the effect of HDACi on different tumour sub-populations in APL.

Valproic acid treatment of APL mice extend their survival, but shortly after interruption of treatment the disease relapses, leading to death; studies to measure the effect of VPA on LICs showed that the number of LICs was not affected by the treatment, consistently with the observed relapse. Valproic acid was therefore selectively acting on the bulk of leukaemic blasts, pushing their differentiation and apoptosis without affecting the self-renewal potential of LICs (Leiva *et al*, 2012) (Figure 3, upper panel). It remains to be seen if HDACi with different specificities show a different behaviour, and whether genetic experiments (knockdown) will be consistent with the pharmacological observations. We speculate that other HDACs and HDAC-containing complexes not tackled by VPA are essential for LIC maintenance, and therefore HDACs may act differentially in specific tumour cell sub-populations (different tumour 'space').

If we analyse the effect of other drugs on LICs used in APL patients, arsenic treatment was shown to act strongly on LICs (Figure 3, middle panel), and RA showed a dose-dependent phenotype, being more effective on bulk cells at lower doses (Nasr *et al*, 2008).

It will be of great interest to test in preclinical models the combination of VPA with drugs acting on LICs (such as arsenic trioxide): VPA alone cannot eliminate the disease as it is not targeting LICs, but targeting LICs alone may not necessarily work, as bulk tumour cells can still contribute significantly in the short

term to tumour growth, leading to patient death before the effect of LIC clearance can be appreciated clinically. Only the combined targeting of the entire tumour cell mass (LICs and bulk) could lead to disease eradication effectively and rapidly, and represents the best potential for cure (Figure 3, bottom panel).

## CONCLUSIONS

Pan-HDACi have given favourable results in a small set of patients with selected haematological diseases, but their use in monotherapy has not been satisfactory. The difference in sensitivity to HDACi cannot be easily allocated to a single cause, making it difficult to envision a smart approach to patient stratification. However, we believe that despite the disappointing results, this field deserves further study and remains a promising therapeutic avenue. Soon, we will know whether more selective HDACi will be more effective in the clinics, and with reduced side effects. The studies in murine models of leukaemia suggest that it is necessary to consider not only the differences among different classes of HDACs but also how the same molecules may act in 'time' and 'space', as we have previously illustrated. In particular, we propose that a systematic effort should be performed to study the effects of HDACi and other epidrugs on the stem cell compartment *vs* the rest of tumour cells, to devise treatment schemes that combine more efficiently drugs targeting the different tumour cell subpopulations. Of course, this must not be limited to epidrugs, and combination with other agents such as DNA-damaging chemotherapeutic drugs, or proteasome inhibitors, has already shown promising results that could be reinterpreted based on the studies proposed above. Histone deacetylase inhibitors may well find their optimal 'clinical space' in the end.

## ACKNOWLEDGEMENTS

We thank Giorgio Melloni for the essential help in the mutational analysis of HDACs, and Marina Mapelli for suggestions. Work in SM's laboratory is supported by AIRC (Italian Association for Cancer Research), FIRC (Italian Foundation for Cancer Research), Epigen Flagship Project, European Community (4D Cell Fate and Blueprint) and Fondazione Veronesi. For lack of space, it was not possible to mention all of the original papers. A more complete list of references can be found in the following reviews: Minucci and Pelicci (2006), Botrugno *et al* (2009), Mercurio *et al* (2010), Falkenberg and Johnstone (2014) and West and Johnstone (2014).

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## REFERENCES

Botrugno OA, Santoro F, Minucci S (2009) Histone deacetylase inhibitors as a new weapon in the arsenal of differentiation therapies of cancer. *Cancer Lett* **280**(2): 134–144.

Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discovery* **2**(5): 401–404.

Chiaradonna F, Cirulli C, Palorini R, Votta G, Alberghina L (2015) New insights into the connection between histone deacetylases, cell metabolism, and cancer. *Antioxid Redox signal* **23**(1): 30–50.

Dawson MA, Kouzarides T (2012) Cancer epigenetics: from mechanism to therapy. *Cell* **150**(1): 12–27.

Eot-Houllier G, Fulcrand G, Magnaghi-Jaulin L, Jaulin C (2009) Histone deacetylase inhibitors and genomic instability. *Cancer Lett* **274**(2): 169–176.

Falkenberg KJ, Johnstone RW (2014) Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat Rev Drug Discov* **13**(9): 673–691.

Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Perez-Rosado A, Calvo E, Lopez JA, Cano A, Calasanz MJ, Colomer D, Piris MA, Ahn N, Imhof A, Caldas C, Jenuwein T, Esteller M (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* **37**(4): 391–400.

Haberland M, Montgomery RL, Olson EN (2009) The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat Rev Genet* **10**(1): 32–42.

Huang ME, Ye YC, Chen SR, Chai JR, Lu JX, Zhoa L, Gu LJ, Wang ZY (1988) Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia. *Blood* **72**: 567–572.

Insinga A, Monestiroli S, Ronzoni S, Carbone R, Pearson M, Pruneri G, Viale G, Appella E, Pelicci P, Minucci S (2004) Impairment of p53 acetylation, stability and function by an oncogenic transcription factor. *EMBO J* **23**(5): 1144–1154.

Kebede AF, Schneider R, Daujat S (2015) Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest. *FEBS J* **282**(9): 1658–1674.

Kreso A, Dick JE (2014) Evolution of the cancer stem cell model. *Cell Stem Cell* **14**(3): 275–291.

Leiva M, Moretti S, Soilihi H, Pallavicini I, Peres L, Mercurio C, Dal Zuffo R, Minucci S, de The H (2012) Valproic acid induces differentiation and transient tumor regression, but spares leukemia-initiating activity in mouse models of APL. *Leukemia* **26**(7): 1630–1637.

Matthews GM, Mehdipour P, Cluse LA, Falkenberg KJ, Wang E, Roth M, Santoro F, Vidacs E, Stanley K, House CM, Rusche JR, Vakoc CR, Zuber J, Minucci S, Johnstone RW (2015) Functional–genetic dissection of HDAC dependencies in mouse lymphoid and myeloid malignancies. *Blood* **126**(21): 2392–2403.

Mercurio C, Minucci S, Pelicci PG (2010) Histone deacetylases and epigenetic therapies of hematological malignancies. *Pharmacol Res* **62**(1): 18–34.

Minucci S, Pelicci PG (2006) Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nat Rev Cancer* **6**(1): 38–51.

Müller BM, Jana L, Kasajima A, Lehmann A, Prinzler J, Budczies J, Winzer KJ, Dietel M, Weichert W, Denkert C (2013) Differential expression of histone deacetylases HDAC1, 2 and 3 in human breast cancer – overexpression of HDAC2 and HDAC3 is associated with clinicopathological indicators of disease progression. *BMC Cancer* **13**: 215.

Nasr R, Guillemin MC, Ferhi O, Soilihi H, Peres L, Berthier C, Rousselot P, Robledo-Sarmiento M, Lallemand-Breitenbach V, Gourmel B, Vitoux D, Pandolfi PP, Rochette-Egly C, Zhu J, de Thé H (2008) Eradication of acute promyelocytic leukemia-initiating cells through PML-RARA degradation. *Nat Med* **14**(12): 1333–1342.

Rincon-Arano H, Halow J, Delrow JJ, Parkhurst SM, Groudine M (2012) UpSET recruits HDAC complexes and restricts chromatin accessibility and acetylation at promoter regions. *Cell* **161**(6): 1214–1228.

Ropero S, Fraga MF, Ballestar E, Hamelin R, Yamamoto H, Boix-Chornet M, Caballero R, Alaminos M, Setien F, Paz MF, Herranz M, Palacios J, Arango D, Orntoft TF, Aaltonen LA, Schwartz Jr S, Esteller M (2006) A truncating mutation of HDAC2 in human cancers confers resistance to histone deacetylase inhibition. *Nat Genet* **38**(5): 566–569.

Saji S, Kawakami M, Hayashi S, Yoshida N, Hirose M, Horiguchi S, Itoh A, Funata N, Schreiber SL, Yoshida M, Toi M (2005) Significance of HDAC6 regulation via estrogen signaling for cell motility and prognosis in estrogen receptor-positive breast cancer. *Oncogene* **24**(28): 4531–4539.

Santoro F, Botrugno OA, Dal Zuffo R, Pallavicini I, Matthews GM, Cluse L, Barozzi I, Senese S, Fornasari L, Moretti S, Altucci L, Pelicci PG, Chiocca S, Johnstone RW, Minucci S (2013) A dual role for Hdac1: oncosuppressor in tumorigenesis, oncogene in tumor maintenance. *Blood* **121**(17): 3459–3468.

Subramanian S, Bates S, Wright J, Espinoza-Delgado I, Piekarz R (2010) Clinical toxicities of histone deacetylase inhibitors. *Pharmaceuticals* **3**: 2751–2767.

Tallman MS, Andersen JW, Schiffer CA, Appelbaum FR, Feusner HJ, Ogden A, Schepherd L, Willman C, Bloomfield CD, Rowe JM, Wiernik PH (1997) All-*trans*-retinoic acid in acute promyelocytic leukemia. *N Engl J Med* **337**(15): 1021–1028.

Valente S, Mai A (2014) Small-molecule inhibitors of histone deacetylase for the treatment of cancer and non-cancer disease: a patent review (2011–2013). *Expert Opin Therap Patents* **24**(4): 401–415.

Wagner GR, Payne RM (2013) Widespread and enzyme-independent *Nε*-acetylation and *Nε*-succinylation of proteins in the chemical conditions of the mitochondrial matrix. *J Biol Chem* **288**(40): 29036–29045.

Wang Z, Zang C, Cui K, Schones DE, Barski A, Peng W, Zhao K (2009) Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**(5): 1019–1031.

West AC, Johnstone RW (2014) New and emerging HDAC inhibitors for cancer treatment. *J Clin Invest* **124**(1): 30–39.

Winter M, Moser MA, Meunier D, Fischer C, Machat G, Mattes K, Lichtenberger BM, Brunmeir R, Weissmann S, Murko C, Humer C, Meischel T, Brosch G, Matthias P, Sibilia M, Seiser C (2013) Divergent roles of HDAC1 and HDAC2 in the regulation of epidermal development and tumorigenesis. *EMBO J* **32**(24): 3176–3191.

Yang XJ, Seto E (2008) The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and man. *Nat Rev Mol Cell Biol* **9**(3): 206–218.



Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

# Self-renewal of tumor cells: epigenetic determinants of the cancer stem cell phenotype

Roberto Ravasio[1], Elena Ceccacci[1,2] and Saverio Minucci[1,2,3]

Among the functional subpopulations that coexist within the tumor, 'cancer stem cells' are characterized by increased self-renewal and the ability to derive all of the other subpopulations of tumor cells ('bulk'). The functional heterogeneity among cancer stem cells and bulk cells must reflect distinct cellular epigenetic landscapes, but — due to the difficulty to isolate bona fide cancer stem cells with a high degree of purity — those different epigenetic landscapes, and the molecular mechanisms underlying them, remain largely unknown. Cues of intratumor phenotypic plasticity complicate the interpretation of the cancer stem cell phenotype: we contend that, however, the concept of cancer stem cell has crucial therapeutic implication, and remains a key target for the exploration of the cancer epigenome.

**Addresses**
[1] Department of Experimental Oncology, European Institute of Oncology, via Adamello 16, 20139 Milan, Italy
[2] Department of Biosciences, University of Milan, via Celoria 26, 20100 Milan, Italy
[3] Drug Development Program, European Institute of Oncology, via Ripamonti 435, 20100 Milan, Italy

Corresponding author: Minucci, Saverio (saverio.minucci@ieo.eu)

**Current Opinion in Genetics & Development** 2016, **36**:92–99

This review comes from a themed issue on **Cancer genomics**

Edited by **Luciano Di Croce** and **Ali Shilatifard**

For a complete overview see the Issue and the Editorial

Available online 7th May 2016

http://dx.doi.org/10.1016/j.gde.2016.04.002



## Introduction: the cancer stem cell phenotype

Cancer is characterized by extensive genetic and functional heterogeneity. The progress of sequencing technologies allows now to appreciate the complex mutational landscape of the tumor cell mass that consists of a large number of genetically distinct subclones, progressively evolving through Darwinian selection mechanisms both intrinsic, and related to therapy [1]. The degree of genetic complexity is somewhat unexpected, and bears daunting implications in view of the potential approaches to cancer therapy.

Long before these studies, however, differences in the behavior of cancer cells that cannot be ascribed to genetic mutations alone have been identified as an additional source of heterogeneity. Here, in contrast to the heterogeneity observed in normal cells that follow ordered developmental programs and respond to defined environmental pressures, the functional heterogeneity of cancer cells is assumed to be largely disordered and following more self-autonomous, scarcely controlled regulatory mechanisms.

Cancer stem cells (CSCs) are supposedly the cells within the tumor cell mass that retain self-renewing capacity and lie at the root of the hierarchy that defines the heterogeneity of all cells that comprise the tumor. Many authors define CSCs on the basis of their functional ability to initiate a tumor in a transplantation assay in mice (when human tumor cells are used, the assay is performed in immunocompromised mice). In many cases (leukemias, breast, lung and colon cancer among the best studied), the number of cells able to initiate a tumor (functional CSCs) is small compared to the total number of tumor cells, similarly to the normal stem cell compartment of most tissues (<1% of total cells). However, in other cancer types (such as melanoma) most of the cells are able to initiate a tumor in those assays that implies that a distinct CSC subpopulation does not exist. Additional studies suggest that distinct subpopulations are characterized by different grades of stemness, but technical hurdles remain in the interpretation of these assays. A more in-depth characterization of CSCs dynamics has revealed in many cases clues of phenotypic plasticity, suggesting a dynamic transition between a 'CSC state' and a 'non-CSC state'. Using a breast cancer inducible model of oncogenesis, in which putative CSCs can be sorted on the basis of expression of the CD44 cell surface marker ($CD44^{high}$ as compared to the non-CSCs that are $CD44^{low}$), Iliopoulos and colleagues showed that while — as expected — CSCs are able to give rise to non-CSCs, there is also, though to a lesser extent, reversion of non-CSCs to CSCs [2]. This observation has been confirmed also in primary human mammary breast cells [3]. According to these findings, CSC and non-CSC do not represent distinct and stable cell phenotypes, but functional subpopulations. Starting from the observation that each isolated subpopulation is able to generate the other subpopulation restoring the dynamic equilibrium among the two functional states, a Markov model of cell-state dynamics has been proposed [4]: this model assumes that interconversion rates depend on the cell current state, highlighting that intercellular signals are not necessarily required for the equilibrium. Whatever the exact nature of CSCs (functional state or

defined hierarchical component of the tumor), the transition from a CSC to a non-CSC cell/state must involve epigenetic changes and an understanding of those changes may be of critical relevance to identify distinct molecular landscapes of tumor cells.

## Epigenetic players recently involved in controlling the cancer stem cell phenotype

In melanomas, Roesch and colleagues described the presence of a slow cycling population characterized by high expression levels of JARID1B, a member of the family of Jumonji/ARID1 (JARID1) histone 3 lysine 4 (H3K4) demethylases [5]. JARID1B is not needed for melanoma initiation, but it is essential for tumor maintenance, both *in vitro* and *in vivo*, and for metastatic progression. Consistent with the findings described above, the phenotype of high expression of JARID1B is dynamic, and cells expressing low levels of Jarid1B may turn to high levels of expression. Dependence for high JARID1B expression in these stem-like melanoma cells support the idea that CSC and non-CSC cells/states are characterized by distinct and dynamic epigenetic landscapes, and depend on distinct epigenetic players. This is of capital importance, because it implies that we should explore the role of epigenetic players in the CSC/non CSC cell-states. JARID1B is overexpressed also in putative breast CSCs, suggesting a more general role of this enzyme in controlling cancer stem cell phenotypes. Interestingly, in $ER^+$ luminal breast cancer cells JARID1B colocalizes with the nuclear protein CTCF, which might modulate its histone demethylase activity and chromatin binding. In these cells, knockdown of JARID1B has growth-inhibitory effects [6].

JARID1B is also highly expressed in hematopoietic stem and progenitor cells (HSPCs) and in many acute myeloid leukemias (AMLs), but its role in controlling self-renewal of murine AMLs is not consistent with the role identified in human solid tumors [7,8]. Using a murine model of AML induced by the MLL-AF10 oncogene, and in which leukemic stem cells (LSCs) and non-LSCs are distinguished by the presence or absence of c-kit expression respectively [9], Wong and colleagues report that H3K4 dimethylation and trimethylation levels are significantly higher in LSCs (c-$kit^+$) than non-LSCs (c-$kit^-$) [10••]. Furthermore, JARID1B transcript levels in c-$kit^+$ cells are lower than c-$kit^-$ and ectopic overexpression of JARID1B markedly suppresses *in vitro* growth of MLL-leukemic cells, but does not impact the growth of non-MLL leukemias, suggesting a tumor suppressive role for JARID1B. It remains to be seen if this is a specific feature of MLL-driven AMLs, harboring an alteration of the H3K4 methyltransferase activity of MLL that leads to a different role of JARID1B. In contrast, Harris *et al.* report that KDM1A, a flavin adenine dinucleotide (FAD)-dependent demethylase with monomethyl and dimethyl H3K4 and H3K9 substrate specificity, is overexpressed in MLL-AF9 LSCs and drives their leukemogenic potential, making KDM1A a potential therapeutic target in MLL-driven leukemias [11]. The finding that two epigenetic factors with a similar enzymatic function can have opposing roles in the same disease (JARID1B and KDM1A in MLL-driven leukemias) is not unique (see the histone H3K27me3 demethylases JMJD3 and UTX in T-cell acute lymphoblastic leukemia [12]), and is strongly suggestive that specific (rather than global) changes in chromatin states (and the histone/DNA modifications underlying them) are responsible for the observed stem-cell phenotypes.

As reported in melanoma, a quiescent drug-tolerant subpopulation has been found in a non-small cell lung cancer cell line (PC9) [13]. This subpopulation expresses CD133 and CD24, previously described markers of CSCs [14,15], and is able to resume normal proliferation and to give rise to $CD133^-/CD24^-$ cells once drug treatment is stopped. Interestingly, these cells present high levels of JARID1A, another member of the Jumonji/ARID1 H3K4 demethylase family, and display reduced H3K4 trimethylation and dimethylation levels [13].

## Aberrant DNA-methylation of tumor cells: an old player with new roles in cancer stem cell functions

While aberrant roles of histone-modifying enzymes have been studied more recently, abnormalities in DNA methylation in cancer cells were observed since 1980s [16–19]: these abnormalities appear in the early phases of tumor transformation, but their role in defining tumor heterogeneity has not been studied extensively [20,21]. Starting from evidences of a cellular hierarchical organization in AMLs and that LSC gene expression signatures are predictive of clinical outcome in multiple cohorts of AML patients [22], Jung and colleagues defined the DNA methylation profile of LSCs (defined on their engrafting ability in a xenotransplantation assays) and their downstream non-engrafting blasts, identifying more than 3000 differentially methylated regions (DMRs), of which almost 92% are hypomethylated in LSCs [23••]. Integrating DNA methylation with gene expression profile, they identified a LSC epigenetic signature and showed that genes involved in leukemogenesis and proliferation of hematopoietic stem and progenitor cells (HSPCs), such as *REC8* and the *HOXA* cluster, are hypomethylated and overexpressed in LSCs; moreover, they demonstrated that human AMLs presenting the LSC-like epigenetic signature are associated with a worse clinical outcome compared to blast-like AMLs [23••], confirming the clinical prognostic potential of DNA-methylation analysis of AML samples [24]. Similarly, in the CSC signature of breast cancer cells, DMRs have lower methylation levels in CSCs than in non-CSCs, and tumors enriched in CSCs are associated with a decrease in relapse-free survival [25].

## Alterations of chromatin states: bivalent domains and phenotypic transitions of cancer stem cells

Abnormal activity of pathways that control normal stem cell self-renewal, embryonic development and differentiation, such as WNT, SHH and NOTCH, have been found in CSCs. As these pathways are active both in normal and cancer stem cells, the difference between a normal and a malignant cell must lie in the mechanisms (including epigenetic alterations) that regulate those pathways. Importantly, it has been shown that in normal (both embryonic and adult) stem cells, promoters of many differentiation-related genes are characterized by the simultaneous presence of both active (H3K4me3) and repressed (H3K27me3) histone marks, by low levels of DNA methylation and low transcription [26]. This chromatin state is termed 'bivalent', and during differentiation this state is resolved into active (H3K4me3) or inactive (H3K27me3) marks, leading to higher levels of active transcription or to gene silencing, respectively. During malignant transformation, the active state of differentiation-related promoter genes can be reverted to bivalent or even hypermethylated at CpG islands to achieve a permanent silenced state [27]; indeed, DNA-methylation constitutes a tighter mode of gene silencing than bivalency or H3K27me3 occupation. Moreover, promoters of developmental regulator genes, which are characterized by bivalent domains in embryonic or adult stem/progenitor cells, have the highest probability to be found DNA-hypermathylated in cancer [28•]. This is a relevant difference between normal and cancer (stem) cells: the inability to actively transcribe those genes prevents cancer cells from reaching a terminally differentiated state, leaving them in an intermediate and phenotypically plastic state (Figure 1). On the other hand, genes repressed in differentiated cells can revert to a bivalent state in cancer cells and this bivalency could underlie the phenotypic plasticity of cancer cell subpopulations, as reported by Chaffer *et al.* [29•]. In particular, the authors identified ZEB1 as a key mediator of spontaneous $CD44^{low}$ (non-CSCs) to $CD44^{high}$ (CSCs) conversion, acting through repression of the MIR200 family of microRNAs, and noted that the promoter of ZEB1 in $CD44^{high}$ cells display a chromatin methylation pattern associated with active transcription (H3K4me3 and H3K79me2 marks), while in $CD44^{low}$ cells it maintains a bivalent chromatin configuration and is transcribed weakly (H3K4me3 and H3K27me3 marks) [29•].

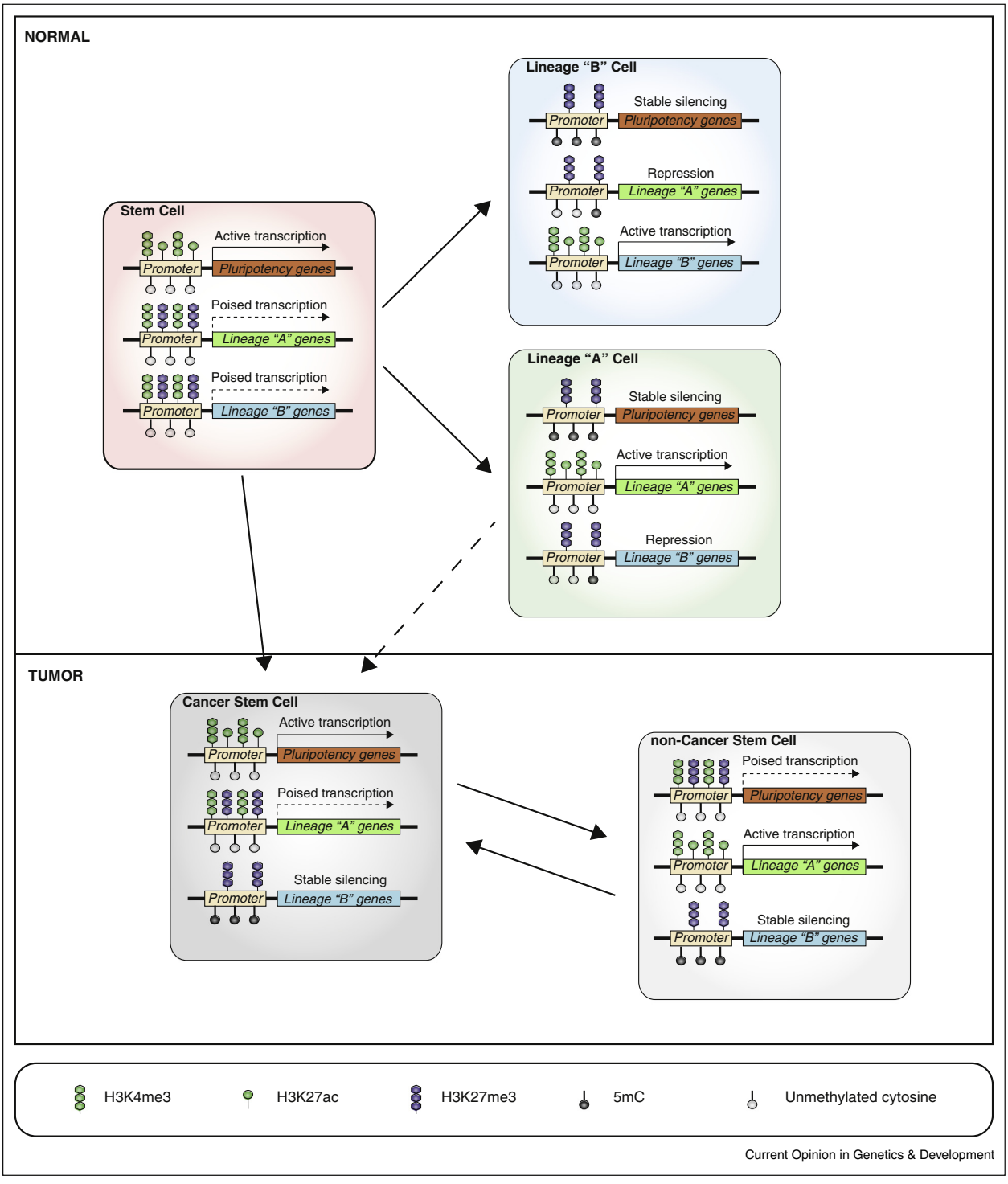
Rheinbay *et al.* examined the epigenetic landscape of Glioblastoma Multiforme (GBM) CSCs, and compared it with those of normal human astrocytes (NHAs) [30]. Interestingly, a major feature distinguishing CSCs in GBM from normal cells lies in a lack of the repressive H3K27me3 mark from promoters normally repressed or bivalent in NHAs: this results in deregulation of several transcription factor networks and in an altered WNT signaling, essential for CSC maintenance and tumor initiation. These findings suggest a critical role for Polycomb complexes (responsible for deposition of H3K27me3) and their regulation for the establishment and maintenance of the CSC phenotype [31].

In the context of hematopoietic malignancies, we performed a bioinformatic analysis of the available datasets of LSC-enriched c-$kit^+$ versus non-LSCs c-$kit^-$ AML cells from one study discussed above [10••]. We found a significant fraction of genes in c-$Kit^+$/LSC cells characterized by the bivalent mark ($N = 2261$), and noticed that the majority of them (>80%) undergo changes in their chromatin state upon transition to a non/LSC, c-$kit^-$ phenotype (Figure 2a). We divided therefore genes that present the bivalent mark in c-$Kit^+$ or in c-$Kit^-$ cells into four distinct classes, based on their epigenetic switching during the c-$Kit^+$/c-$Kit^-$ transition (Figure 2b–e). Class 1 (bivalent marks in LSC that lose H3K4me3 methylation and maintain the repressive H3K27me3 mark in non-LSC) is the most represented and, as expected, contains several genes related to pluripotency and strictly involved in cancer (e.g. *Wnt*, *Notch*, *Tgf-β*, *Id2*, *Zeb1*, *Axin1*, *Gli*). Examples of class 2 (bivalent in LSC, and maintaining the H3K4me3 active mark in non-LSC) and class 3 genes (presenting only H3K27me3 in LSC, and becoming bivalent in non-LSC), which show a more repressive chromatin state in LSC than in non-LSC, are: *Ip3r*, required for active proliferation [32], *Ptp* genes [33], genes involved in retinoic acid driven differentiation (*Pka* and *Rarα*), and genes controlling adipogenesis/metabolism (such as *Pparg*, *Bmal1*, *Clock*, *S6k1*) that potentially outline the metabolic circuitry responsible for differences in bioenergetics between CSCs and non-CSCs. Finally, class 4 (genes marked by H3K4me3 in LSC, found bivalent in non-LSC), comprehends genes involved in the CSC phenotype, such as *Mpo*, *Pi3k* [34], *Vegf*, important for self-renewal, survival and HSC niche maintenance [35] and *Smad* genes, which are component of *Tgf-β* signaling, hypothesized to be a cardinal regulator of HSC quiescence, maintaining a slow cycling state in HSCs [36]. These observations suggest a critical role for the molecular machineries responsible for the histone modifications involved in the bivalent chromatin state (methyltransferase — including MLL and Polycomb complexes — and respective demethylase activities) in imposing the CSC phenotype and governing the CSC to non-CSC transition, consistent with the available findings discussed above.

## New approaches to dissect tumor heterogeneity: single-cell epigenomics

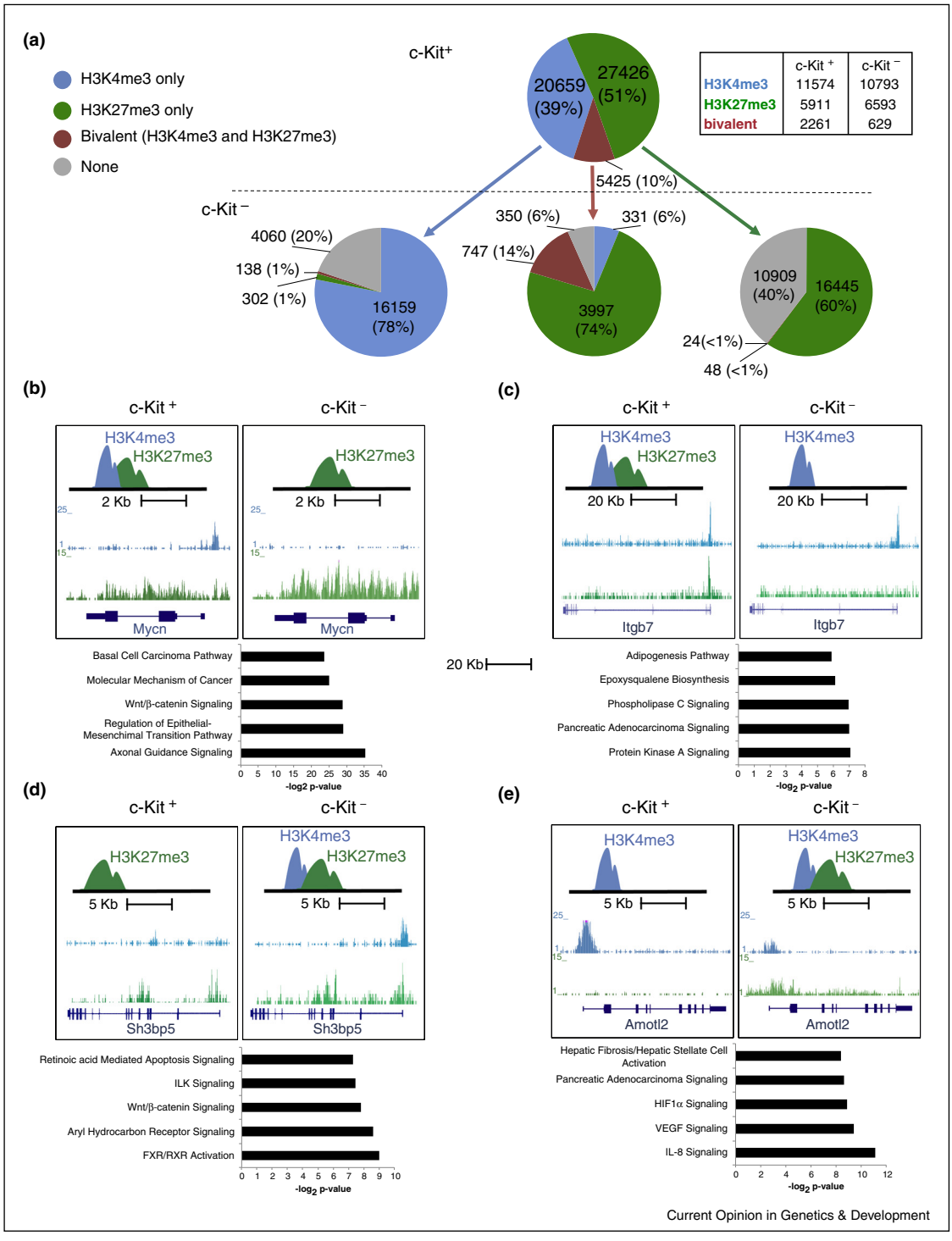
Epigenomic studies usually rely on the analysis of large (millions) to relatively small (thousands) cell fractions. Given the complexity of the stem cell phenotype in cancer, and the lack of definitive markers to isolate a 100% pure cancer stem cell fraction, single-cell resolution

**Figure 1**




Bivalent domains in normal and cancer cells. In normal stem cells, the presence of H3K4me3, H3K27ac and the absence of DNA methylation at pluripotency genes make them active, while differentiation-related gene promoters are usually characterized by bivalent chromatin that makes them poised. During normal differentiation, pluripotency genes are stable silenced by DNA methylation, while bivalent domains resolve into an active or a repressed chromatin state. Cancer cells show a more variegated pattern of chromatin changes in the transition from CSCs to non-CSCs (see Figure 2 for a specific case): they might maintain bivalent domains in differentiation-related genes (CSCs) and pluripotency genes (non-CSCs), or acquire irreversible silencing of differentiation-related genes (and tumor-suppressor genes) by DNA methylation. The presence of bivalent marks instead of stable repression may contribute significantly to the phenotypic plasticity of tumor cells.

**Figure 2**




Changes in the bivalent domain mark associated with the transition of CSCs to a non-CSC state in MLL-AF10 murine acute myeloid leukemia cells. **(a)** The pie charts show the number of genomic regions carrying the histone modifications associated with bivalency in c-Kit$^+$ (LSC-enriched) and c-Kit$^-$ cells. In the charts of c-Kit$^-$ cells, 'none' marks regions that lose both histone marks (histone H3 K4me3 and K27me3) in the transition from c-Kit$^+$ cells. The total number of genes associated with the genomic regions is shown in the table inset. **(b)**–**(e)** Genes characterized by bivalent domains are divided in four classes (composed of 1663, 173, 46 and 124 genes, respectively) as indicated and described in details in the main text. For each class, in the panels we show tracks of the histone marks for a representative gene, and the analysis of the corresponding IPA gene ontology.

could help to achieve a better definition of the epigenome associated with the cancer stem cell phenotype, and even lead to the identification of better markers.

Jaitin *et al.* developed an automated massively parallel RNA single-cell sequencing framework (MARS-Seq) to dissect in an unsupervised manner cellular heterogeneity of the hematopoietic system [37•], while single-cell expression profiling of five freshly resected and dissociated human glioblastomas proved cellular heterogeneity at the transcriptional level in primary tumors and, regardless of the dominant cellular subtype, highlighted the presence in each tumor of stem-like proneural cells [38•]. These and other studies [39] demonstrate the value of single-cell approaches and further suggest that single-cell epigenomics would incredibly increase our knowledge in cell-to-cell variation of heterogeneous cell populations. Indeed, recent studies suggest that the analysis of at least some aspects of the epigenome can be achieved at a single cell resolution. For DNA methylation, a single-cell genome-wide bisulfite sequencing method has been shown to be able to deconvolve mixed cell populations and to individuate rare cells in heterogeneous populations [40]. Using a different single-cell whole-genome bisulfite sequencing protocol (scWGBS), Farlik and colleagues demonstrated that is possible to infer cell-state dynamics through DNA methylation profiling in three *in vitro* models of cellular differentiation and pluripotency [41]. Nucleosome positioning, chromatin accessibility, transcription factor occupancy and gene expression are closely related [42]. Many methods to interrogate chromatin conformation have been developed [43]: among these methods, the assay for transposase-accessible chromatin (ATAC-Seq) allow simultaneous identification of regions of open chromatin, nucleosome-bound and nucleosome-free positions in regulatory elements, and fingerprint of DNA-binding proteins [44]. Recently, two different protocols for single-cell ATAC-Seq have been published [45,46].

Interestingly, recent studies allow the identification and potential isolation of primitive cancer initiating cells [47]. Analysis of these cells with the techniques outlined above is technically challenging, but offers fascinating perspectives to our understanding of the cancer stem cell phenotype.

## Conclusion and perspectives for cancer therapy

Though the cancer stem cell concept has been intensively explored in cancer research for the past 15 years, several critical aspects remain unresolved. It is not surprising therefore that a mechanistic analysis of epigenetic alterations underlying the cancer stem cell phenotype suffers from several limitations, and waits for further studies (we have highlighted the single-cell analyses as a new technological advance that may surely lead to advances in this field).

Another limitation is that, while epigenetic alterations defining differences among CSC and non-CSC tumor cell populations are being described, much less clear is how these differences are being achieved, and what is causing the dynamic equilibrium between the two states.

Nevertheless, we suggest that it is critical to continue to explore the epigenome of cancer stem cells, considering its potential relevance for therapy. As aforementioned, CSCs and non-CSCs have a different sensitivity to both cytotoxic and epigenetic drugs. Chronic myeloid leukemia (CML) is one of the best-described examples, being CML stem cells resistant to Imatinib treatment, in contrast to the dramatic sensitivity of bulk CML cells [48]. Acute promyelocytic leukemia (APL) represents another example, where APL stem cells are less sensitive than bulk leukemic cells to retinoic acid (RA) differentiation stimulus and only the combination of RA with other drugs (arsenic, chemotherapy) is able to eradicate the CSC compartment [49]. In the same disease, the class I HDAC inhibitor valproic acid (VPA) acts mainly on the bulk of tumor cells, showing only a mild effect on CSCs [50]. These examples outline the importance of targeting the right cellular subpopulations, including CSCs. A therapeutic strategy based only on CSCs eradication is unlikely to succeed: reducing the percentage of CSCs below a threshold could trigger a massive non-CSC to CSC phenotypic switch [51]. Therefore, a rational pharmacological approach could be a combination of debulking drugs (chemotherapy, cytotoxic drugs) with CSC targeted therapy, maybe directed to those epigenetic modifiers previously described (JARID1A, JARID1B, KDM1A). A better understanding of the epigenetic mechanisms controlling the stem cell functions of tumor cells will therefore allow the definition of additional therapeutic approaches: we surmise that epigenetic drugs will represent a critical weapon in the arsenal of drugs targeting the CSC phenotype.

## Acknowledgements

Work in SM's lab is supported by AIRC, FIRC, CNR Epigen Flagship Project, Ministry of Health and European Community (Blueprint, 4D Cell Fate, DRAMA).

## References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Nowell PC: **The clonal evolution of tumor cell populations**. *Science* 1976, **194**:23-28.

2. Iliopoulos D, Hirsch HA, Wang G, Struhl K: **Inducible formation of breast cancer stem cells and their dynamic equilibrium with non-stem cancer cells via IL6 secretion**. *Proc Natl Acad Sci U S A* 2011, **108**:1397-1402.

3. Chaffer CL, Brueckmann I, Scheel C, Kaestli AJ, Wiggins PA, Rodrigues LO, Brooks M, Reinhardt F, Su Y, Polyak K *et al.*: **Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state**. *Proc Natl Acad Sci U S A* 2011, **108**:7950-7955.

4. Gupta PB, Fillmore CM, Jiang G, Shapira SD, Tao K, Kuperwasser C, Lander ES: **Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells**. *Cell* 2011, **146**:633-644.

5. Roesch A, Fukunaga-Kalabis M, Schmidt EC, Zabierowski SE, Brafford PA, Vultur A, Basu D, Gimotty P, Vogt T, Herlyn M: **A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth**. *Cell* 2010, **141**:583-594.

6. Yamamoto S, Wu Z, Russnes HG, Takagi S, Peluffo G, Vaske C, Zhao X, Moen Vollan HK, Maruyama R, Ekram MB *et al.*: **JARID1B is a luminal lineage-driving oncogene in breast cancer**. *Cancer Cell* 2014, **25**:762-777.

7. Cellot S, Hope KJ, Chagraoui J, Sauvageau M, Deneault E, MacRae T, Mayotte N, Wilhelm BT, Landry JR, Ting SB *et al.*: **RNAi screen identifies Jarid1b as a major regulator of mouse HSC activity**. *Blood* 2013, **122**:1545-1555.

8. Stewart MH, Albert M, Sroczynska P, Cruickshank VA, Guo Y, Rossi DJ, Helin K, Enver T: **The histone demethylase Jarid1b is required for hematopoietic stem cell self-renewal in mice**. *Blood* 2015, **125**:2075-2078.

9. Somervaille TC, Cleary ML: **Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia**. *Cancer Cell* 2006, **10**:257-268.

10. •• Wong SH, Goode DL, Iwasaki M, Wei MC, Kuo HP, Zhu L, Schneidawind D, Duque-Afonso J, Weng Z, Cleary ML: **The H3K4-methyl epigenome regulates leukemia stem cell oncogenic potential**. *Cancer Cell* 2015, **28**:198-209.
In this study, the researchers define the epigenetic landscape of LSC and non-LSC cells and highlight the role of KDM5B in defining the LSC phenotype.

11. Harris WJ, Huang X, Lynch JT, Spencer GJ, Hitchin JR, Li Y, Ciceri F, Blaser JG, Greystoke BF, Jordan AM *et al.*: **The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells**. *Cancer Cell* 2012, **21**:473-487.

12. Ntziachristos P, Tsirigos A, Welstead GG, Trimarchi T, Bakogianni S, Xu L, Loizou E, Holmfeldt L, Strikoudis A, King B *et al.*: **Contrasting roles of histone 3 lysine 27 demethylases in acute lymphoblastic leukaemia**. *Nature* 2014, **514**:513-517.

13. Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, McDermott U, Azizian N, Zou L, Fischbach MA *et al.*: **A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations**. *Cell* 2010, **141**:69-80.

14. Hemmati HD, Nakano I, Lazareff JA, Masterman-Smith M, Geschwind DH, Bronner-Fraser M, Kornblum HI: **Cancerous stem cells can arise from pediatric brain tumors**. *Proc Natl Acad Sci U S A* 2003, **100**:15178-15183.

15. Vermeulen L, Todaro M, de Sousa Mello F, Sprick MR, Kemper K, Perez Alea M, Richel DJ, Stassi G, Medema JP: **Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity**. *Proc Natl Acad Sci U S A* 2008, **105**:13427-13432.

16. Feinberg AP, Vogelstein B: **Hypomethylation distinguishes genes of some human cancers from their normal counterparts**. *Nature* 1983, **301**:89-92.

17. Herman JG, Latif F, Weng Y, Lerman MI, Zbar B, Liu S, Samid D, Duan DS, Gnarra JR, Linehan WM *et al.*: **Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma**. *Proc Natl Acad Sci U S A* 1994, **91**:9700-9704.

18. Esteller M, Herman JG: **Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours**. *J Pathol* 2002, **196**:1-7.

19. Kulis M, Esteller M: **DNA methylation and cancer**. *Adv Genet* 2010, **70**:27-56.

20. Baylin SB, Jones PA: **A decade of exploring the cancer epigenome — biological and translational implications**. *Nat Rev Cancer* 2011, **11**:726-734.

21. Esteller M: **Epigenetic lesions causing genetic lesions in human cancer: promoter hypermethylation of DNA repair genes**. *Eur J Cancer* 2000, **36**:2294-2300.

22. Eppert K, Takenaka K, Lechman ER, Waldron L, Nilsson B, van Galen P, Metzeler KH, Poeppl A, Ling V, Beyene J *et al.*: **Stem cell gene expression programs influence clinical outcome in human leukemia**. *Nat Med* 2011, **17**:1086-1093.

23. •• Jung N, Dai B, Gentles AJ, Majeti R, Feinberg AP: **An LSC epigenetic signature is largely mutation independent and implicates the HOXA cluster in AML pathogenesis**. *Nat Commun* 2015, **6**:8489.
Here the authors show different DNA methylation signatures between AML LSCs and their blast progeny. Moreover, they use these signatures to identify distinct subgroups of AML, likely reflecting the cell of origin.

24. Bartholdy B, Christopeit M, Will B, Mo Y, Barreyro L, Yu Y, Bhagat TD, Okoye-Okafor UC, Todorova TI, Greally JM *et al.*: **HSC commitment-associated epigenetic signature is prognostic in acute myeloid leukemia**. *J Clin Invest* 2014, **124**:1158-1167.

25. El Helou R, Wicinski J, Guille A, Adelaide J, Finetti P, Bertucci F, Chaffanet M, Birnbaum D, Charafe-Jauffret E, Ginestier C: **Brief reports: a distinct DNA methylation signature defines breast cancer stem cells and predicts cancer outcome**. *Stem Cells* 2014, **32**:3031-3036.

26. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K *et al.*: **A bivalent chromatin structure marks key developmental genes in embryonic stem cells**. *Cell* 2006, **125**:315-326.

27. Easwaran H, Tsai HC, Baylin SB: **Cancer epigenetics: tumor heterogeneity, plasticity of stem-like states, and drug resistance**. *Mol Cell* 2014, **54**:716-727.

28. • Easwaran H, Johnstone SE, Van Neste L, Ohm J, Mosbruger T, Wang Q, Aryee MJ, Joyce P, Ahuja N, Weisenberger D *et al.*: **A DNA hypermethylation module for the stem/progenitor cell signature of cancer**. *Genome Res* 2012, **22**:837-849.
The investigators show that in tumor cells DNA hypermethylation preferentially targets developmental regulators genes that are PcG target and in normal stem cells are characterized by bivalent chromatin; this may contribute to the stem-like state of cancer cells.

29. • Chaffer CL, Marjanovic ND, Lee T, Bell G, Kleer CG, Reinhardt F, D'Alessio AC, Young RA, Weinberg RA: **Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity**. *Cell* 2013, **154**:61-74.
Here the researchers demonstrate that non-CSCs of human basal breast cancer can switch to CSC state and this switching is dependent on ZEB1, whose promoter maintain a bivalent chromatin conformation in non-CSCs and an active conformation in CSCs.

30. Rheinbay E, Suva ML, Gillespie SM, Wakimoto H, Patel AP, Shahid M, Oksuz O, Rabkin SD, Martuza RL, Rivera MN *et al.*: **An aberrant transcription factor network essential for Wnt signaling and stem cell maintenance in glioblastoma**. *Cell Rep* 2013, **3**:1567-1579.

31. Voigt P, Tee WW, Reinberg D: **A double take on bivalent promoters**. *Genes Dev* 2013, **27**:1318-1338.

32. Ouyang K, Leandro Gomez-Amaro R, Stachura DL, Tang H, Peng X, Fang X, Traver D, Evans SM, Chen J: **Loss of IP3R-dependent $Ca^{2+}$ signalling in thymocytes leads to aberrant development and acute lymphoblastic leukemia**. *Nat Commun* 2014, **5**:4814.

33. Arora D, Kothe S, van den Eijnden M, Hooft van Huijsduijnen R, Heidel F, Fischer T, Scholl S, Tolle B, Bohmer SA, Lennartsson J *et al.*: **Expression of protein-tyrosine phosphatases in acute myeloid leukemia cells: FLT3 ITD sustains high levels of DUSP6 expression**. *Cell Commun Signal* 2012, **10**:19.

34. Fransecky L, Mochmann LH, Baldus CD: **Outlook on PI3K/AKT/mTOR inhibition in acute leukemia**. *Mol Cell Ther* 2015, **3**:2.

35. Kampen KR, Ter Elst A, de Bont ES: **Vascular endothelial growth factor signaling in acute myeloid leukemia**. *Cell Mol Life Sci* 2013, **70**:1307-1317.

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

37. • Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretsky I, Mildner A, Cohen N, Jung S, Tanay A *et al.*: **Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types**. *Science* 2014, **343**:776-779.
This report demonstrates for the first time single-cell RNA-seq as an effective tool for comprehensive cellular decomposition of complex tissues.

38. • Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, Cahill DP, Nahed BV, Curry WT, Martuza RL *et al.*: **Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma**. *Science* 2014, **344**:1396-1401.
The investigators profile primary glioblastoma samples using single-cell RNA-seq and found stem-like proneural cells (putative CSCs) in each sample analyzed.

39. Sandberg R: **Entering the era of single-cell transcriptomics in biology and medicine**. *Nat Methods* 2014, **11**:22-24.

40. Smallwood SA, Lee HJ, Angermueller C, Krueger F, Saadeh H, Peat J, Andrews SR, Stegle O, Reik W, Kelsey G: **Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity**. *Nat Methods* 2014, **11**:817-820.

41. Farlik M, Sheffield NC, Nuzzo A, Datlinger P, Schonegger A, Klughammer J, Bock C: **Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics**. *Cell Rep* 2015, **10**:1386-1397.

42. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B *et al.*: **The accessible chromatin landscape of the human genome**. *Nature* 2012, **489**:75-82.

43. Tsompana M, Buck MJ: **Chromatin accessibility: a window into the genome**. *Epigenet Chromatin* 2014, **7**:33.

44. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ: **Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position**. *Nat Methods* 2013, **10**:1213-1218.

45. Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, Steemers FJ, Trapnell C, Shendure J: **Epigenetics. Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing**. *Science* 2015, **348**:910-914.

46. Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, Chang HY, Greenleaf WJ: **Single-cell chromatin accessibility reveals principles of regulatory variation**. *Nature* 2015, **523**:486-490.

47. Devkota S: **MICROBIOME. Prescription drugs obscure microbiome analyses**. *Science* 2016, **351**:452-453.

48. Corbin AS, Agarwal A, Loriaux M, Cortes J, Deininger MW, Druker BJ: **Human chronic myeloid leukemia stem cells are insensitive to imatinib despite inhibition of BCR-ABL activity**. *J Clin Invest* 2011, **121**:396-409.

49. Ablain J, de The H: **Revisiting the differentiation paradigm in acute promyelocytic leukemia**. *Blood* 2011, **117**:5795-5802.

50. Leiva M, Moretti S, Soilihi H, Pallavicini I, Peres L, Mercurio C, Dal Zuffo R, Minucci S, de The H: **Valproic acid induces differentiation and transient tumor regression, but spares leukemia-initiating activity in mouse models of APL**. *Leukemia* 2012, **26**:1630-1637.

51. Sellerio AL, Ciusani E, Ben-Moshe NB, Coco S, Piccinini A, Myers CR, Sethna JP, Giampietro C, Zapperi S, La Porta CA: **Overshoot during phenotypic switching of cancer cell populations**. *Sci Rep* 2015, **5**:15464.