UNIVERSITÀ DEGLI STUDI DI UDINE
hic sunt futura
www.uniud.it

PhD Course in Biomedical Sciences and Biotechnology

XXXIV cycle

BIOINFORMATICS ANALYSES TO ELUCIDATE THE CLASS IIA HDACs-MEF2 ONCOGENIC POTENTIAL IN COLORECTAL CANCER AND LEIOMYOSARCOMA

PhD candidate:
Massimo Faggiani

Supervisor:
Prof. Claudio Brancolini
Co-supervisor:
Dr. Emiliano Dalla PhD

2022

Abstract

During my doctoral studies I had the opportunity to participate in several bioinformatics projects under the supervision of professor Brancolini, to address pressing scientific questions regarding the involvement of the class IIa histone deacetylases (HDACs) epigenetic regulators, which exert a repressive function on transcription, in cancer. Firstly, my efforts were dedicated to exploring colorectal cancer (CRC) clinical samples, in addition to profiling the gene expression of the tumor microenvironment's cell populations and the hypoxic milieu, two key factors influencing CRC patients' prognosis, to estimate their oncogenic contribution. Secondly, another analysis was centered on investigating the impact of HDAC4 knockout on the transcriptome of leiomyosarcoma (LMS) cells, through the quantitative assessment of the H3K27ac histone mark on several putative class IIa HDACs and MEF2 target genes, particularly those displaying an unconventional drop in the acetylation signal after HDAC4 depletion. Furthermore, the effect of the HDAC inhibitor NKL54, a PAOA (pimeloylanilide o-aminoanilide) derivative, in leiomyosarcoma cells was also examined by ChIP-seq data analysis, following the promising anti-proliferative action observed *in vitro*, with the aim of understanding its effect on the genomic occupancy of class IIa HDACs, MEF2 and on the changes in H3K27ac distribution. Overall, these results helped expand the horizon about the oncogenic potential of the class IIa HDACs-MEF2 axis in two different tumoral contexts, CRC and LMS, towards a better understanding of epigenetic-driven carcinogenesis.

Summary

Introduction

In the last decade, our research group has dedicated significant efforts in studying and comprehending the contribution to oncogenesis of the class IIa HDACs epigenetic regulators, describing an increasingly complex framework. In particular, after the first demonstration of the oncogenic potential of a hyper-mutated form of HDAC4 in fibroblasts and mice or and the pro-proliferative action of HDAC7 in the mammary gland, the focus shifted to those tumor types where a specific HDAC-MEF2 signature could justify further study and dedication for the clinical benefit. These tumors emerged to be the soft tissue sarcomas, a heterogeneous group of tumors which includes leiomyosarcoma (LMS), where the MEF2 signature appears suppressed and can correlate either with the PTEN oncosuppressor or with HDAC4, showcasing the MEF2 transcription factors (TFs) as a converging hub for HDAC4-driven oncogenic transformation (Di Giorgio et al., 2013).

To extend the horizon to other tumoral contexts, we decided to investigate if class IIa HDACs could exert a role in one of the increasingly prevalent cancers associated with an important clinical-societal burden, colorectal cancer (CRC), in the context of the COLONACT project, aimed at culturing patient-derived organoids for the optimization of treatment and an effective personalized medicine, also taking into account the tumor microenvironment (TME) conditions.

1.1 THE MYOCYTE ENHANCER FACTOR 2 (MEF2) FAMILY

The MEF2 TFs were identified the first time in differentiating myoblasts as being able to bind the muscle-specific creatine kinase enhancer (Gossett et al., 1989). Subsequent studies have delineated their crucial role during early myogenesis for the diverse muscle types (smooth and striated), although subordinated to the 4 basic helix–loop–helix (bHLH) myogenic master regulators, particularly myogenin (Cserjesi and Olson, 1991); however, they appear to be expressed in a variety of other tissues, including the developing nervous system. Currently, the MEF2 family has been involved in a multitude of heterogeneous processes, for example muscle differentiation (Taylor and Hughes, 2017), promotion of neuronal cell survival (Yin et al., 2012), response to stress stimuli and growth factors (Di Giorgio et al., 2018), inflammation (Tóth et al., 2018) and, finally, cancer (Potthoff and Olson, 2007; Di Giorgio et al., 2017; Chen et al., 2017).

1.1.1 ORIGIN AND GENOMIC ORGANIZATION OF THE MEF2 TRANSCRIPTION FACTORS

MEF2 can be found in simple organisms such as the yeast Saccharomyces cerevisiae and Drosophila melanogaster, which possess only a copy of the gene, while in Vertebrates several gene duplication events gave rise to 4 MEF2 paralogues (MEF2A, MEF2B, MEF2C and MEF2D), among which MEF2B is the most divergent due to a reduced selective pressure (Hobson et al., 1995, Wu et al., 2011). The coding region appears well conserved, with 9 common exons and 2 mutually exclusive ones (the ubiquitous α-1 and the muscle-specific α-2), whereas greater variability exists in the 5'-UTR region, comprising 7 non-coding exons along 60 kb (Ramachandran et al., 2008). Functionally, MEF2 belongs to the MADS-box family (MCM1 agamous deficiens SRF), and presents the homonym conserved A/T-rich DNA binding motif (YTA(A/T)4TAR), known as the MEF2 response element (MRE), localized at the N-terminus region in proximity to the MEF2 domain, responsible for the interaction with other cofactors; these two domains constitute an unstructured N-terminal tail formed by 3 α-helixes (H1-3) and 3 β-sheets, where H1 and H3 are involved in DNA binding and H2 seems concerned with mutually exclusive partner interactions (Wu et al, 2010). On the other side, the C-terminus appears poorly conserved among the 4 MEF2 members and presents 2 transcription activation domains (TADs) and the nuclear localization signal (NLS); besides, this region is subjected to alternative splicing events (Potthoff and Olson, 2007).

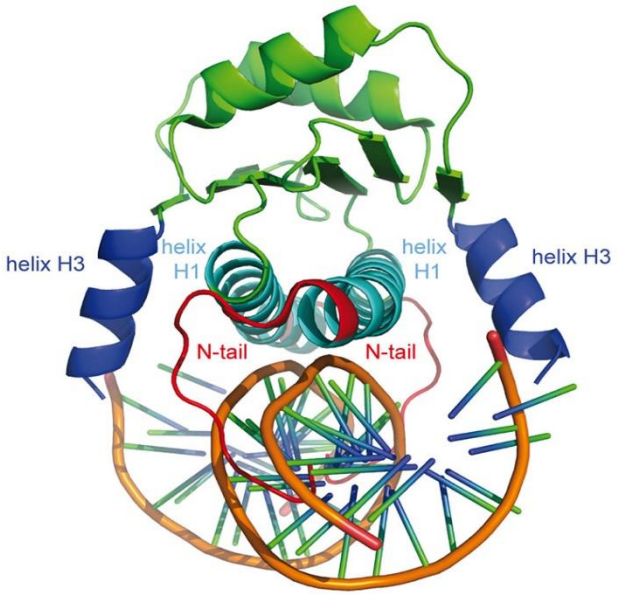


Fig.1.1.1 Structure of MEF2A MADS-box/MEF2 domains bound to DNA (PDB 3KOV) (modified from Di Giorgio et al., 2018).

1.1.2 MECHANISMS OF MEF2 REGULATION

Several actors contribute to the control of the MEF2 function at multiple levels, which can be summarized as follows:

1. binding of direct/indirect transcriptional corepressors, primarily class IIa HDACs (Miska et al., 1999), but also HDAC3, Cabin1, Smad3, or coactivators like p300 and Ash2L (Rampalli et al., 2007);
2. post-transcriptional modifications (PTMs), such as phosphorylation and SUMOylation which lead to positive/negative and negative effects on transcription, respectively (Gregoire et al., 2005), as well as acetylation and methylation with positive and negative effects, respectively (Ma et al., 2005; Choi et al., 2014);
3. control of the MEF2 transcript stability by means of miRNAs (Townley-Tilson et al., 2010);
4. regulation of the MEF2 protein turnover via different mechanisms: caspase and protease cleavage (Tang et al., 2005), the ubiquitin-proteasome pathway (Butts et al., 2005), autophagy (Yang et al., 2009).

In particular, class IIa HDACs represent the most prominent partners of MEF2 TFs, through different types of actions exemplified by the recruitment of the repressive complex N-CoR/SMRT/HDAC3 (Grégoire et al., 2007, Clocchiatti et al., 2013). The very first evidence of class IIa HDAC-MEF2 close relationship emerged with the discovery of MITR, a splicing variant of HDAC9 able to bind the MEF2 domain without interfering with the binding to DNA (Sparrow et al., 1999). Soon after, HDAC4 yeast's HDA1 homolog was characterized as a new repressor of the MEF2A activity (Miska et al., 1999), an association under the influence of calmodulin/CAMK (calmodulin-dependent protein kinase) and which appears important in triggering cardiomyocyte hypertrophy (Youn et al., 2000; Lu et al., 2000). Among the PTMs, phosphorylation at multiple sites plays a central role in the regulation of MEF2, usually resulting in an increased MEF2 activity, and a number of protein kinases are involved: casein kinase II (CKII), that acts on a conserved amino acid (Ser59 in MEF2C) (Molkentin et al., 1996); ERK5, belonging to the mitogen-activated protein kinases (MAPK) family, activates MEF2C after the addition of fetal serum in cultured cells (Kato et al., 1997). Moreover, the ERK5-MEF2 axis seems crucial for neuronal survival in response to the brain-derived neurotrophic factor (BDNF) during early stages of the central nervous system (CNS) development (Liu et al., 2003), and also in the inhibition of TRAIL-induced apoptosis in $Her2^{+}$ mammary cancer cells (Borges et al., 2007). Another important MAPK is p38, a potent MEF2A

activator which is involved in inflammation and infection responses (Han et al., 1997, Suzuki et al., 2004), and also in the re-establishment of the differentiation program in rhabdomyosarcoma (RMS) cells mediated by the activation of MEF2 (Puri et al., 2000). Additionally, GSK3β (glycogen synthase kinase 3β) regulates MEF2 activity, too, although indirectly through a crosstalk with p38 both in the skeletal and cardiac muscle, acting as a suppressor of myogenesis and cardiomyocyte hypertrophy (Dionyssiou et al., 2013). However, in some cases phosphorylation is associated with a suppressive function, as it is the case for CDK5 (cyclin-dependent kinase 5), where it precedes the caspase cleavage of MEF2A/D in primary neurons after neurotoxic damage (Tang et al., 2005), while the resulting N-terminal fragments bring to NMDA-induced neuronal apoptosis (Okamoto et al., 2002). SUMOylation appears generally associated with the negative regulation of MEF2, an effect reinforced by HDAC4 (although this is inhibited by the sumoylation of HDAC4 itself) and reversed by SENP3 and ERK5 (Gregoire and Yang, 2005). Furthermore, also acetylation affects MEF2 activity, as trichostatin A and nicotinamide upregulate this modification in MEF2D, and HDAC3 can catalyze the deacetylation by interacting with MADS box, unlike class IIa HDACs, thus inhibiting myogenesis *in vitro* and *in vivo* (Grégoire et al., 2007). Regarding the regulation by alternative splicing, MEF2C and MEF2D undergo tissue-specific splicing events, presenting an ubiquitously expressed isoform (α1) and a muscle-specific one (α2); a ChIP-seq experiment (Sebastian et al., 2013) revealed that these MEF2D isoforms bind overlapping sets of genes but only the latter is capable of activating late muscle differentiation gene expression, and this difference may be attributed to the escape from protein kinase A (PKA) inhibitory phosphorylation due to exon switching, thus abrogating the PKA binding site in MEF2D-α2 and aiding the recruitment of Ash2L for the transactivation of muscle genes. Similarly, MEF2C-α1 displays no myogenic activity while MEF2C-α2 is required for muscle differentiation; moreover, the α2/α1 ratio is down-regulated in rhabdomyosarcoma (RMS) and the over-expression of MEF2C-α2 mediated by SRPK3 promotes the differentiation of RMS cells inhibiting proliferation (Zhang et al., 2015). MEF2D appears also involved in cerebellar neurons' survival in response to DNA damage, being phosphorylated and activated by ATM (ataxia telangiectasia mutated); in fact, the MEF2D-silenced neurons are more susceptible to cell death induced by DNA disruption (Chan et al., 2014). It has been also demonstrated that in some circumstances the MEF2 activity is post-transcriptionally regulated by miRNAs: for example, in the heart miR-1 is able to block MEF2A and calmodulin, consequently inhibiting the hypertrophic process (Ikeda et al., 2009); by contrast, in skeletal muscle miR-1 promotes

myogenesis through the targeting of HDAC4, a well-acknowledged blocker of muscle-related gene expression (Chen et al., 2005).

With respect to the MEF2 protein stability, at least in myocytes and neurons these transcription factors are highly expressed and have extended life, due to the fully differentiated state of these cell types (Yu, 1996). In rat cerebellar granule neurons the induction of apoptosis causes the phosphorylation of MEF2A/D (but not MEF2B/C), diminished transcriptional activity and caspase-dependent cleavage, leading to the loss of their pro-survival role (Li et al., 2001). An analogous phenomenon happens also through the caspase-dependent cleavage of HDAC4, whose amino-terminal resulting fragment acts as a repressor of MEF2C independently from the HDAC domain and it induces cell death because of the release of mitochondrial cytochrome c (Paroni et al., 2004). Autophagy, the cellular mechanism by which unnecessary cellular components are actively eliminated, seems to be involved at least in the neuron in the turnover of MEF2D through the chaperone protein Hsc70, while the inhibition of this process causes the accumulation of MEF2D in the cytoplasm, pointing to a potential link with Parkinson's disease (Yang et al., 2009).

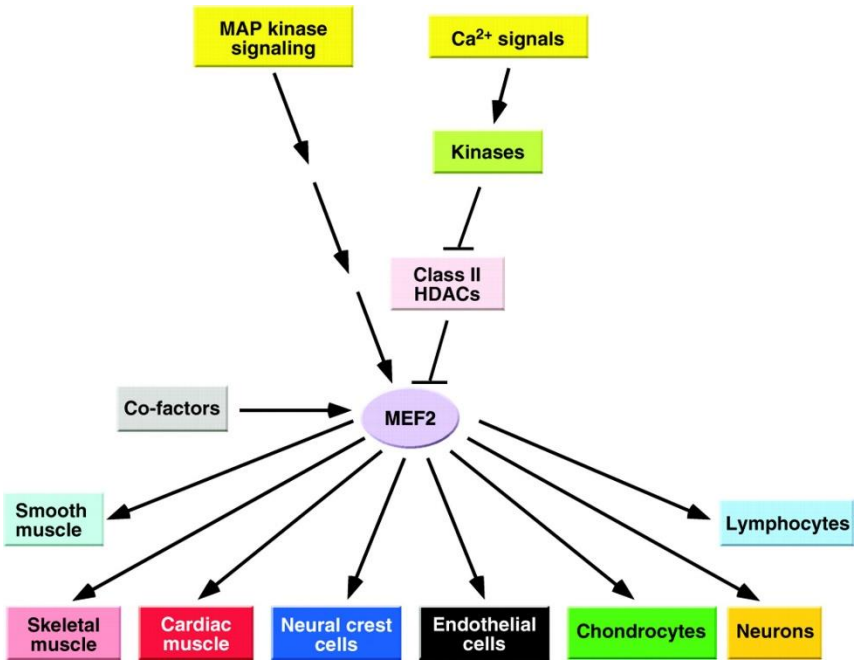


Fig.1.1.2 Signaling and differentiation pathways converging on MEF2 TFs in several cell types (Potthoff and Olson, 2007).

1.1.3 FUNCTIONS OF MEF2 FROM EMBRYOGENESIS TO THE ADULT

As anticipated, the MEF2 TFs represent important contributors in the myogenesis process, as firstly reported in Drosophila melanogaster, where Mef2 is induced via enhancer by the bHLH transcription factor Twist in the early mesoderm, this event corresponding to the initiation of the myoblast fusion (Baylies and Bate, 1996; Cripps et al., 1998). Further genome-wide studies have underlined the central role of MEF2 in all stages of muscle development, from flies to humans, for instance in the neuromuscular junction and ionic transport (Sandmann et al., 2006). In vertebrates, the 4 bHLH myogenic factors (MyoD, Myf5, Myogenin and MRF4) have been elevated to master regulators of skeletal muscle development, as they are necessary and sufficient to induce terminal myogenic differentiation also if expressed ectopically (Olson, 1990). In particular, MyoD and Myf5 appear firstly in proliferating myoblasts, while Myogenin and MRF4 are expressed only after the exit from the cell cycle (Olson and Klein, 1994; Rudnicki and Jaenisch, 1995). The smooth and the heart muscle types share several genes with the skeletal muscle but in these contexts the bHLH regulators are absent, suggesting the presence of an additional actor, MEF2, which alone is incapable of fully determining the muscle phenotype (Cserjesi and Olson, 1991; Yu et al., 1992). Recently, the MEF2 activity emerged being regulated by the cAMP-dependent signaling in myocyte-like cells, preventing the pathological activation by hypo-phosphorylated HDAC5 (He et al., 2020). Besides, the MEF2 activity is enhanced in cardiac hypertrophy and additionally seems to be involved in heart failure (Tobin et al., 2017). The specific functions of the different MEF2 members can be evaluated in mice, where animals deficient for MEF2A or MEF2C result in embryonic and soon-after-birth lethality, respectively, due to severe cardiac abnormalities, whereas MEF2D-null mice appear unaffected (Lin et al., 1997; Naya et al., 2002; Arnold et al., 2007). However, since the pattern of expression of MEF2 members is overlapping, as well as their function, researchers have been conducting studies with conditional mutants, hyper-active forms (VP16) and super-repressive ones (ENGRAILED) in order to delineate more accurately the singular roles (Karamboulas et al., 2006). The MEF2-HDAC7 interaction is important in maintaining vascular integrity by targeting metalloproteinase-10 (MMP10); indeed, HDAC7-null mice are embryonic lethal because of vascular collapse and also MEF2 appears to be essential in vascular remodeling through the ERK5 mediation (Chang et al., 2006; Hayashi et al., 2004). Neurons, second to the muscle cells, present the highest MEF2 levels, particularly in the cerebellum, the cortex, and the hippocampus. During development, the MEF2 expression pattern follows the neuronal differentiation gradient, peaking at fully differentiated cells (Lyons et al., 1995).

Furthermore, MEF2 TFs play a role in other developmental processes, i.e. in the bone and in particular in the craniofacial formation, as well as in the thymocytes maturation. For example, MEF2C deletion in the neural crest leads to severe craniofacial defects and consequently choke-related neonatal lethality in mice (Verzi et al., 2007). MEF2C is also involved in chondrocyte hypertrophy preceding endochondral ossification (Arnold et al., 2007), while MEF2D in cooperation with HDAC7 negatively regulates the survival of double-positive $CD4^+$ $CD8^+$ thymocytes through the pro-apoptotic factor Nur77, following T cell receptor (TCR) activation (Dequiedt et al., 2003). It can be argued that in adult organisms MEF2 mostly mediates adaptive and remodeling responses, such as the skeletal and cardiac muscular hypertrophy (Kolodziejczyk et al., 1999; Potthoff e Olson, 2007). Additionally, a SUMOylation-induced repressive form of MEF2A seems to participate in the synapse development and plasticity (Shalizi et al., 2006), reinforcing the idea of an anti-apoptotic function of MEF2 in the neuron (Okamoto et al., 2000), where in the presence of high levels of activity and soaring intracellular calcium levels, the CAMK-dependent phosphorylation of HDACs leads to the MEF2-HDAC dissociation, turning MEF2 into transcriptional activators (Shalizi and Bonni, 2005). MEF2 TFs also play a role in neuronal migration and differentiation, axon guidance, dendrite formation and remodeling (Ma and Telese, 2015; Latchney et al., 2015). MEF2 activity is not limited to neurodevelopmental processes, though, as data are emerging for a potential role of MEF2C in different mental disorders (Assali et al., 2019).

1.1.4 EMERGING EVIDENCE OF A MEF2 ROLE IN CANCER

The function played by MEF2 in cell fate and adaptive programs is well-acknowledged. However, in relation to cancer the evidence is currently scarcer but nonetheless increasingly established (Pon and Marra, 2016), with several MEF2 genetic alterations occurring in various cancers (Di Giorgio et al., 2018). This correlation was firstly suggested by the discovery, after serum stimulation, of a MEF2D-mediated activation of c-Jun involved in the cell cycle progression (Han and Prywes, 1995); furthermore, also receptors like the G protein-coupled receptors (GPCRs) and the epidermal growth factor receptor (EGFR) seem to converge on MEF2 for the activation of c-Jun (Coso et al., 1997; Clarke et al, 1998). Putative positive mediators for the MEF2 oncogenic potential are calmodulin and some MAPK members, whose pathways are subjected to reciprocal crosstalk and reinforcement; for instance, calmodulin activates calmodulin-dependent protein kinase IV (CAMKIV) causing the nuclear extrusion of HDAC4, permitting the interaction between MEF2 and the coactivator p300 (Lu et al., 2000).

On the other side, ERK5 and p38 directly activate MEF2 through the phosphorylation of the TAD (Passier et al., 2000). Overall, these pathways converging on MEF2 effectively contribute to the hypertrophy of cardiomyocytes. Indeed, in satellite cells the depletion of MEF2A/C/D does not impair proliferation but impacts negatively on the regenerative potential after muscle damage derived from disease or injury, in which HDAC4 is also involved (Liu et al., 2014; Choi et al., 2014). In the liver, however, the MEF2 expression associates positively with the proliferation of stellate cells (HSC) *in vitro* and fibrosis *in vivo*, inducing at the same time α-smooth muscle actin (α-SMA) and collagen I, through the mediation of p38 (Wang et al., 2004); in the same context, the inhibition of HDAC4 effectively reduces HSC activation markers and blocks proliferation (Mannaerts et al., 2013). Moreover, the MEF2D level in hepatic carcinoma is higher than in the corresponding normal tissue, it represents a prognostic marker and, finally, also seems involved in epithelial to mesenchymal transition (EMT) through TGF-β (Ma et al., 2014; Yu et al., 2014). Moving forward, in RMS cells MEF2D appears repressed and its overexpression promotes the differentiation by up-regulating p21 (Zhang et al., 2013). In concert with class IIa HDACs, MEF2 TFs form a complex which is often deregulated in soft tissue sarcomas (STS), and in a subset of these tumors the MEF2 signature is suppressed, pointing to an oncogenic function of HDAC4 (Di Giorgio et al., 2013). Leiomyosarcomas (LMS) constitute about 10% of STS cases, with few therapeutic options and poor prognosis; in LMS, the MEF2 expression inversely correlates with overall survival and the coexistence of activator and repressor complexes with HDAC9 can influence tumor aggressiveness, driving either pro-oncogenic or onco-suppressor activities *in vitro* (Di Giorgio et al., 2017, 2018). More recently, MEF2D emerged as the privileged partner of class IIa HDACs in LMS, leading to a vicious circle that is able to sustain cell proliferation through the formation of repressive complexes with HDAC4 and HDAC9, preferentially at distal regions, acting as super-enhancers (SE) (for a description of super-enhancers see paragraph 1.3) (Di Giorgio et al., 2020). These data support the view that in LMS the HDAC IIa-MEF2 axis represents a promising target, even though the MEF2 TFs themselves are undruggable, pointing the attention on the research of effective and selective class IIa HDACs inhibitors.

1.2 THE CLASS IIA HDACs EPIGENETIC REGULATORS

Histone deacetylases (HDACs) constitute an ancient protein superfamily found also in prokaryotes and, in contraposition to histone acetylases (HATs), they exert an important regulatory function on gene expression, bringing to a reset of the chromatin to a condensed and

inaccessible state associated with transcriptional repression (Bannister and Kouzarides, 2011). However, members of class IIa HDACs (HDAC4, HDAC5, HDAC7, HDAC9) represent to some extent an exception, characterized by an almost negligible action on histone tails due to a single-point mutation occurred in vertebrates (Lahm et al., 2007). In contrast to other HDAC sub-families, class IIa HDACs possess a bipartite structure with an extended N-terminal domain (residues 450-600) where reside several phosphorylation target sites for the nuclear-cytoplasmic shuttling mediated by 14-3-3 chaperons and a polyQ domain responsible for numerous interactions, like the one with with Runx2 (Clocchiatti et al, 2011); the C-terminus, instead, contains a conserved deacetylase domain and the nuclear export sequence (NES) for cytoplasmic accumulation. Partners of this class include chromatin remodeling factors (CtBP, HP1 and the SUV39H1 methyltransferase), development-associated factors (Runx 2/3, GATA 1/2, FOXP3, Nkx2-5) and hypoxia-inducible factor-1 α (HIF-1α) which is stabilized in hypoxic conditions by HDAC7 (Kato et al., 2004), but the list is on the rise. The most studied partners, however, are represented by the MEF2 TFs, as reported more than 20 years ago with the first characterization of HDAC4 and its repressive impact on the MEF2A signature (Miska et al., 1999). Turning back to the nuclear extrusion, this event constitutes the prevalent mechanism of class IIa HDAC inactivation and is influenced by several kinases, for example CAMK I-IV, PKD, MARK2-3, SIK1 and Dyrk1B, which make this HDAC sub-class responsive to a variety of signaling pathways, from the response to mechanical stress in the heart to T cell apoptosis (Martin et al, 2007). On the other hand, phosphorylases restore the HDAC IIa activity taking them back to the nucleus: in this regard, PP2A is the best characterized (Paroni et al, 2007).

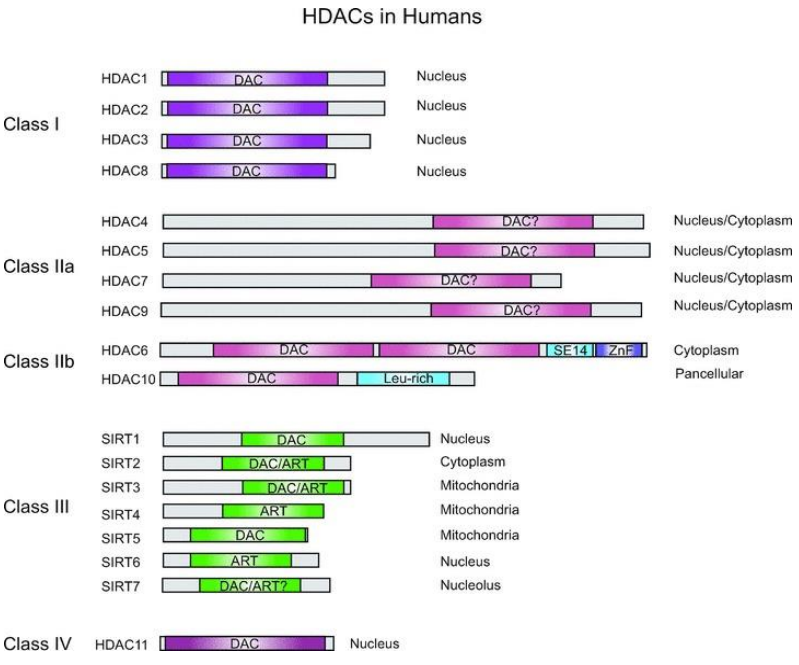


Fig.1.2 The histone deacetylase family is subdivided in subfamilies (Clocchiatti et al., 2011).

1.2.1 HDACs IIA IN HEALTH AND DISEASE

Dysregulation of class IIa HDAC has been implicated in a wide variety of diseases, as a consequence of the complex regulatory networks and multifaceted activities exerted by this class, notably in cardiovascular development, immune response, cancer, and neurodegenerative diseases (Zhang et al, 2002, Clocchiatti et al, 2011, Mielcarek et al, 2013), leading to an increasing interest towards their therapeutic targeting.

Currently, most biological functions of class IIa HDACs can be linked to the axis with the MEF2 TFs, exemplified during the embryogenesis stage in which the detachment from MEF2 is fundamental for beginning the myogenic development, reinforced by the existence of negative feedback loops (McKinsey et al., 2000, Haberland et al., 2006). In primary cardiomyocytes, HDAC5/9 counteract the hypertrophic adaptation guided by MEF2A: in fact, HDAC9 knockout mice become more sensitive to hypertrophic stimuli (Zhang et al., 2002). During bone formation, HDAC4 negatively regulates MEF2C-mediated chondrocyte hypertrophy, and the HDAC4-null phenotype results lethal within 2 weeks due to severe bone malformations and ectopic ossification (Vega et al, 2004, Arnold et al., 2007). HDAC7, on the other hand, plays crucial roles in vasculogenesis: it is specifically expressed in the vascular endothelium during embryogenesis, where it participates in the maintenance of vascular integrity through the repression mediated by MEF2 of the expression of matrix metalloproteinase-10 (MMP-10), a secreted endoproteinase that degrades the extracellular matrix (ECM); in fact, HDAC7-null mice are embryonic lethal because of the rupture of blood vessels (Chang et al., 2006). HDAC5, upon phosphorylation by protein kinase D (PKD), dissociates from GATA1 because of erythropoietin influence, in this manner affecting red cell development (Delehanty et al., 2012). Additionally, HDAC7 is highly expressed in $CD4^+$ $CD8^+$ double-positive thymocytes, where it is involved in the T cell receptor (TCR)-mediated apoptosis as a result of nuclear extrusion and Nur77 expression (Dequiedt et al., 2003). Evidence is emerging for a role of class IIa HDACs also in diabetes pathogenesis, since these epigenetic regulators seem to regulate the production of hormones involved in blood glucose levels and insulin resistance, particularly in skeletal muscle, through the modulation of GLUT4 mediated by HDAC5 (McGee et al., 2008).

Regarding the involvement of class IIa HDACs in cancer development and progression, evidence is still limited, as most studies have focused the attention on class I HDACs; nonetheless, it seems to be strictly context dependent (Ropero and Esteller, 2007). For example, HDAC4, which was characterized as a transcriptional repressor more than 20 years ago (Wang

et al., 2000), has been implicated in the DNA damage response when it is recruited at repair foci by interacting with 53BP1 (Kao et al., 2003) and also in the regulation of cell proliferation as a component of the p53 pathway, since its suppression induces resistance to p53-dependent cell cycle arrest (Berns et al., 2004). Moreover, HDAC4 is involved in triggering apoptosis as a result of caspase cleavage and release of cytochrome c (Paroni et al., 2004). In STS sub-groups, HDAC4 has been established to have a pro-oncogenic role dependent on the repression of some MEF2 targets (Di Giorgio et al., 2013) and, specifically, the pro-survival effect is particularly effective in leiomyosarcoma in concert with HDAC9 and MEF2D complexes which exert a dominant positional effect, mostly from distant intergenic regions (Di Giorgio et al., 2020). The involvement of HDAC7 in the control of proliferation and angiogenesis is a smoking gun for its role in cancer (Clocchiatti et al., 2011); for instance, higher HDAC7 (and HDAC9) expression is correlated with poor prognosis in childhood acute lymphoblastic leukemia (ALL), but an alteration of HDACs expression appears as a common characteristic of many tumors (Moreno et al., 2010). More recently, HDAC7 together with HDAC1 have been found to be required for the maintenance of cancer stem cells (CSCs) in both breast and ovarian tumors, and the inhibition of HDAC1/3 resulted in the downregulation of HDAC7 leading to decreasing H3K27ac levels at the TSS and super-enhancers (SE) (discussed in the paragraph 1.3), selectively in CSCs (Caslini et al., 2019). One peculiar characteristic of HDAC9 is represented by the extensive alternative splicing events it undergoes; in particular, MEF2-interacting transcription repressor (MITR) was the first to be characterized, lacking the C-terminal domain but maintaining the capability of MEF2 repression (Sparrow et al., 1999); furthermore, HDAC9 is itself a target of MEF2 and a negative feedback loop exists that is particularly effective during muscle differentiation (Haberland et al., 2006). HDAC9 appears to be also involved in the Foxp3-dependent suppressive function of T_{reg} cells after treatment with a HDAC inhibitor (Tao et al., 2007). Moreover, HDAC9 is highly expressed in the nervous system where in developing cortical neurons it seems to regulate dendritic growth (Sugo et al., 2010). HDAC9 (and HDAC5) high expression has been found to be correlated with poor overall survival in medulloblastoma patients and silencing of these HDACs led to decreased cell viability *in vitro* (Milde et al., 2010). HDAC9 is over-expressed in breast cancer cells, particularly in the basal subtype, where it exerts a pro-tumorigenic function by targeting SOX9 and decreasing the efficacy of HDACs inhibitors (Lapierre et al., 2016). In mammary epithelial cells, HDAC7 contributes to cell proliferation and stem-like status by regulating the surrounding microenvironment by repressing a number of cytokines including IL-24 (Cutano et al., 2019). The broad-range HDAC inhibitor vorinostat has been approved for the treatment

of T cell lymphoma and several other inhibitors are under study, as for instance the selective class IIa HDAC inhibitor TMP195 (Lobera et al., 2013; Wu et al., 2019).

1.2.2 TARGETING THE MEF2-HDAC IIA AXIS WITH SMALL MOLECULES

The consolidated data about the dysregulation of several HDAC members in different cancers have highlighted the possibility of interfering with HDACs function to restore the altered epigenome. The discovery of small molecule inhibitors, either class-specific or broad-targeting, which have relevant biological and clinical effects demonstrates how promising this research field is, as an adjuvant anticancer treatment in a combinatorial regimen with conventional chemotherapeutics to reach synergistic effects (West and Johnstone, 2014; Hontecillas-Prieto et al., 2020). Based on the chemical structure, HDAC inhibitors (HDACi) are grouped in distinct classes: hydroxamic acids (trichostatin A, vorinostat), carboxylic acids (valproate, butyrate), aminobenzamides (entinostat, mocetinostat), cyclic peptides (apicidin, romidepsin), epoxy ketones (trapoxins), and hybrid molecules. The first HDACi were identified many decades ago, when the aspecific inhibitor butyrate was found to induce cell cycle arrest and differentiation *in vitro* (Candido et al., 1978), followed in later years by the characterization of the antibiotic trichostatin A (Yoshida et al., 1995) and SAHA (suberoylanilide hydroxamic acid) as pan HDAC inhibitors. SAHA became the first FDA-approved HDACi for the treatment of advanced cutaneous T cell lymphoma, followed by romidepsin, whereas many others are clinically tested (Xu et al., 2007). These compounds target the Zn^{2+}-dependent catalytic site, thus not appearing particularly effective and specific against class IIa HDACs; moreover, the discovery of class-specific HDACi would greatly benefits in terms of safety with less side effects, although HDACi are generally well-tolerated in spite of the deleterious consequences of experimentally removing single HDAC members, due to the transient nature of pharmacological action and, probably, the non-dissociation of the repressive complexes HDACs take part in (Haberland et al., 2009). To selectively affect class IIa HDACs, three main targets potentially exist: the zinc binding domain (ZBD), the N-terminal domain and the nuclear-cytoplasmic shuttling; these strategies could directly or indirectly interfere the binding of HDACs with their partners, most importantly MEF2, thus repressing the pro-oncogenic function, for example in leiomyosarcoma cells (Di Giorgio et al., 2013). However, the clinical relevance of these molecules is not limited to cancer, extending to endothelial dysfunctions, angiogenesis, skeletal abnormalities, and neurodegenerative diseases. Likely, all these

conditions present an altered epigenome which could be susceptible of pharmacological treatment to restore the chromatin to a pre-disease state (Haberland et al., 2009).

1.3 INVESTIGATING EPIGENETIC MARKERS AND TRANSCRIPTOMIC DATA

Epigenetics is referred to as the potentially heritable genetic and phenotypic patterns which do not involve a change of the DNA sequence (Kouzarides, 2007). This additional layer of information adds to the complexity of the genome, since in contrast to the singularity of the latter a multitude of epigenomes exist, representing a key contributor to cell identity and gene expression regulation.

Development, intended in general terms, is an epigenetic phenomenon, since epigenetic changes have the great advantage of being flexible and reprogrammable; in particular, histone marks offer a valuable and short-term medium to accomplish cell diversity and gene regulation, in contrast to the more stable changes in DNA methylation patterns (Reik, 2007). Epigenomics aims to study global epigenetic changes, benefiting from an unbiased approach and the application of the most recent high-throughput technologies. These epigenetic changes are involved in virtually all cellular processes to some extent, both physiological and disease-related, constituting a "chromatin-based signaling" (Jenuwein and Allis, 2001; Audia and Campbell, 2016).

The study of epigenetics has greatly benefited from the implementation of high-throughput next-generation sequencing technologies, such as ChIP-seq and Hi-C, which have made possible the unbiased and hypothesis-free analysis of chromatin structure and histone modifications (Rivera and Ren., 2013). Relevant large-scale projects, like ENCODE (Encyclopedia of DNA Elements, 2003), have also contributed to mapping a reference of the epigenome and functional elements (The ENCODE Project Consortium, 2012; Roadmap Epigenomics Consortium, 2015).

A high number of epigenetic markers exist: acetylation, methylation, phosphorylation, ubiquitination and other classes of modifications of the core histones offer a varied and fine-tuned control on gene expression, the so-called histone code which appears not entirely deciphered yet. Together with DNA methylation, acetylation and methylation of lysine residues found in the histone tails are the most frequently investigated epigenetic markers, as they can correlate with the formation of chromatin domains, specifically euchromatin and heterochromatin. Some histone marks such as H3K27me3 or H3K9me2 are generally associated with condensed chromatin and are enriched for example in the inactive X

chromosome, while others (H3K27ac and H3K4me2) are often found at promoters or enhancers of transcriptionally active genes (Bartova et al., 2008). Specific enzymes are responsible for these modifications and must achieve an equilibrium, as is the case of histone deacetylases (HDACs) which antagonize the action of histone acetylases (HATs) removing an acetyl group to the N-terminal tails of histones (Bannister and Kouzarides, 1996). Epigenetic alterations can be found in many diseases, including cancer, where the methylation of tumor suppressor genes' promoters can represent a mechanism leading to their silencing, and treatment with HDACi and demethylating agents is able to reactivate the expression of these important genes (Yamashita et al., 2018). In particular, H3K27ac is one of the most studied histone marks in gene expression profiling and developmental studies, as it is associated with decondensed and accessible chromatin, thus representing a robust epigenetic marker of gene expression activation, also helping distinguishing active from inactive enhancers, which are delimited genomic regions (50–1500 bp) promoting gene transcription in *cis* via distal mechanisms (up to 1 Mb), either up- or down-stream from the target gene; in the human genome hundreds of thousands of enhancers are estimated to exist, scattered across the non-coding DNA (Creyghton et al., 2010; Pennacchio et al., 2013). The term super-enhancer (SE) is applied to describe clusters of enhancers in genomic proximity with high levels of Mediator (Med1) and RNApol II binding, eRNAs (unstable RNAs transcribed bidirectionally from enhancer loci and correlated with their activity), chromatin marks and other features. SEs span large genomic regions, generally an order of magnitude larger than traditional enhancers; moreover, in cancer cells SEs are enriched at oncogenes, and some identified translocations can bring SEs in proximity of oncogenes (Pott and Lieb, 2015). For this reason, new therapeutic strategies aim at targeting SEs to hamper the expression of oncogenes, such as Myc in hematologic malignancies (Jia et al., 2019).

Transcriptomics accounts for the qualitative and quantitative analysis of the whole transcriptome (i.e. the sum of all RNA transcripts), allowing a thorough analysis of the gene expression profile of an organism, of a specific tissue or cell (Lowe et al., 2017). Before the advent of transcriptomics, the studies focused for many decades mainly on individual transcripts; in the 1980s, the conception of the Sanger sequencing was used to sequence random fragments of individual transcripts called expressed sequence tags (ESTs), useful for the identification of new genes without the need of sequencing the whole genome (Marra et al., 1998); other widely applied quantitative methods, although laborious and low-throughput, were the Northern blotting and the reverse Transcriptase quantitative PCR (RT-qPCR) (Alwine et al., 1977; Becker-André et al., 1989). The term 'transcriptomics' was initially adopted in the

1990s, when SAGE (serial analysis of gene expression) was introduced, which took advantage of the Sanger sequencing of concatenated random transcript fragments, where the transcripts were quantified by matching those fragments to known genes (Velculescu et al., 1995). The advent of microarrays (oligonucleotide arrays and Affymetrix's GeneChip platforms were the most widely diffused) represented a breakthrough (Schena et al., 1995), allowing the determination of transcript abundance via the hybridization of thousands of fluorescently-labeled transcripts on an array of complementary probes (short nucleotide oligomers), becoming the method of choice until the late 2000s. Technological advances in manufacturing and fluorescence detection greatly improved over time the specificity and sensitivity of microarrays, in particular for low abundance transcripts (McLachlan et al., 2005). Thanks to the increasing adoption of next-generation sequencing, which progressively overcame Sanger sequencing, RNA-sequencing (RNA-seq) could be developed, based on the counting of sequenced transcripts cDNAs, with the earliest reports published in 2006 (Bainbridge et al., 2006) and increasing its popularity after 2008, allowing the quantification of the entire human transcriptome (Mortazavi et al., 2008). The analysis of the transcriptome has broad applications, from diagnostics to gene function annotation and study of the non-coding portion of the genome. For example, RNA-seq allows the dissection of regulatory features, splicing variants, SNPs, allele-specific expression, and gene fusions associated with diseases (Ozsolak and Milos, 2010; Khurana et al., 2016). The great majority of transcriptomic experiments is devoted to the analysis of protein-coding messenger RNAs (mRNAs); however, the same techniques can be also applied to non-coding RNA, primarily miRNAs and long non-coding RNAs (lncRNAs), which are involved in many important cellular biological processes, from RNA splicing to protein translation, potentially contributing to the onset of several diseases, including cancer (Hüttenhofer et al., 2005; Esteller et al., 2011).

1.3.1 THE CHIP-SEQ TECHNOLOGY ALLOWS THE STUDY OF EPIGENETIC CHANGES AND PROTEIN-DNA INTERACTIONS

The ChIP-seq technique was first introduced more than 10 years ago; combining chromatin immunoprecipitation with massively parallel next-generation sequencing; this powerful methodology makes possible to implement the genome-wide discovery of transcription factor binding sites and the prevalence of histone marks, replacing and improving microarray-based ChIP-chip thanks to lower starting material requirements and offering higher resolution and greater coverage at a lower cost; nonetheless, the implementation of ChIP-seq has required the development of new and still evolving analytical tools (Zhang et al., 2008; Park, 2009). The

essential concept at the basis of ChIP-seq is the enrichment of DNA segments bound by a protein (TF) of interest or harboring a specific histone mark; the DNA fragments are sequenced (most frequently single-end reads from 5' end only) and then aligned to a reference genome to investigate for example the TF binding strength and/or pattern between two experimental conditions (differential binding), unearthing the underlying molecular insights regarding a biological effect. Currently, this technique has become the gold standard for dissecting DNA-protein binding events and large datasets have been generated for several TFs and histone modifications, which can be harnessed to test hypotheses about the mechanisms of gene regulation. Given the importance of DNA-binding proteins in many processes, for example in transcription, replication and DNA repair, ChIP-seq experiments and data analysis have become a critical asset for many research groups worldwide. Depending on the research question, ChIP-seq experiments can be essentially classified in point-source, broad-source, and mixed-source, based on the nature of the signal: if highly localized (TFs), spanning large regions of the DNA (some histone modifications such as H3K36me3) or displaying both patterns (RNApol II), respectively. The peak calling is the computational method used to discover the genomic areas in which the alignment of several reads has produced a pileup (enrichment), corresponding to a putative TF binding site, an indicator of chromatin state or DNA methylation. The selection of the most suitable peak calling tool, as for example the popular MACS2, represents a signal to noise problem and the chosen parameters strongly affect the outcome; additionally, the researcher must decide if prioritize specificity or sensitivity. The choice of the peak finder will also depend on the type of ChIP signal, as well as the sequencing depth to be used, since a higher depth (optimally 20-40M reads) allows the detection of a higher number of sites with reduced enrichment, as it is usually the case with a broad signal (Furey, 2012). Finally, careful planning must be done while preparing experimental controls (input, lacking the immunoprecipitation step and required for signal normalization) and biological replicates, since at least two are required as the ideal condition to reach a greater robustness of results. This large-scale information can be further complemented with additional functional genomics assays (e.g. DNAse I footprinting, FAIRE-seq and Hi-C) to predict the chromatin state, gene expression patterns and long-range interactions (Jiang and Mortazavi, 2018). Moreover, the recently developed single-cell ChIP-seq allows to investigate intra-tissue cellular diversity and intra-tumor heterogeneity (Nakato and Sakata, 2021).

1.3.2 RNA-SEQ OFFERS AN UNBIASED QUANTITATIVE TRANSCRIPTOME PROFILING TOOL

RNA-sequencing represents a revolutionary method to quantitatively assess the level of transcripts by using next-generation sequencing (NGS) technologies, overcoming the intrinsic limits of microarrays (e.g. the limited dynamic range and the higher amount of starting material (micrograms rather than nanograms) (Wang et al., 2009). Over the last decade, RNA-seq has become an invaluable tool for transcriptome-wide analysis and differential gene expression profiling, evolving with the advancement of NGS technologies, and extending nowadays to the study of several aspects of RNA biology, from its structure to spatial transcriptomics (spatialomics) and single-cell gene expression (scRNA-seq) (Stark et al., 2019). RNA-seq is most commonly used to study mRNA as a proxy for protein expression, but the vast RNA output from noncoding regions, like long non-coding RNAs (lncRNA), small regulatory RNAs (sRNA) and enhancer RNAs (eRNA), has not to be neglected, virtually permitting the simultaneous profiling of all RNA classes subsequently distinguished *in silico* (Westermann et al., 2016). RNA-seq is also becoming an important tool in the clinical routine, allowing for example the detection of novel splicing variants potentially linked to genetic diseases (Marco-Puche et al., 2019). This technological evolution has paralleled the development of better computational tools for data analysis and interpretation of increasingly complex datasets, from read mapping and batch effect correction to normalization (Chen et al., 2019). Since a typical RNA-seq experiment generates a large amount of raw data, the implementation of robust algorithms is essential to process it and several bioinformatics tools exist, each designed to address a specific analytic purpose depending on the experimental design. Four main processing stages can be identified: quality control, alignment, quantification, and differential expression (Van Verk et al., 2013). In particular, the quantification of sequence alignments (over the reference genome) may be performed at the gene, transcript or exon level, generating a table of read counts for each feature, ready for the differential gene expression analysis associated with data normalization and statistical assessment, providing probability estimates of these differences. A common and accurate method employed for transcript abundance quantification is RSEM, derived from the homonym package and useful for both single-end and paired-end RNA-seq data, without requiring a reference genome (Li and Dewey, 2011). In the end, transcriptomic analyses can be also finally validated by independent techniques, such as quantitative PCR (qPCR) *in vitro* and knockdown studies *in vivo* (Lowe et al., 2017).

1.4 FOCUS ON CANCER

1.4.1 GENETIC ALTERATIONS AT THE BASIS OF CANCER

The maintenance of genome integrity is a prerequisite for the proper functioning of multicellular organisms, and this requires the intervention of several safe-guard mechanisms for repairing the damaged DNA or inducing apoptosis when necessary. However, the events at the basis of the carcinogenic process pose a great challenge to the correct preservation of the genome, resulting in varying degrees of mutation burden and genetic instability, which themselves represent a hallmark of cancer with a direct causal relationship (Dixon and Kopras, 2004; Armaghany et al., 2012). Among the major somatic tumor alterations of the mammalian genome, gene deletions or amplifications and single nucleotide variants (SNV) contribute mostly to the malfunctioning of the cancer cell, together with the still poorly characterized epigenomic alterations which are, however, more difficult to profile routinely in a clinical setting. Moreover, large-scale chromosomal alterations and rearrangements occur in cancer as well, as it was firstly documented decades ago for leukemia patients through cytogenetic studies (Balmain, 2001). Overall, the model of proto-oncogene activation and tumor suppressor inactivation as principal drivers of tumorigenesis in diverse tissues is well-established, as it mechanistically explains the accelerated behavior of cancer cells. In fact, molecular alterations drive phenotypic changes in cancer cells and their microenvironment, as it is usually assessed through histopathology. Nowadays, the traditional methodologies can potentially be aided by machine learning, as recently demonstrated with a deep learning algorithm able to predict a wide range of molecular alterations from routine, paraffin-embedded histology slides, potentially favoring the implementation of personalized cancer therapy (Kather at al., 2020).

1.4.2 THE TCGA RESOURCE FOR CANCER RESEARCH

The Cancer Genome Atlas (TCGA) represents a collection of genomic, transcriptomic, and clinical data covering a multitude of cancer types. The pilot project was launched in 2005 by the National Institutes of Health (NIH), as a public funded project to catalog cancer-related genome alterations in large cohorts of clinical samples, initially limited to brain, lung and ovarian cancers, while in a later phase (2014) extending to more than 30 types of tumors and counting about 11 thousand profiled patients. Several centers contribute at managing the well-organized structure of TCGA: different Tissue Source Sites (TSSs) collect the required biospecimens and send them for sample processing to the Biospecimen Core Resource (BCR) for cataloging and quality control, followed by the submission of clinical data to the Data

Coordinating Center (DCC); subsequently, molecular analytes are delivered to the Genome Characterization Centers (GCCs) and Genome Sequencing Centers (GSCs) for high-throughput sequencing and bioinformatics analyses. The genomic data are made available to the research community and Genome Data Analysis Centers (GDACs). The GDACs implement information-processing and visualization tools for a broader usage of TCGA data. The aim is to provide publicly available cancer genomic datasets that will allow the improvement of diagnostic methods, treatment standards and ultimately cancer prevention (Tomczak et al., 2015). Several investigators have already taken advantage of this plethora of information, publishing highly cited papers which have contributed to the advancement of cancer biology understanding (Cancer Genome Atlas Network, 2008, 2011, 2012, 2014, 2015, 2017).

1.4.3 COLORECTAL CANCER IS ONE OF THE MOST PREVALENT TUMORS WORLDWIDE

Colorectal cancer (CRC) is the third most common cancer diagnosed every year globally, with approximately 1.2 million cases and over 600 thousand deaths. Its incidence is strongly correlated with aging, as the median age at diagnosis is 70 years in developed countries and is typically linked to a so-called Western lifestyle. Most cases are sporadic despite the existence of predisposing hereditary conditions, such as familial adenomatous polyposis and Lynch syndrome, but these account for only 5% of CRC cases. However, a large twin study twenty years ago seemed to suggest that a substantial heritable component exist, with 35% of CRC risk which may be attributable to heritable factors, albeit the underlying genetics is still poorly characterized (Brenner et al, 2014; Lichtenstein et al., 2000). In more recent years, the prognosis has steadily improved, although a disparity exists between high- and low-income countries. CRC staging indicates the degree of extension of the tumor and is obtained through the TNM system, which accounts for local invasion (T), lymph node involvement (N) and metastasis spreading (M); as with most tumors, the stage at diagnosis represents the most critical prognostic factor, with 5-year survival dramatically dropping from stage I (over 90%) to stage IV (about 11%); additionally, the stage affects the therapeutic decision. There is not a single risk factor responsible for most CRC cases, however the following have been well-recognized: age, male sex, familiarity (especially first-degree relatives) and inflammatory bowel disease, while others appear modifiable: smoking, excessive alcohol consumption, obesity, diabetes, Helicobacter pylori infection and high intake of red and processed meat. Among the protective factors, there are physical activity, hormone replacement therapy, aspirin and a healthy dietary pattern in addition to endoscopy for the removal of precancerous lesions

(Brenner et al., 2014). The molecular pathogenesis of CRC is heterogenous yet clinically relevant, as it can affect the prognosis and the response to treatment. The traditional model was proposed in the 80s and today is well-established, represented by the adenoma-carcinoma sequence, a slow and multistep process evolving through 10 or even more years, with the APC (adenomatous polyposis coli) oncosuppressor mutation occurring as one of the earliest events disrupting the Wnt signaling and leading to dysplastic benign (pre-malignant) adenomas; APC functions in a multitude of cellular processes, from mitosis, cell migration, regulation of genome stability and stabilization of CTNNB1 (β-catenin), a key component of the Wnt pathway (Dunn et al., 2013). The further progression to invasive carcinoma is promoted by the activating mutation of the KRAS oncogene and the inactivation of the TP53 tumor suppressor, often accompanied by chromosomal instability. Indeed, APC restoration in mice can revert the tumoral phenotype into differentiated normal crypt cells, independently from the KRAS and TP53 mutation status (Dow et al., 2015). However, about 10-15% of sporadic cases lacking APC alteration cannot be explained by this model, pointing to additional molecular events: for example, the presence of genomic instability which can take the form of chromosomal instability, microsatellite instability or CpG methylator phenotype (CIMP), and furthermore epigenetic alterations, involving in particular promoter hypermethylation of tumor suppressor genes. Such alterations are commonly associated with the BRAFV600E mutation and serrated precursor lesions, which are more difficult to diagnose (Armaghany et al., 2012; Brenner et al., 2014).

1.4.4 LEIOMYOSARCOMA IS A RARE TUMOR, CURRENTLY DIFFICULT TO TREAT

Soft tissue sarcomas (STS) are a heterogenous group of rare tumors of mesenchymal origin, which includes over 50 histological subtypes, with the exclusion of parenchymatous organs and Kaposi sarcomas. These tumors are generally aggressive and with poor prognosis, indeed about one patient in two will develop relapse or metastasis (Cormier and Pollock, 2004). In particular, leiomyosarcoma appears as an aggressive STS with currently few therapeutic options; this tumor can be diagnosed at various sites, mostly the retroperitoneum, the uterus, the dermis and the blood vessels. The precise tumor initiating cell (TIC) is still a mystery, also because a corresponding benign form does not exist, even if presumably it is represented by a mesenchymal stem cell which failed the proper differentiation program into a smooth muscle cell (Danielson et al, 2010); some evidence highlighted the role of TP53 alterations in this context, as the re-expression of wild-type TP53 reduces tumorigenicity in uterine LMS (ULMS)

cells (Rubio et al, 2012; Pollock et al, 1998). Moreover, high grade LMS presents 3 times more amplifications compared to low grade LMS, in which deletions are more common, suggesting that the inactivation of oncosuppressors may precede the activation of proto-oncogenes in tumor initiation and progression (Yang et al., 2009). Noteworthy, some organ-specific alterations occur, like the over-expression of cell cycle-related genes in ULMS (Barlin et al., 2015). Being scarcely responsive to chemo- and radiotherapy, chirurgical resection represents the main solution for affected patients, with 5-year survival approaching 50%. Due to the complex karyotype and the absence of recurrent genetic alterations, a precise medicine treatment for LMS is still missing. An RNA-seq analysis revealed a molecular sub-classification associated with different clinical outcomes, although not clearly corresponding to histological markers, complicating its interpretation (Guo et al., 2015). Myocardin seems to exert an onco-suppressive function in ULMS (Kimura et al., 2010), and this is similarly true when over-expressing MEF2, which overcomes the HDAC4-mediated repressive function and oncogenic potential (Di Giorgio et al., 2013). However, recent evidence supports the idea that MEF2 in LMS can have either pro-oncogenic or onco-suppressive roles, depending on the interaction with co-activator or co-repressor (class IIa HDACs) complexes, which influence the underlying genetic programs involved; in particular, class IIa HDACs are over-expressed in about one quarter of LMS, where high expression of MEF2, HDAC4 and HDAC9 inversely correlate with overall survival. In this context, HDAC9 knockout inhibits the transformative potential of LMS cells, making HDAC9 a potential therapeutic target (Di Giorgio et al., 2017). More recently, the MEF2D-HDAC9 relationship, under a negative feedback loop deregulated in LMS, gained attention as a promoter of cell survival through the repression of FAS, interestingly presenting mainly distant binding at enhancers, a new modality of gene regulation described for class IIa HDACs (Di Giorgio et al., 2020).

1.5 THE TUMOR MICROENVIRONMENT RECLAIMS A PROMINENT ROLE IN MOST SOLID TUMORS

Cancer cells do not sustain their uncontrolled proliferation by themselves but depend on the establishment of complex interactions with the surrounding non-transformed stromal and immune cells, which overall constitute the tumor microenvironment (TME), a niche that can nurture the growth and invasion of the tumor after losing its homeostatic function. Indeed, the TME is not static but dynamically changes with cancer progression, with either pro-tumorigenic or anti-tumorigenic roles, thus emerging as a promising therapeutic target (Quail and Joyce, 2013). One of the first pieces of evidence came from the increased incidence of cancer in tissues

subject to chronic inflammation, where an accumulation of activated cells leads to a malfunctioning tissue homeostasis, as for instance is the case of cirrhosis and hepatocellular carcinoma or inflammatory bowel disease and CRC (Grivennikov et al., 2010). Furthermore, retrospective studies have shown that the incidence of multiple cancer types can also be correlated with an impaired immune response, for example in transplant recipients, suggesting a protective effect exerted by a properly functioning immune system. Several actors in the TME take part in promoting tumorigenesis, most notably cancer-associated fibroblasts and tumor-associated macrophages (CAFs and TAMs, respectively), as a result of a phenotypic switch driven by the growing tumor (Qian and Pollard, 2010). For instance, TAMs are the result of the functional plasticity (polarization) of macrophages and promote the invasive potential of tumor cells by releasing proteases and cytokines (EGF, CSF-1), through a paracrine signaling loop (Quail and Joyce, 2013). CAFs, abundant in the TME, are activated by a multitude of cytokines and, at least in breast cancer, have the capability to induce a mesenchymal phenotype which favors the occurrence of metastasis (Dumont et al., 2013). There are additional players in the tumor milieu, from the myeloid as well as the lymphoid lineages, such as myeloid-derived suppressor cells (MDSCs) and regulatory T cells (T_{reg}). MDSCs are immature mixed myeloid cells with immunosuppressive properties, which are mobilized and then infiltrate the developing tumor enhancing vascularization and metastasis, disrupting the mechanisms of immunosurveillance (Gabrilovich et al., 2012). Regulatory T cells in physiological conditions coordinate in a complex manner the activity of T and B cells, besides that of innate cytotoxic lymphocytes, thus their effect on diverse types of tumors may result the opposite, correlating with either poor (breast and HCC) or improved survival (CRC) (Whiteside et al., 2012, Frey et al., 2010). Overall, the presence of these immunosuppressive cell populations in the TME significantly contributes to halt an effective immune response targeted against the tumor, and this needs to be considered when developing immunotherapeutic strategies (Ohue and Nishikawa, 2019).

1.6 HYPOXIA IS A HALLMARK OF THE TUMOR MILIEU

The TME is characterized by a decreasing gradient of oxygen moving from the external to the inner zones of the tumor, resulting in dynamically changing levels of hypoxia (when oxygen demand is superior to the supply) in about half of the solid tumors and representing a negative prognostic factor (Harris, 2002). Hypoxia poses a strong selective pressure on cancer cells, which must adapt to this adverse condition acquiring more malignant characteristics, thus

favoring the emergence of apoptosis-deficient or TP53-mutated cells; this adaptation is associated with increased genomic instability and mutagenesis, resistance to cytotoxic drugs, metastatic dissemination and poor prognosis (Bristow and Hill, 2008; Luoto et al., 2013; Dhani et al., 2015). Moreover, a hypoxia-driven metabolic reprogramming is another key stone, with a rewiring of cellular energetic networks towards a glycolytic metabolism, which contributes to sustain the rapid proliferation and growth of cancer cells (the Warburg effect), a discovery that goes back to 1930 (Phan et al., 2014). Hypoxia inducible factor 1α (HIF-1α), which is specifically expressed and stabilized during hypoxia, is the master regulator of this phenomenon and supervises a high number of downstream genes crucial for cancer progression, among which VEGF and erythropoietin, that help in promoting oxygen delivery through new blood vessel formation (Semenza, 2003; Lee et al., 2004). Consequently, hypoxia has emerged as a promising therapeutic target, for instance through bioreductive prodrugs activated only in the hypoxic tissue or by inhibiting HIF-1α (Wilson and Hay, 2011); moreover, since hypoxia is known to down-modulate a proper immune response in the TME, targeting hypoxia could also benefit the efficacy of immunotherapy regimens (Wang et al., 2021).

Aims

Different scientific questions have been addressed in this thesis, through *in silico* data analysis tools. In the first part, we wanted to investigate the potential involvement of class IIa HDACs and MEF2 transcription factors in one of the most prevalent, yet hardly treatable and genetically complex tumors, represented by colorectal cancer (CRC). This analysis took great advantage from the TCGA resources to study gene expression, its prognostic value, and several clinical associations in a large cohort of CRC patients. Since the tumor microenvironment (TME) seems to play an essential role in the progression of the disease together with the strictly correlated hypoxia phenomenon, published gene signatures concerning these components were used to better decipher the gene expression profiles available for these patients, and potentially to assess a clinical-oriented value.

Secondly, we focused on the in-depth characterization of ChIP-seq datasets obtained from leiomyosarcoma (LMS), an aggressive and incurable cancer of mesenchymal origin in which class IIa HDACs seem to exert a pro-tumorigenic function by repressing the MEF2 signature. This analysis extends from recently published data by our research group (Di Giorgio et al., 2020), which described the complex genetic programs under the supervision of the HDAC IIa repressive complexes, specifically HDAC4 and HDAC9 in cooperation with MEF2D. By taking a divergent approach, we focused on those genes displaying a reduction of H3K27ac levels (a marker of active promoters and enhancers) in HDAC4 KO cells, implying a transcriptional activation exerted by class IIa HDACs, an overlooked behavior in most studies, given their well-acknowledged repressive role.

A third part of this thesis was devoted to dissecting the effects of the putative HDAC inhibitor NKL54, recently under study by our research group in the EPIC project together with other similar compounds, to selectively target the class IIa HDAC-MEF2 axis in LMS, aiming to investigate the H3K27ac changes occurring after NKL54 treatment in LMS cells.

Results

3.1 HDAC IIA AND MEF2 IN COLORECTAL CANCER - A TCGA ANALYSIS

3.1.1 GENETIC ALTERATIONS OF CLASS IIA HDACs AND MEF2s IN CRC

To investigate the potential roles played by class IIa HDACs and MEF2 TFs in colorectal cancer (CRC), we first explored their gene alterations and mutational landscapes in the TCGA-COADREAD cohort of patients (n=640). We observed a very low frequency of mutations and copy number variations in these genes, with HDAC9 and MEF2C being the most frequently altered in the respective families (data available at cBioPortal and summarized in Table 1).

Gene	Single Nucleotide Variants (SNV)	Deletions	Amplifications
HDAC4	4	1	-
HDAC5	2	1	2
HDAC7	4	-	-
HDAC9	11	-	3
MEF2A	2	1	2
MEF2B	1	3	-
MEF2C	7	5	-
MEF2D	1	-	3

Table 1 Gene alteration occurrence of class IIa HDACs and MEF2s transcription regulators in the TCGA-COADREAD cohort.

3.1.2 CLASS IIA HDACs AND MEF2D GENE EXPRESSION CHANGES WITH TUMOR PROGRESSION AND WITH RESPECT TO NORMAL COLON

Given the limited plausible impact on CRC tumorigenesis of the genetic alterations occurring in these gene families, our focus moved to the changes in gene expression levels and, in particular, we restricted the analysis to the RNA-seq profiled samples (n=382), evaluating if and how gene expression levels varied in CRC and with respect to the tumor progression. The stratification based on tumor stage showed that most class IIa HDACs exhibit a significant increase in mRNA expression during CRC advancement, with the only exception represented by HDAC4; by contrast, only MEF2D displays a strong increase among the MEF2 family members (Fig.1). Moreover, absolute expression levels widely change depending on the gene, with the highest mRNA levels found in the case of HDAC5, HDAC7 and MEF2D, whereas HDAC9 and MEF2B are the lowest expressed (Fig.1). Expression data were also obtained for matched normal colon samples from Firebrowse (n=50). The gene expression comparison clearly showed that HDAC7 is the only gene under investigation significantly upregulated in the tumor tissue, while the other HDACs IIa and MEF2 genes exhibit the opposite trend, being significantly downregulated in CRC compared to normal colon (Fig.2A). Secondly, as represented in the two correlograms (Fig.2B-C), in some cases the correlation coefficients between HDACs IIa and MEF2 gene expression vary substantially between normal and cancer tissues, as in the case of HDAC7, HDAC4 and MEF2A, which are strongly associated in normal tissue but more mildly associated in CRC. By contrast, in several other cases there seems to be only a slight difference, with the highest frequency of non-significant correlations found in normal tissue (presumably due to the lower number of profiled patients).

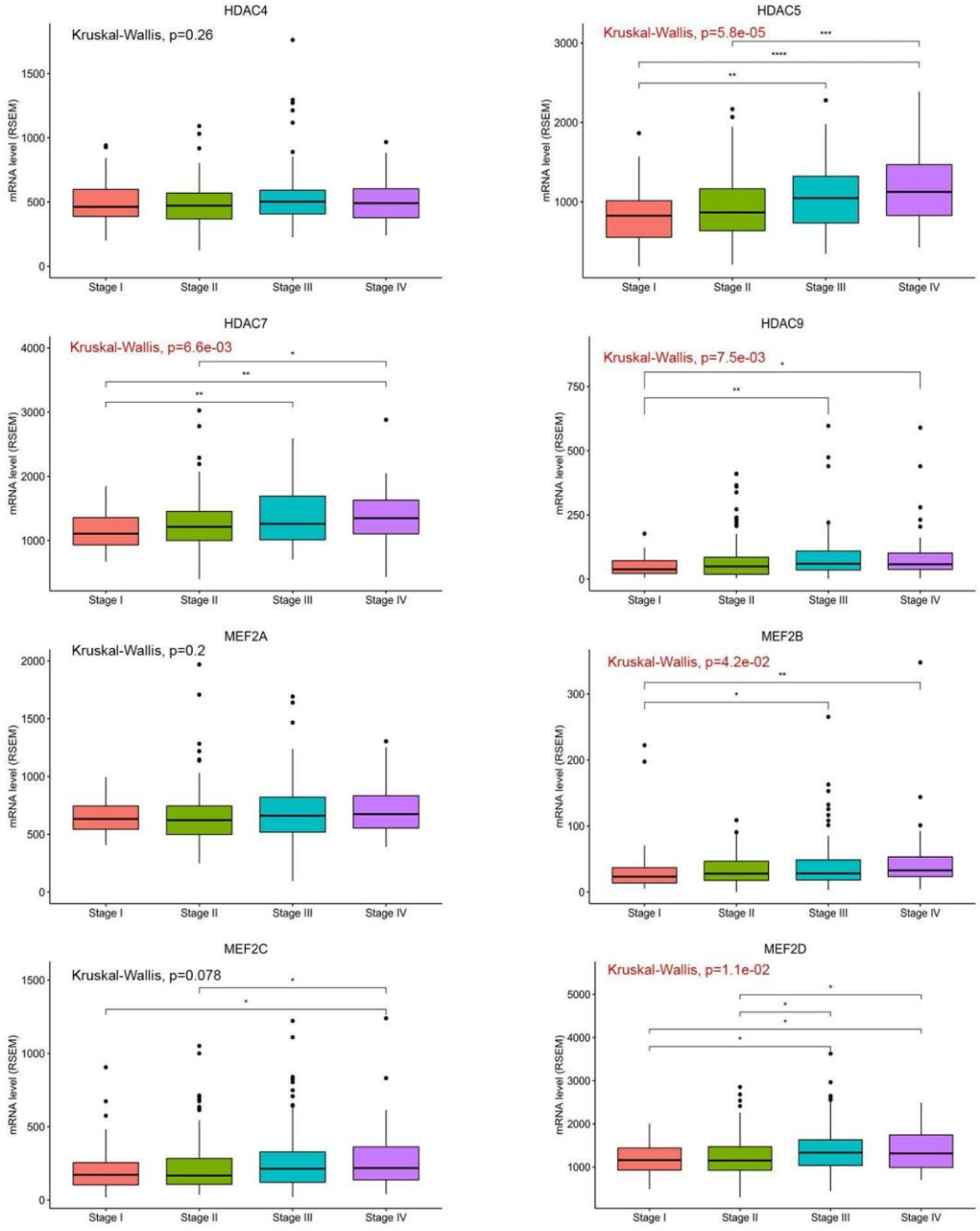


Fig.1 Boxplots showing class IIa HDACs and MEF2 mRNA expression levels (RSEM) in CRC patients stratified by tumor stage (Kruskal-Wallis global test, P<0.05 (colored in red when statistically significant), Wilcoxon paired test, *P<0.05, **P<0.01, ***P<0.005, ****P<0.001).

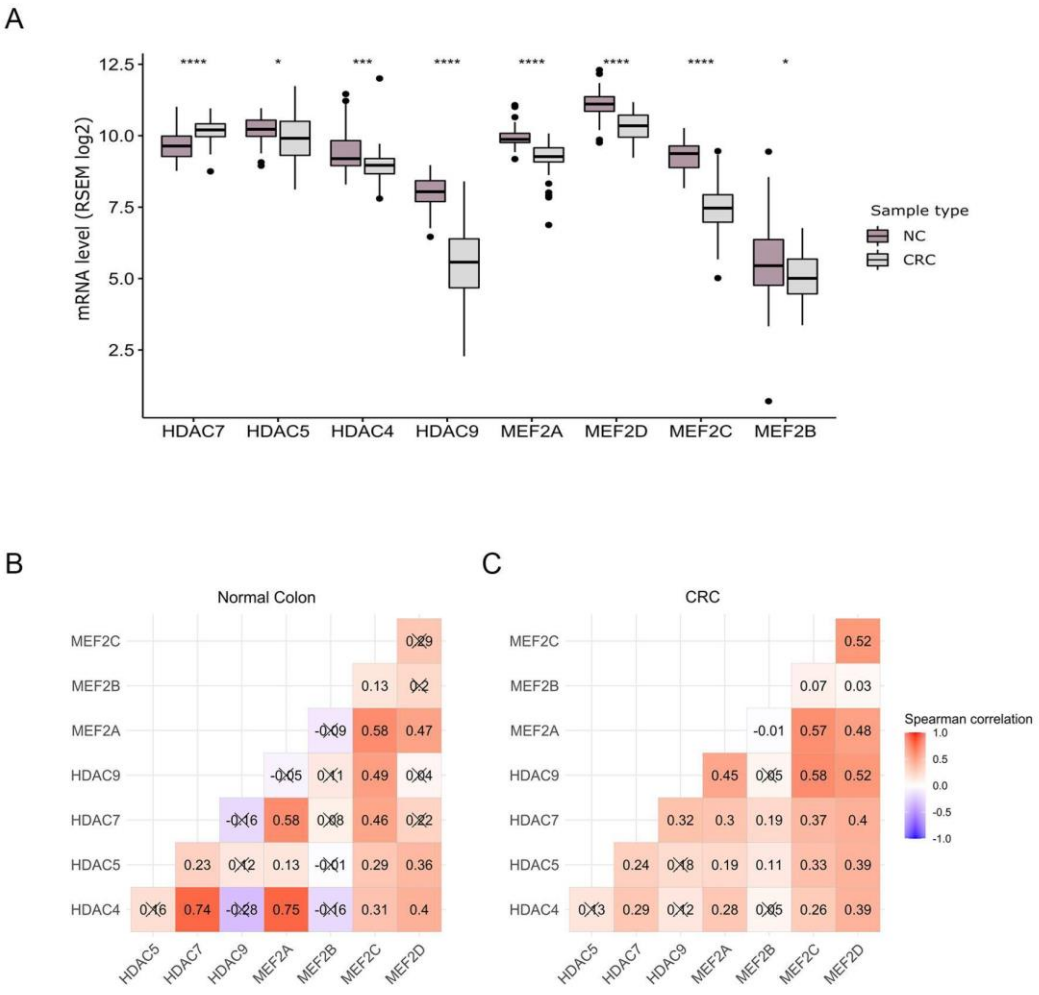


Fig.2 (A) Comparative boxplots showing the gene expression difference (RSEM log2) between normal colon (NC) and matched colorectal cancer (CRC) samples (n=50) (Wilcoxon paired test, *P<0.05, **P<0.01, ***P<0.005, ****P<0.001). Correlation matrices of class IIa HDACs and MEF2s family members' RNA-seq expression levels in (B) normal colon (n=51) and (C) colorectal cancer (n=382); a strikethrough box indicates a non-significant Spearman correlation (P>0.05).

3.1.3 EFFECTS OF CLASS IIA HDACs AND MEF2 GENE EXPRESSION LEVELS ON PATIENTS' SURVIVAL

Afterwards, we assessed the effect of HDACs IIa and MEF2 transcriptional regulators on the survival of CRC patients, adopting the median expression level of each gene as the cutoff to stratify patients in high and low expression groups. We found that HDAC7, MEF2A and MEF2C have a prognostic value, with high expression associated with significantly reduced overall survival (logrank test, $P < 0.05$, Fig.3). On the other hand, statistical significance was also obtained evaluating disease-free survival for HDAC7 and MEF2A but not MEF2C (Fig.3) We also performed a univariate and multivariate analysis of CRC data, finding that only in the case of HDAC7 the prognostic significance seems to be unaffected by the tumor stage and the fraction of genome altered (FGA), which are indeed strong prognostic attributes (Table 2). Additionally, the correlation of gene expression levels between the three prognostic genes in comparison to all the other family members varies substantially and appears as statistically significant in all cases, as represented in Table 3, with the strongest association found for HDAC9 and MEF2C, besides MEF2A and MEF2C.

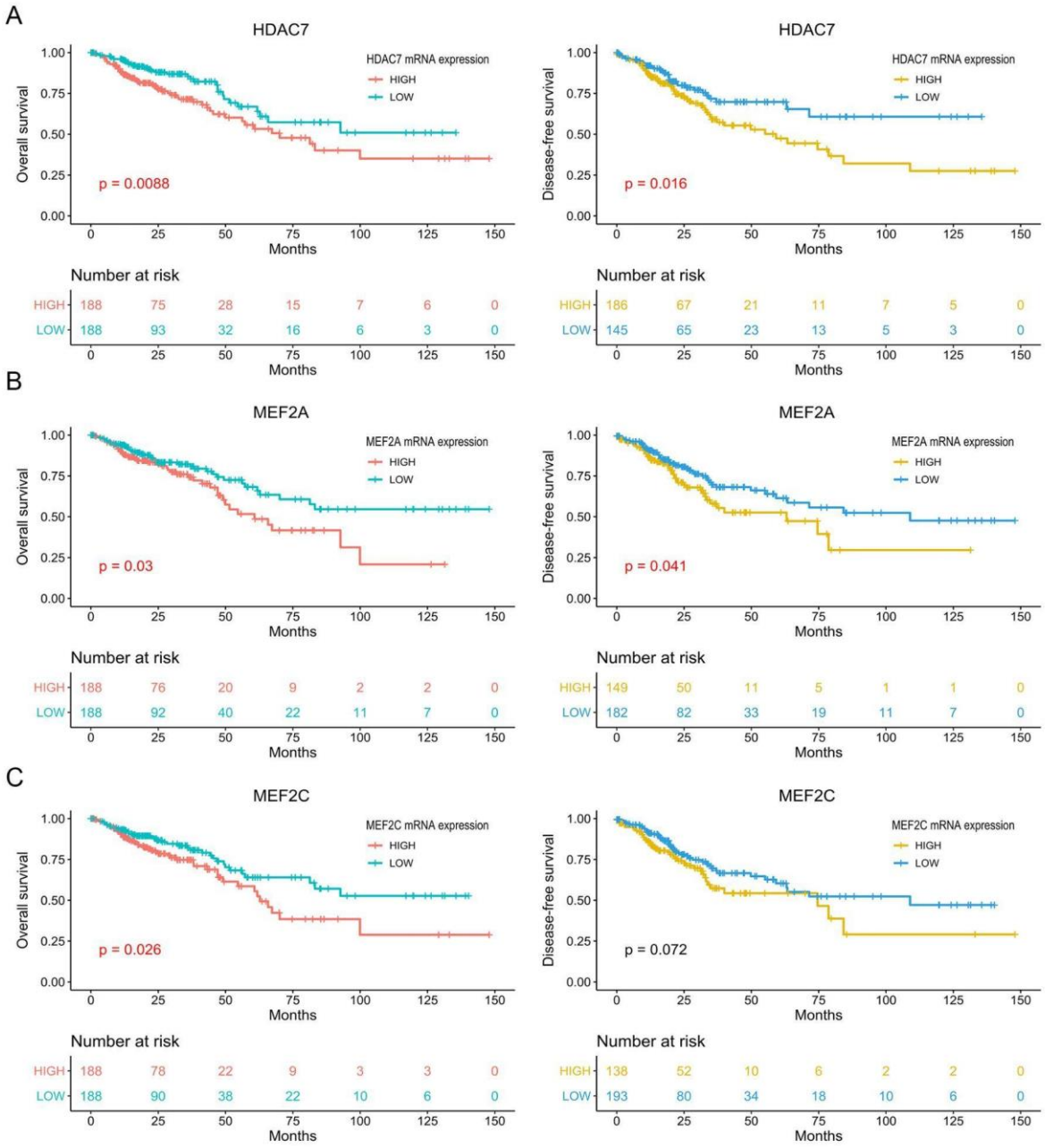


Fig.3 Survival analysis showing the Kaplan-Meier plots (overall survival (left) and disease-free survival (right), logrank test, P<0.05) of patients having high and low mRNA expression levels (RSEM) of HDAC7 (A), MEF2A (B) and MEF2C (C) genes, respectively. In the lower right plot the K-M curves partially cross, potentially violating the assumption of proportional hazards, essential in a Cox model; however, it does not appear relevant as the survival analysis does not reach statistical significance.

Variable	Univariate			Multivariate			Multivariate (step)		
	n	HR (95% CI)	P-value	n	HR (95% CI)	P-value	n	HR (95% CI)	P-value
HDAC7 (low vs high)	376	0.5646 (0.366-0.871)	0.00977	373	0.5837 (0.3738-0.9115)	0.0179	373	0.5437 (0.350-0.8422)	0.00636
MEF2A (low vs high)	376	0.6492 (0.4239-0.9941)	0.0469		0.7626 (0.4878-1.1921)	0.2344			
MEF2C (low vs high)	376	0.6222 (0.406-0.9535)	0.0294		0.7850 (0.5004-1.2314)	0.2919			
AJCC Stage (low vs high)	376	0.3285 (0.2078-0.5194)	1.9e-06		0.3688 (0.2309-0.5891)	2.99e-05		0.3558 (0.223-0.5668)	1.36e-05
Fraction of genome altered (low vs high)	373	0.5518 (0.3567-0.8536)	0.00756		0.5762 (0.3692-0.8992)	0.0152		0.5914 (0.379-0.9211)	0.02014

Table 2 Univariate and Multivariate analysis of gene expression and clinical data from colorectal cancer patients. Gene expression levels (low vs high) of prognostic genes HDAC7, MEF2A and MEF2C were considered, as well as tumor stage (low (Stage I and II) vs high (Stage III and IV)) and fraction of genome altered (FGA). HR indicates the Hazard Ratio, CI the confidence interval. A backward-forward step procedure was applied to optimize the multivariate model with the most informative variables.

	HDAC4	HDAC5	HDAC7	HDAC9	MEF2A	MEF2B	MEF2C	MEF2D
HDAC7	0.29 (1.98E-05)	0.23 (3.18E-05)	1	0.31 (2.1E-05)	0.30 (4.47E-15)	0.19 (4.19E-05)	0.37 (1.42E-19)	0.39 (1.20E-16)
MEF2A	0.28 (9.57E-08)	0.19 (0.0043)	0.30 (4.47E-15)	0.45 (4.84E-13)	1	-0.01 (0.04)	0.57 (7.34E-53)	0.48 (1.17E-29)
MEF2C	0.25 (3.80E-06)	0.33 (5.82E-10)	0.37 (1.42E-19)	0.58 (3.06E-13)	0.57 (7.34E-53)	0.07 (2.99E-08)	1	0.52 (3.95E-28)

Table 3 Spearman correlation between the genes with a prognostic value and all the other class IIa HDACs and MEF2 members.

3.1.4 A TUMOR MICROENVIRONMENT GENE SIGNATURE IS VARIABLY EXPRESSED AND CORRELATED WITH HDACs IIA AND MEF2 GENES

We examined a published transcriptomic signature derived from a wide set of tumor microenvironment (TME) associated cell populations (Becht et al., 2016), since from our previously published data class IIa HDACs, and in particular HDAC7, emerged to have an inhibitory function towards inflammatory signals derived from the extracellular environment. To evaluate the expression level of these signatures in CRC patients, their gene expression data were represented as the median level of expression of their constituting genes. These signatures were expressed at different levels in TCGA samples (Fig.4A), with the fibroblasts signature (composed of only 8 genes, half encoding collagen chains) having a higher level in comparison to most immune-associated gene signatures, the only exception being represented by the monocytic lineage. Secondly, the fibroblasts signature expression increased significantly with the progression of the tumor, particularly from Stage I to III, although only after Z-score normalization of the expression data as these data are available as well (Fig.4B). The increasing expression trend, in this case RSEM values, was observed also for the endothelial cells signature, comprising 27 genes (Fig.4C). Moreover, we found that high expression of the fibroblasts signature is associated with significantly reduced overall survival (Fig.4D, logrank test, $P< 0.05$) and disease-free survival (using the optimal cut-point approach, Fig.4E). The Spearman correlation between each signature and HDACs IIa/MEF2 genes, as depicted in the heatmap (Fig.5), showed on average a positive correlation, especially with respect to MEF2 transcription factors (excluding MEF2B), HDAC7 and HDAC9. The strongest overall

correlation with the TME signatures was found for MEF2C, particularly with the stromal component (fibroblasts and endothelial cells) and the monocytic lineage. This result agrees with the increasing MEF2C expression from initial to later stages of CRC and, additionally, it seems to confirm the fact that the MEF2 TFs are subject to a regulation exerted by the tumor microenvironment (Di Giorgio et al., 2018).

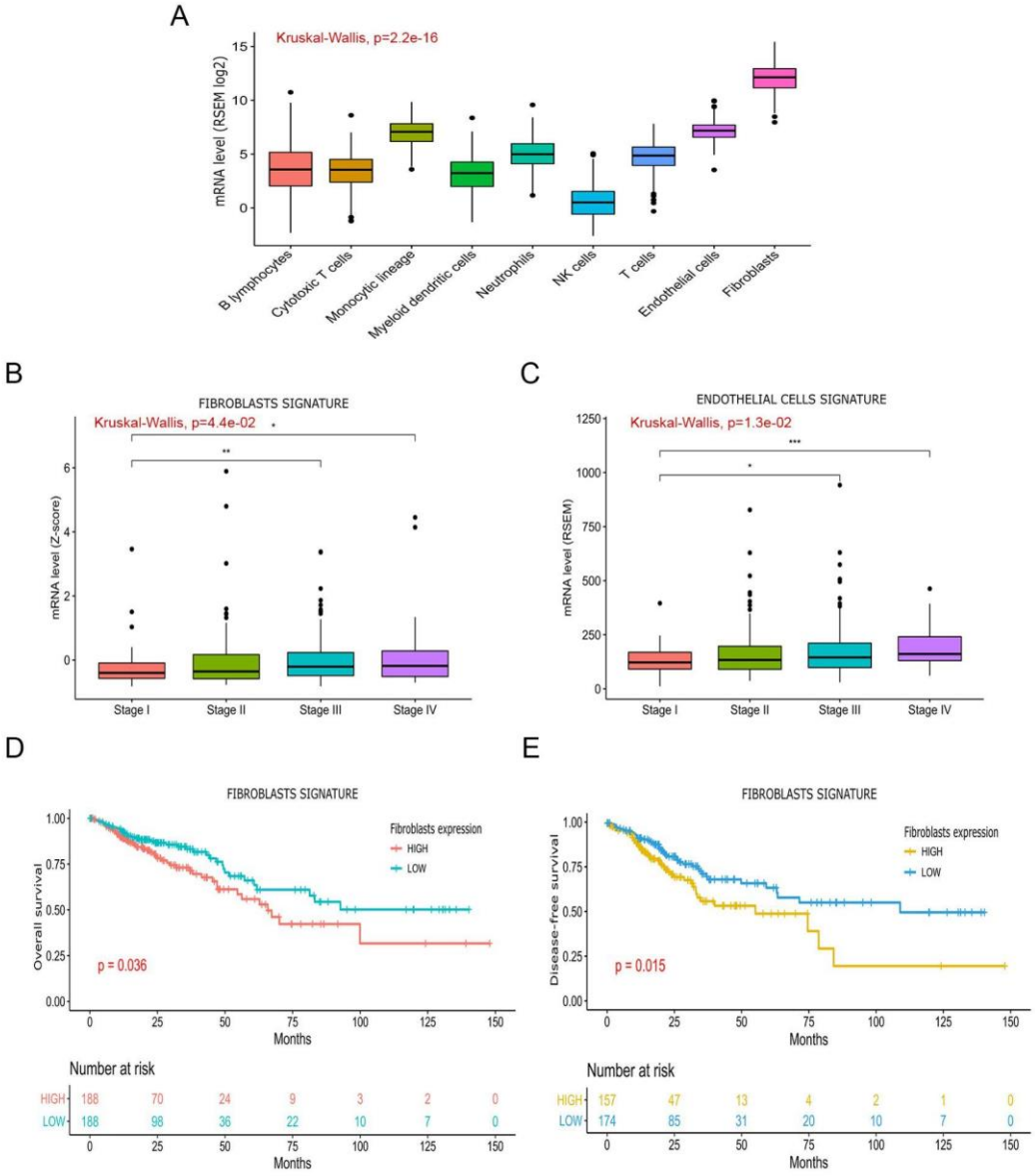


Fig.4 (A) Boxplots representing the immune and stromal gene signatures mRNA expression levels (RSEM log2), summarized by the respective median expression values (Kruskal-Wallis global test, P<0.05). (B) Boxplots showing the increasing gene expression trend (Z-scores) of the fibroblasts signature with tumor stage progression

(Kruskal-Wallis global test, P<0.05, Wilcoxon paired test, *P<0.05, **P<0.01). (C) Boxplots representing the increasing gene expression trend (RSEM) of the endothelial cells signature with tumor stage progression (Kruskal-Wallis global test, P<0.05, Wilcoxon paired test, *P<0.05, ***P<0.005). Kaplan-Meier plots (overall survival (D) and disease-free survival (E)) of patients having high and low mRNA expression levels of the fibroblasts signature (logrank test, P<0.05).

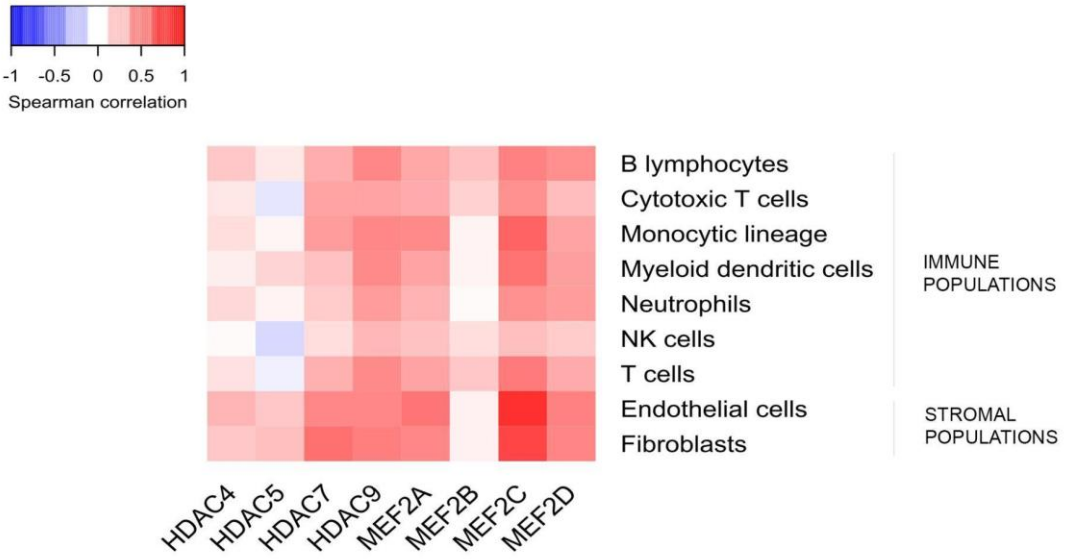


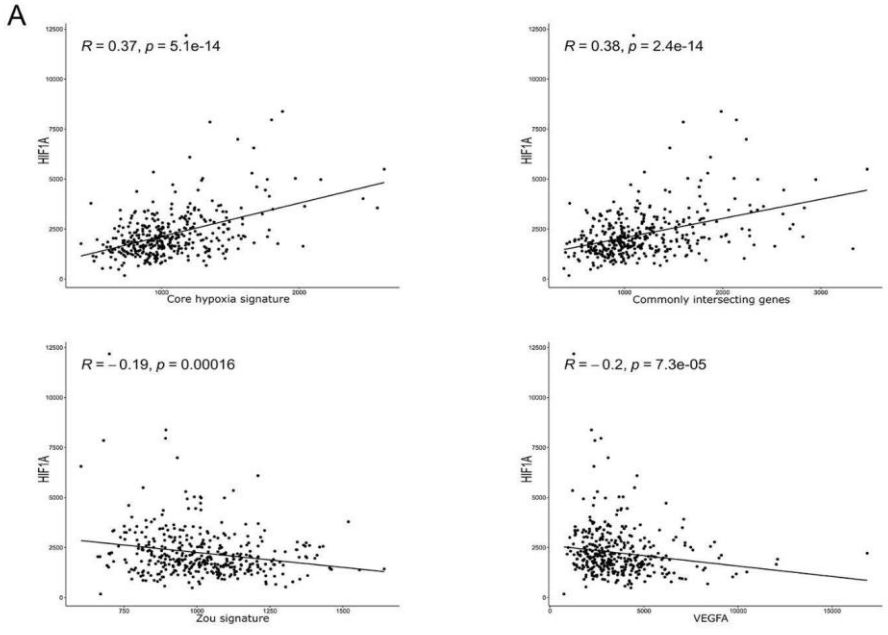
Fig.5 Heatmap depicting the Spearman correlation coefficients calculated between the mRNA expression levels (RSEM) of the HDACs IIa and MEF2 genes and the median expression levels of the Immune and Stromal gene signatures.

3.1.5 EVALUATION OF THE TUMORIGENIC ROLE OF SEVERAL HYPOXIA GENE SIGNATURES IN CRC

To include the contribution of hypoxia, we selected heterogeneous sets of published gene signatures associated with this biological process, to extend the knowledge about the role played by the TME in CRC tumorigenesis and metastasis, and additionally the potential influence exerted by HDACs IIa and MEF2 genes. The chosen hypoxia gene signatures (HGS) include different aspects associated with tumor hypoxia, from the metabolic adaptation to *in vivo* metastasis; we named "core" signature the union of the 20 most frequently occurring genes in 32 published signatures and 10 genes commonly included in 10 clinically-derived and validated hypoxia gene signatures (Lukovic et al., 2019). Secondly, we considered another

global signature resulting from an *in silico* meta-analysis on published microarray datasets, common to a wide range of cancer cell types (Lendahl et al., 2009). The last two signatures are associated with *in vivo* distant metastasis (Hu et al., 2009) or derived from transcriptomic data of CRC patients (Zou et al., 2019), respectively. We also considered for this analysis the union of all gene sets and the commonly intersecting genes, while the only gene in common between all the signatures appeared to be VEGF-A.

Using again the median expression level of each gene belonging to the signatures as a proxy of the signature itself, the core signature appeared as the one most strongly associated with HIF1α expression, followed by the commonly intersecting genes; by contrast, curiously, VEGF-A and the CRC-derived signature are slightly negatively correlated with the hypoxia master regulator (Fig.6A). The expression level of the CRC specific signature significantly decreases during tumor progression, particularly from stage II onwards, while VEGF-A and the metastasis-associated signature exhibit the opposite trend (Fig.6B). As expected, the latter is significantly up-regulated in patients with lymphovascular invasion and metastasis, and VEGF-A expression, too, is higher in the case of metastatic patients (Fig.6C).



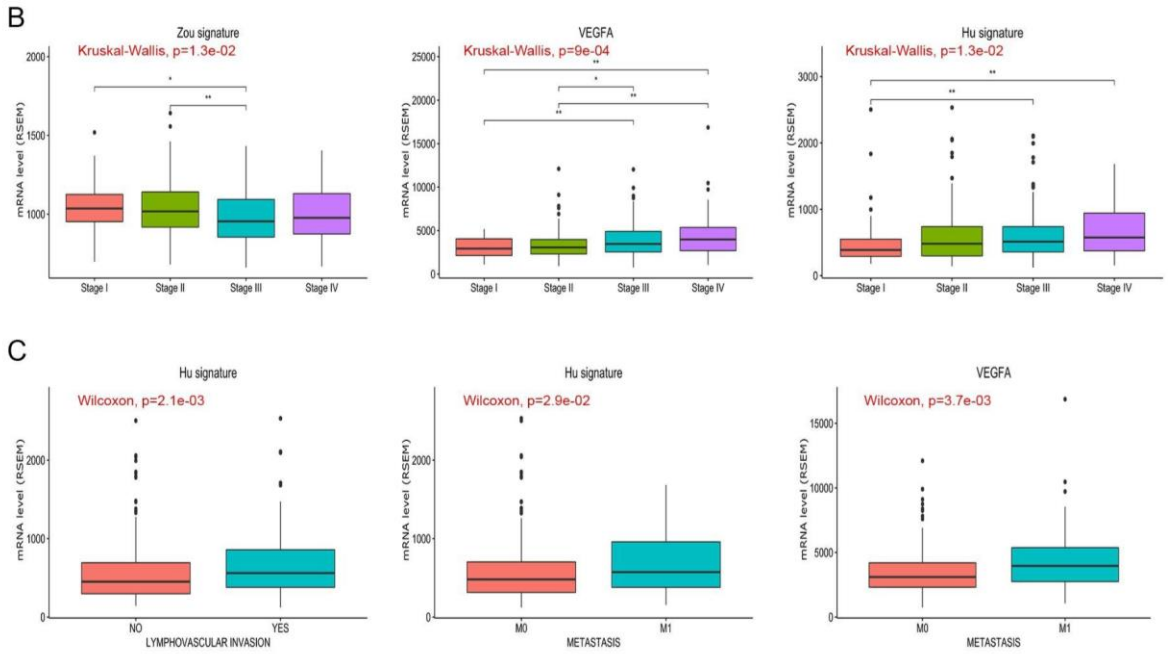


Fig.6 (A) Scatterplots depicting the Spearman correlation between HIF1α and, from upper left to lower right, the core hypoxia signature, the commonly intersecting genes, the Zou signature and VEGF-A. (B) Boxplots showing, from left to right, the Zou signature, VEGF-A and the Hu signature gene expression levels stratified by CRC tumor stage (Kruskal-Wallis global test, P<0.05, Wilcoxon paired test, *P<0.05, **P<0.01). (C) Boxplots representing the gene expression levels of the Hu signature between CRC patients with or without lymphovascular invasion and metastasis (in this case also of VEGF-A) (M0 indicates absence of metastasis, M1 indicates presence of metastasis) (Wilcoxon paired test, P<0.05).

Moreover, when considering the fraction of genome altered (FGA), almost all gene signatures, excluding the one by Lendahl, show a significant lower expression in patients with higher FGA; on the contrary, VEGF-A displays the opposite trend (Fig.7), a singular result considering that hypoxia is known to induce genome instability in cancer.

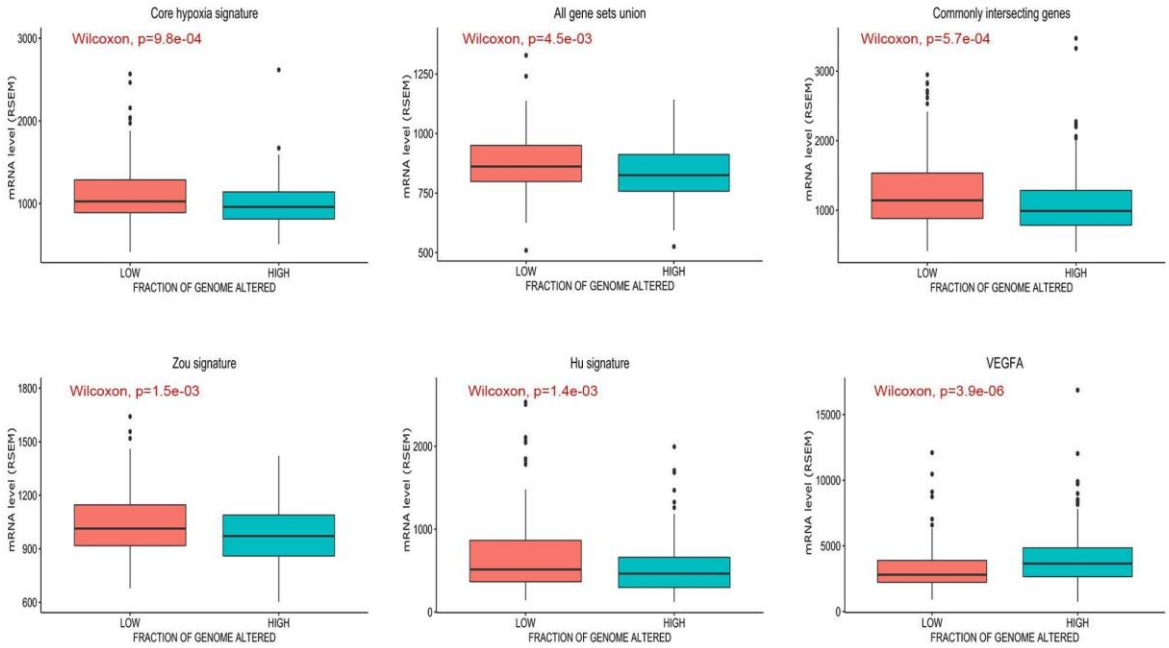


Fig.7 Boxplots representing the gene expression levels of the core hypoxia signature, the union of the gene sets, the commonly intersecting genes, Zou and Hu signatures, and VEGF-A in CRC patients with a high versus low fraction of genome altered (FGA) (Wilcoxon paired test, P<0.05).

We then evaluated how patients' survival is influenced by the high or low expression of the hypoxia signatures. Interestingly, the group of patients with high VEGF-A expression live (almost) significantly lower (P=0.05) and are also associated with significantly higher rates of recurrence or progression of the disease (Fig.8A). Instead, only when considering patients at early stages (I and II) the Lendahl signature and the union of all hypoxia-related genes appear as significant prognostic biomarkers, since the respective high expression groups are associated with both lower overall survival and disease/recurrence-free survival (Fig.8B-C).

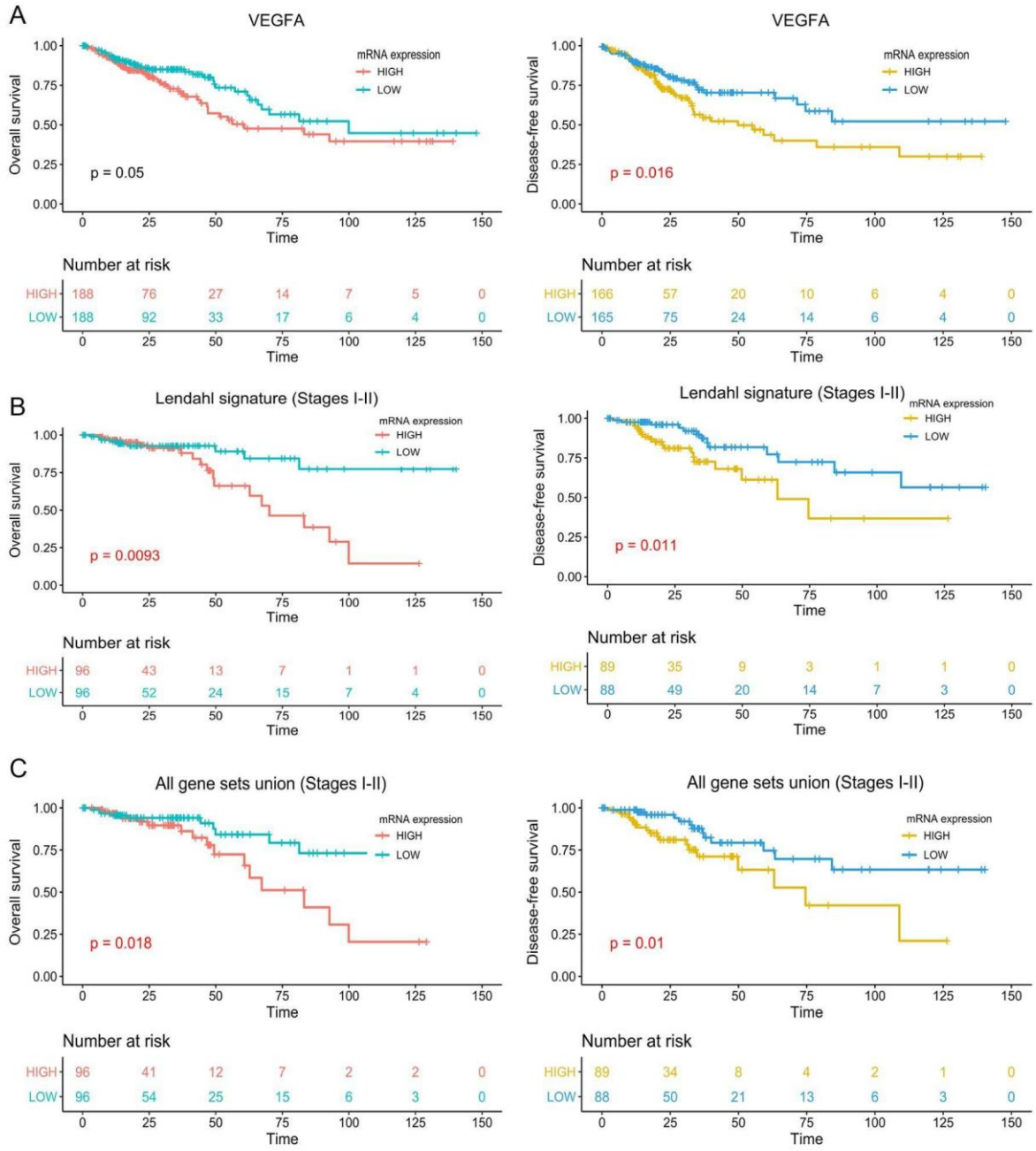


Fig.8 Survival analysis showing the Kaplan-Meier plots of overall survival (left) and disease/progression-free survival (right) of CRC patients having high and low mRNA expression levels (RSEM) of VEGF-A (A), the Lendahl signature (B) and the union of all hypoxia signatures (C) (logrank test, P<0.05).

Finally, we investigated the correlation between the HGS, the TME, HDACs IIa and MEF2 TFs. From the heatmaps, it is possible to observe that in the case of the previously analyzed tumor microenvironment signatures, there is an overall positive correlation between the TME

populations and HIF1α, particularly strong in the case of the stromal component, the monocytic lineage and the neutrophils; by contrast, with VEGF-A the overall correlation is slightly negative and noteworthy, also between the stromal cell populations and the CRC-derived HGS. Moreover, the same signature is particularly negatively associated with HDACs IIa and MEF2 transcriptional regulators, with the only exception being represented by MEF2B. From the same heatmap it is evident that both HDAC4 and HDAC5 appear negatively correlated with all the examined HGS; additionally, with HIF1α the most positive correlation is found for MEF2A and MEF2C, and among HDACs IIa with HDAC7 (Fig.9A-B). Stratifying for early and late tumor stages (Fig.9C), there is an increase in the correlation between MEF2C and HIF1α in the more advanced stages, while with respect to HDAC5 the correlation remains substantially negative. Interestingly, HDAC7 appears negatively associated with the CRC-derived hypoxia signature only at late stages, and overall, there is a slight increase in the correlation between HDACs IIa and MEF2 transcriptional regulators and the core hypoxia signature, with the exclusion of MEF2B, the lowest expressed in CRC patients among the MEF2 family, which is negatively associated with the global signatures, the commonly intersecting genes and also the union of all genes. These results pinpoint an important role of tumor hypoxia in CRC that in some cases, for example with VEGF-A and the union of all gene sets, significantly affects patients' survival. Globally, the interaction between hypoxia, class IIa histone deacetylases and MEF2 deserves further investigation.

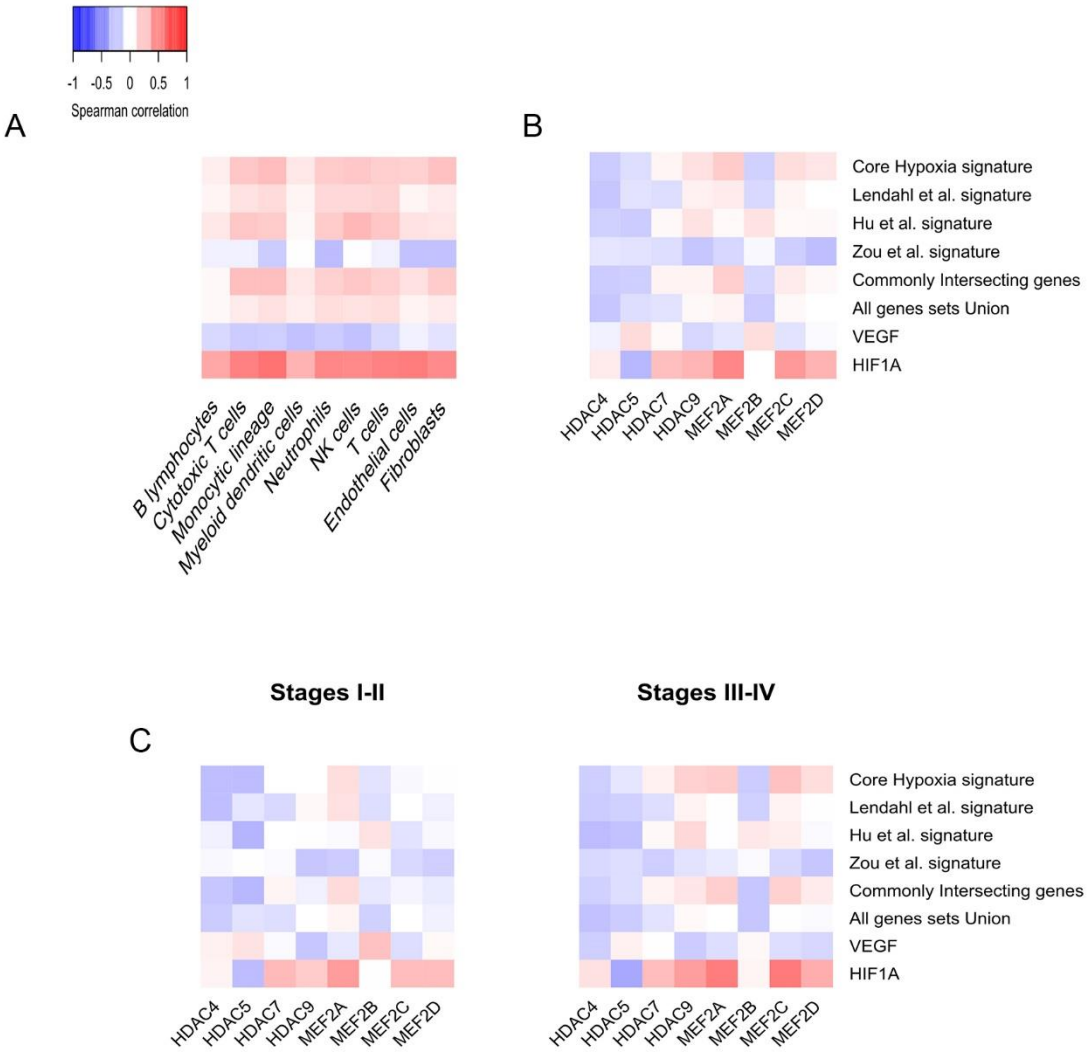


Fig.9 Heatmaps depicting the Spearman correlation coefficients calculated between the median mRNA expression levels (RSEM) of the hypoxia gene signatures and the TME signatures (A), the HDACs IIa and MEF2s genes in all CRC patients (B) and stratified in early (I and II) and late (III and IV) tumor stages (C).

3.2 HISTONE H3K27AC CHANGES IN THE SK-UT-1 LMS CELL LINE

3.2.1 H3K27AC PATTERNS IN PROXIMITY OF HDACs IIA BINDING

Histone acetylation changes reflect the differential binding of transcriptional regulators and, ultimately, induce variations in gene expression. Firstly, the raw counting of enriched peaks identified in the ChIP-seq experiments performed in the SK-UT-1 cell line (derived from a high grade uterine leiomyosarcoma) showed a predominance of HDAC4 binding regions (n=7732) with respect to HDAC9 (n=1257). A substantial proportion of HDAC9 peaks (n=1010) overlaps HDAC4's by at least 1 bp, highlighting the presence of shared supervised networks (Fig.10A-B). Interestingly, when comparing the global normalized acetylation levels in each context, H3K27ac peaks signal rises expectedly in HDAC4 KO but decreases sharply in HDAC9 KO. This trend, however, is not observed when looking at H3K27ac peaks which fall within 5-10 kb from HDACs IIa binding sites; in this case, there is an increase of acetylation peaks, particularly steep for HDAC4 KO (Fig.10C). Using the 'GREAT' tool (from Bejerano Lab, Stanford University) to identify the genes close to the HDAC binding regions associated with a decrease of H3K27ac within 5 kb, it is clear that such regions are found mostly distally (more than 50 kb) from the TSS of said genes, either upstream or downstream (Fig.10D). This is naturally true also for HDAC4 and HDAC9 intersecting regions, but in this case about half (45%) of the regions do not appear associated with any gene (data not shown).

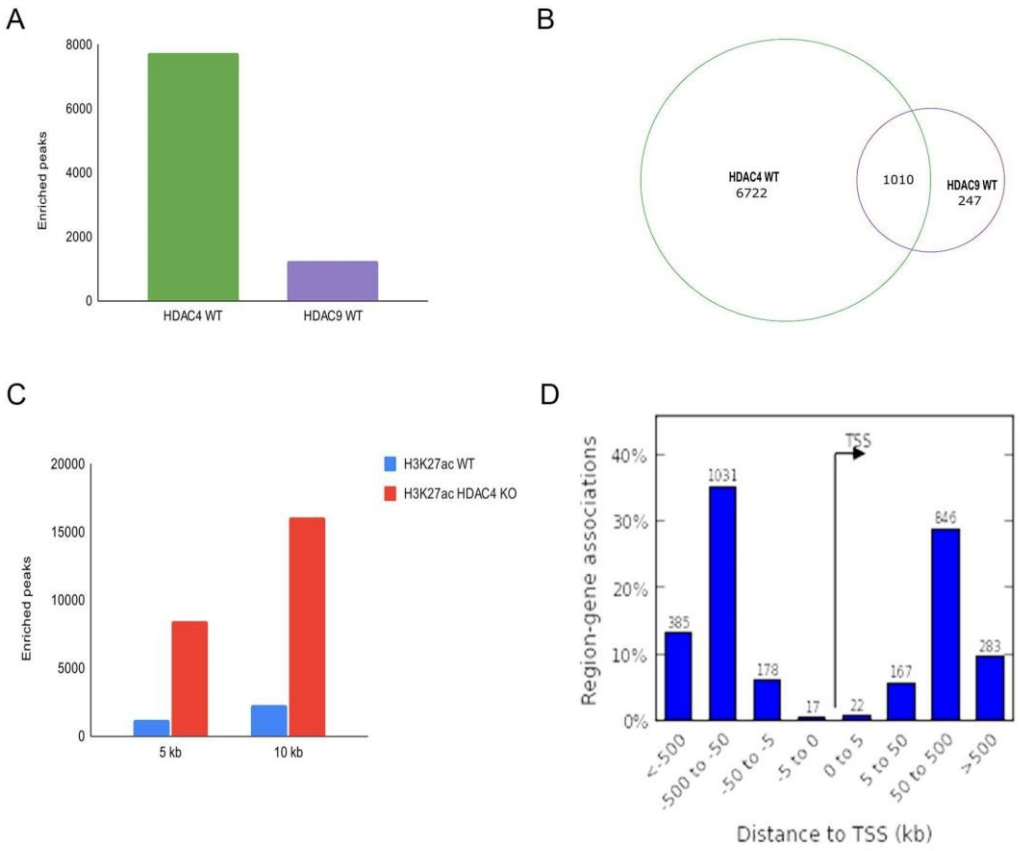


Fig.10 (A) Number of HDAC4 and HDAC9 enriched peaks in SK-UT-1 cells and (B) the number of HDAC4/9 peaks overlapping by at least 1 bp. (C) Comparison of the number of H3K27ac enriched peaks in WT versus HDAC4 KO conditions falling within 5-10 kb from HDAC4 binding. (D) Position, relative to the TSS of the closest annotated genes, of HDAC4 binding associated with a drop of H3K27ac within 5 kb.

Moreover, among the HDAC4/9 common regions, there are n=253 HDAC4 peaks associated with a decrease of H3K27ac levels within 5 kb which show, as reported from 'GREAT' results page, a Gene Ontology (GO) enrichment relative to cellular component (used to describe the cellular compartment in which a gene product performs its function), for the Invadopodium membrane, a structure of the plasma membrane correlated with the prime stages of cancer cell invasion through the ECM (Table 4).

	H3K27ac KO enriched peaks within 5 kb	Enriched peaks	Genes	GO enrichment
HDAC4	↑	253	82	Invadopodium membrane
	↓	14	16	-
HDAC9	↑	117	63	Heart development
	↓	119	65	Multicellular organism growth

Table 4 Enriched HDAC4 and HDAC9 peaks after the knockout characterized by an increase or decrease of H3K27ac enriched peaks within 5 kb, the corresponding annotated genes and the Gene Ontology (GO) functional analysis enriched terms.

As a step forward, we decided to restrict the intersection to only 1 bp (rather than 5 kb), to account for a mechanistically driven direct relationship between HDAC binding and H3K27ac variation. In the case of HDAC4, the binding peaks with an increase/decrease of H3K27ac within 1 bp comprise about half of those found within 5 kb. The 'GREAT' annotation of the different regions associated with an opposite acetylation trend showed a unique set of n=271 genes, when accounting for their overlap. We chose to focus the attention on HDAC4 regions with a proximal decrease in H3K27ac peaks, as we never investigated this subcategory before. Among the set of 38 genes, we could quantify the overall and average acetylation ChIP signal from the sum of the associated H3K27ac peaks. In absolute terms, the peak intensities display a great variability among the genes, with DACH1, MZT1 and ZIC1 showing the greatest drop of acetylation in HDAC4 KO SK-UT-1 cells (Fig.11A). This is true also when examining the average H3K27ac peak intensities, although the genes exhibit less variability due to the different number of peaks per gene. Additionally, the distance of HDAC4 binding with respect to each gene's TSS varies too, according to the results found by 'GREAT', up to a maximum distance set at 1 Mb; with respect to this, some genes display HDAC4 binding relatively close, such as ARID5B, MSRB3 and DPH6 (Fig.11B). ARID5B, MSRB3 and PDK4 also show a

MEF2 binding motif (Fig.11C). However, the exact mechanism responsible for the decrease of H3K27ac following the HDAC4 loss in LMS cells will require additional validation; given the very low catalytic activity of class IIa HDACs, however, it appears plausible to hypothesize that this class is also able to form coactivator complexes.

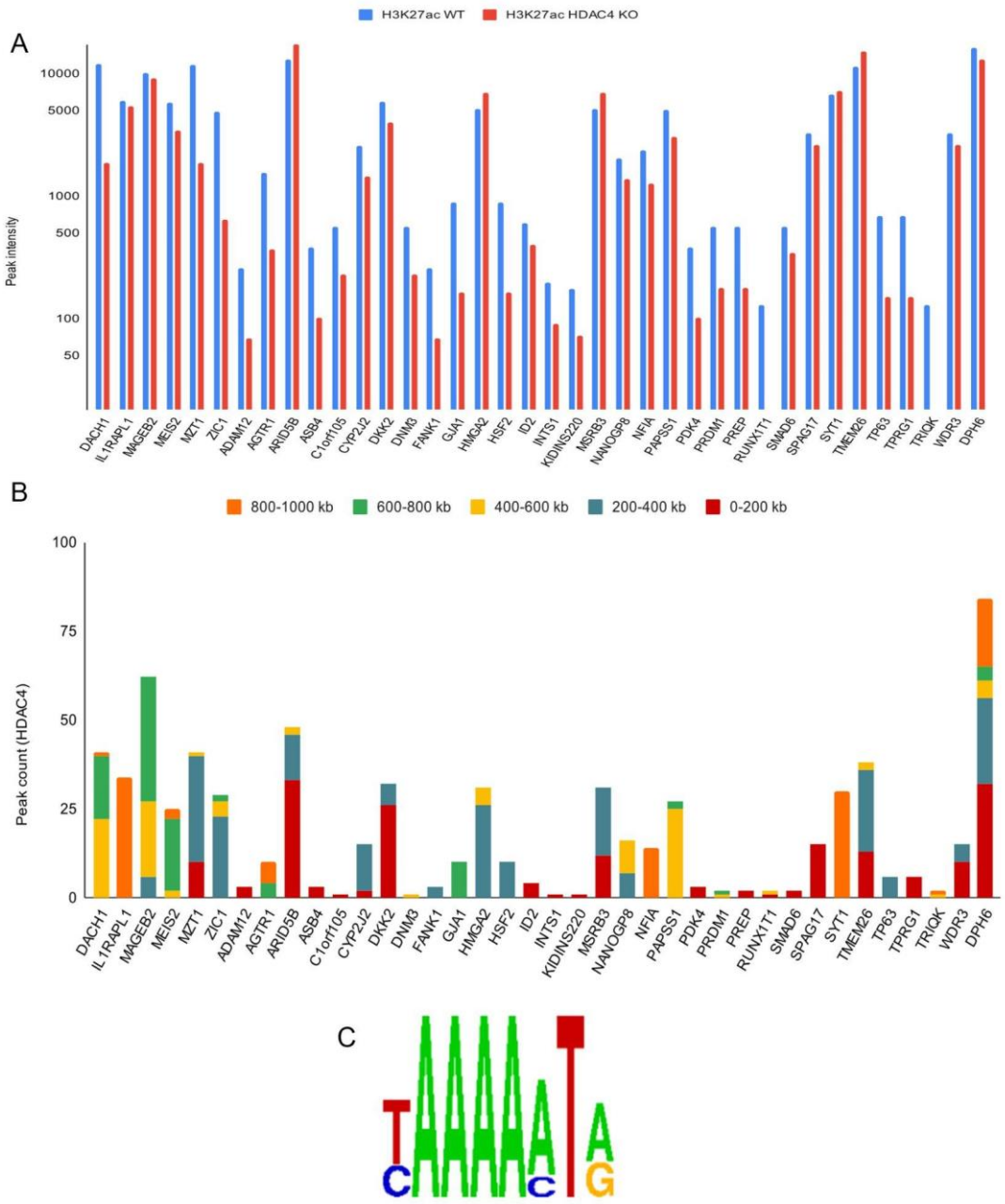


Fig.11 (A) H3K27ac signal intensity for the 38 genes characterized by a decrease in H3K27ac enriched peaks within 1 bp from HDAC4 binding. (B) Distance of HDAC4 binding from the TSS of the 38 genes. (C) Motif discovery performed using the PDK4

genomic coordinates with Trawler (https://trawler.erc.monash.edu.au/) and default options, showing the MEF2 binding motif.

3.2.2 SUPPORTING EVIDENCE FOR A ROLE OF GJA1 (CONNEXIN43) IN LMS

We decided to analyze some available microarray datasets (GSE132569; GSE94416) from SK-UT-1 cells as a means of validating the previous findings; moreover, an evaluation of the gene expression profile in addition to clinical data was also performed in soft tissue sarcomas samples from TCGA for further confirmation. As a starting point, we focused on 26 genes selected based on the number of HDAC4 and H3K27ac associated peaks (at least two) and on the presence of a decreasing acetylation signal in HDAC4 KO. For all but one gene (NANOGP8, a retrogene) the expression data could be retrieved from the microarray datasets; twelve of them showed a significant log2 fold change, among which GJA1 (encoding Connexin-43) exhibited by far the stronger magnitude of repression in HDAC4 (and also HDAC9) KO cells (575-fold decrease), followed by CYP2J2 and NFIA. Gene expression was also evaluated in microarray data in the presence of MEF2 silencing through shRNA, showing in some cases opposing trends of expression in comparison to HDAC4 KO, although rarely statistically significant, except in the case of DACH1, strongly repressed after MEF2 shutdown and also counting the highest number of associated HDAC4 peaks (n=41) (Fig.12A). Then, we looked at the gene expression of these genes in TCGA clinical samples, analyzing the dataset "adult soft tissue sarcomas, TCGA (Cell, 2017)" from the TCGA Research Network (found at https://www.cbioportal.org/datasets), constituted of n=206 soft tissue sarcoma (STS) samples, including 80 leiomyosarcomas (LMS). As shown in Fig.12B, the expression of the twelve genes displays a degree of variability, with GJA1, PDK4 and NFIA having the highest expression in LMS patients.

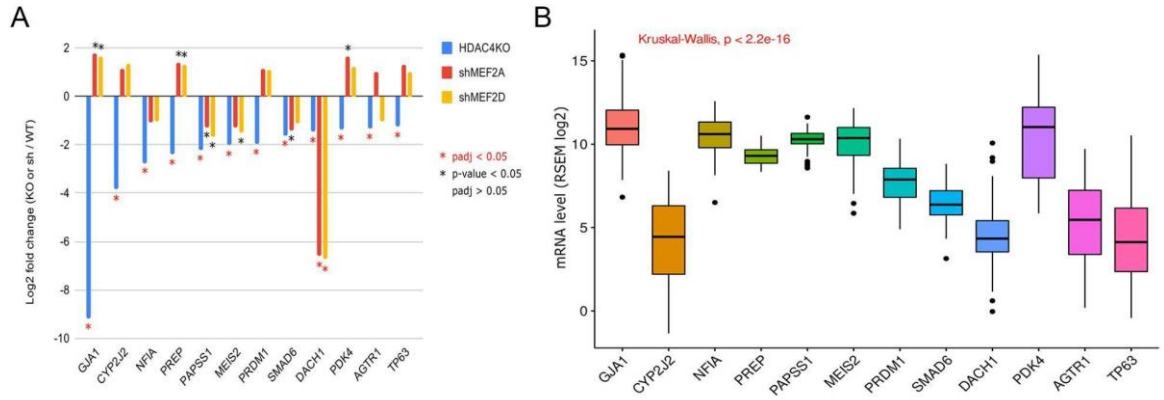


Fig.12 (A) Log2 fold change of gene expression in microarray datasets (*P<0.05, *p-value adjusted for multiple testing). (B) Gene expression level (RSEM log2) in TCGA LMS samples (Kruskal-Wallis global test, P<0.05).

Concerning GJA1, its expression significantly rises with tumor aggressiveness (FNCLCC grading system) in STS but not specifically in LMS, presumably due to the limited sample size and loss of statistical power (Fig.13A); the expression of MZT1 and PREP also rise with tumor aggressiveness in LMS and STS patients, respectively (data not shown); by contrast, the PDK4 gene expression significantly falls with an increase in both STS and LMS aggressiveness (Fig.13B, only LMS shown), and this is occurring also for MEIS2 and MSRB3, in STS and LMS patients respectively (data not shown). Furthermore, we compared the expression of GJA1 between tumor and matched normal tissue and, although the extremely limited sample size (n=2) does not allow to reach statistical significance, the expression in LMS is about 3 times higher than normal tissue in the same patient (Fig.13C). Additionally, the high expression of GJA1 appears correlated with a poor prognosis (disease-free survival), in terms of a significantly higher occurrence of relapse or recurrence, both in STS and in the LMS subgroup, adopting the optimal cut-point approach and interrupting the K-M curves before the sample size becomes too small (Fig.13D-E).

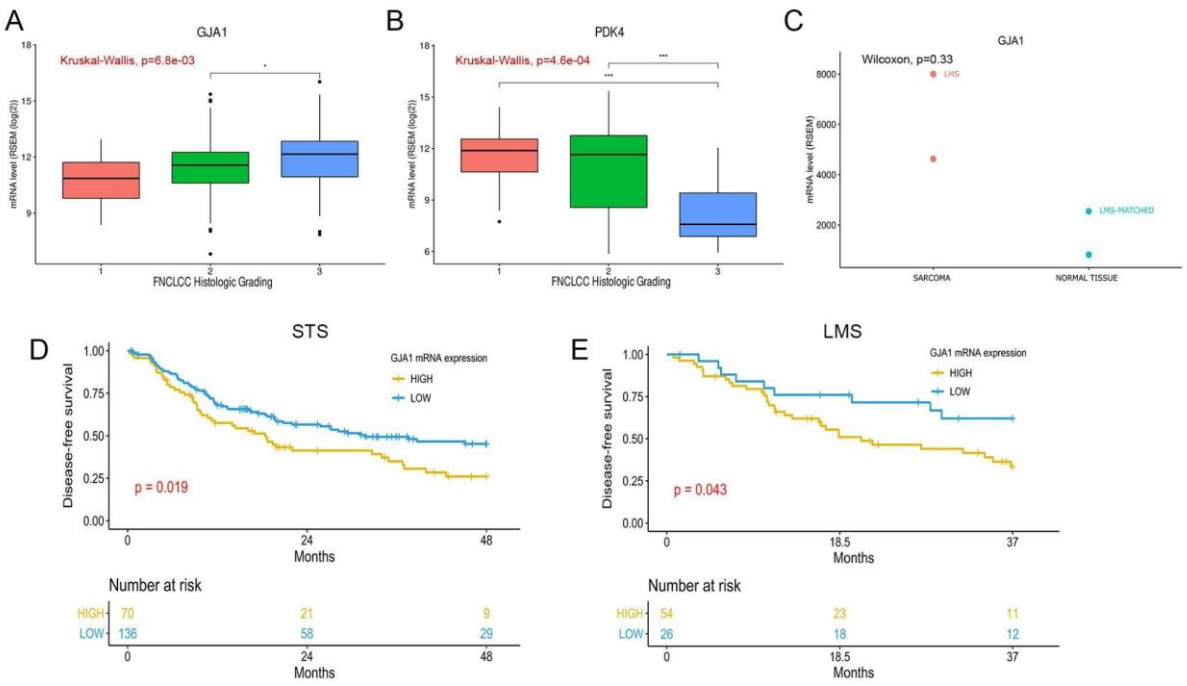


Fig.13 (A) GJA1 mRNA expression (RSEM log2) in TCGA STS patients based on the FNCLCC histologic grade (Kruskal-Wallis global test, P<0.05, Wilcoxon paired test, *P<0.05). (B) PDK4 mRNA expression (RSEM log2) in TCGA LMS patients based on the FNCLCC histologic grade (Kruskal-Wallis global test, P<0.05, Wilcoxon paired test, ***P<0.005). (C) Comparison of the GJA1 gene expression (RSEM) between tumoral and matched normal tissue (Wilcoxon paired test). (F) Survival analysis (disease-free survival) in STS (D) and LMS (E) TCGA patients having high and low mRNA expression levels (RSEM) of GJA1, adopting the optimal cut-point approach (logrank test, P<0.05).

Another interesting point, which can be linked to the pleiotropic actions of connexins, is the influence exerted by GJA1 (Cx43) to the tumor microenvironment (TME); in fact, the GJA1's most co-expressed genes *in vivo* appear associated with functional terms related to the interaction with the extracellular matrix (ECM), like the Syndecan interaction. Finally, a high expression of GJA1 transcript is also associated with a greater leukocyte infiltration (i.e. leukocyte fraction) both in STS and LMS (Fig.14), pointing to a possible facilitating function played by Connexin-43, although the underlying mechanism remains to be elucidated. However, from the analysis of the RNA-seq data available from the TCGA STS dataset, the GJA1 transcript mRNA level appeared poorly correlated with both class IIa HDACs and MEF2

TFs, except for a weak-moderate negative correlation with MEF2A and MEF2D (data not shown), a result which is concordant with the microarray shMEF2 dataset previously reported.

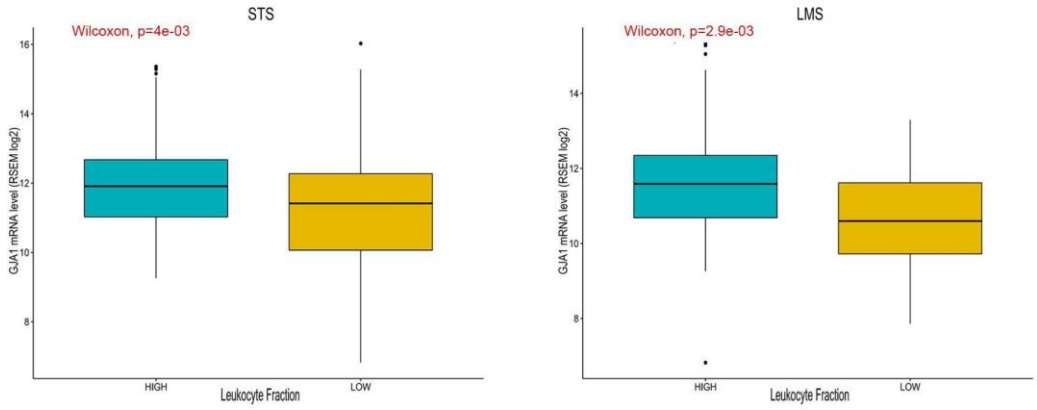


Fig.14 Boxplots showing the GJA1 mRNA expression (RSEM log2) in TCGA STS (left) and LMS (right) patients, based on the stratification in high versus low Leukocyte Fraction levels (Wilcoxon paired test, $P<0.05$).

3.2.3 MAPPING THE GJA1 LOCUS

As described in the previous results, we observed the occurrence of a region enriched in HDAC4 peaks in an interval between 600 and 800 kb from the GJA1's TSS (Fig.15, extension of approximately 1 Mb downstream to the gene). Secondly, the diminishing of the H3K27ac signal happening in HDAC4 KO condition with respect to WT is also evident, both downstream to the gene and in proximity, confirming the previous data. Thirdly, there is also evidence of a moderate MEF2D binding around GJA1's TSS, hinting at the presence of HDACs-bound MEF2, thus forming a complex capable of repressing GJA1 expression. To corroborate this hypothesis, a useful technique to determine in an unbiased manner the presence of distal genomic interactions is represented by Hi-C, an extension of 3C (chromosome conformation capture), which displays a contact matrix, in this case derived from the IMR-90 fibroblast cell line, which shows to some extent the presence of a chromatin loop connecting the distant genome locations, on one side the HDAC4 binding and the reduction of acetylation and on the other the locus of GJA1 with a putative MEF2D binding. Overall, GJA1 (Cx43) appears to be part of the network under the supervision of HDACs IIa, although with a long-range mechanism of regulation and also an unusual positive control evident after the decrease of histone H3K27ac

mark due to the absence of HDAC4 in LMS cells and supported also by the strong down-regulation in microarray. Taken together, these data point to Cx43 as an interesting candidate player in LMS, albeit the interest in connexins is not something new given the complex roles in physiological conditions and in different tumors and their niche.

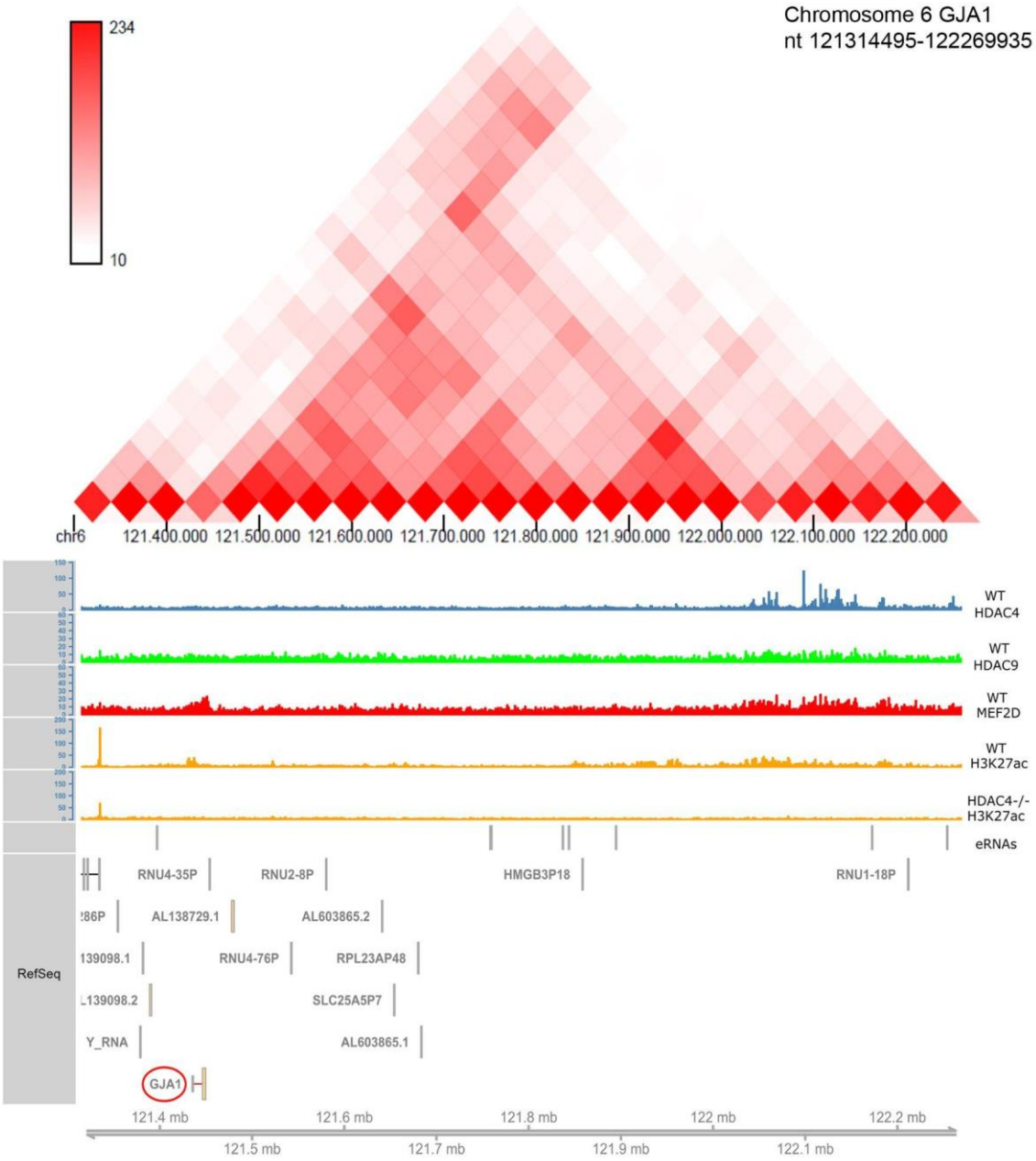


Fig.15 Locus of the GJA1 gene with HDAC4, HDAC9, MEF2D, H3K27ac (WT), H3K27ac (HDAC4 KO) and eRNA tracks (retrieved from SlideBase (Human Enhancers Selector), data

from the FANTOM5 project). In addition, the Hi-C contact heatmap (derived from the IMR-90 cell line) is shown above. Genomic coordinates: chromosome 6, 121314495-122269935.

3.3 EFFECTS OF NKL54 TREATMENT IN SK-UT-1 LMS CELLS

Recently, our laboratory started studying the effects of inhibiting the HDACs IIa-MEF2 axis through a series of small molecules, with the rationale of disrupting the repressive action exerted by histone deacetylases and reactivating the MEF2 differentiation transcriptional program, which could counteract the uncontrolled proliferation in soft tissue sarcomas. In particular, the class of pimeloylanilide o-aminoanilide (PAOA) derivatives emerged as the more promising in terms of inhibiting LMS cells proliferation *in vitro*: from the original compound BML210, originally characterized as a pan-HDAC inhibitor (Savickiene et al., 2006), the corresponding fluorinated analog NKL54 was studied for its capability to bind the hydrophobic groove of MEF2. However, the mechanism of action of this compound needs to be re-assessed, as the previously reported displacement of the MEF2-HDACs IIa interaction has not been confirmed by our more recent results (Minisini, Di Giorgio et al., 2022).

3.3.1 CHIP-SEQ ENRICHED PEAKS ANALYSIS FOLLOWING NKL54 TREATMENT

Chromatin was immunoprecipitated with HDAC4, HDAC9 and MEF2D antibodies, in addition to the H3K27ac mark, after NKL54 treatment. As observed also from *in vitro* experiments, NKL54 can trigger a change in the expression level of these factors; in fact, the number of enriched peaks dramatically increased for MEF2D (from 2214 to 7763) and diminished for HDAC4 (from 7732 to 2153) and, less significantly, also in the case of HDAC9 (from 1257 to 570) (Fig.16A). Furthermore, the overall number of acetylation peaks increases, with approximately thirty-thousand additional H3K27ac enriched peaks identified after the treatment, likely indicating a more accessible chromatin status induced by the compound. Nonetheless, the pairwise intersection (at least 1 bp) of enriched peaks between untreated and treated conditions showed that a substantial part of H3K27ac marks are overlapping (57743). Similarly, only 500 and 201 HDAC4 and HDAC9 peaks, respectively, were unique for the NKL54 treatment condition, while 1653 and 369 were in common with the untreated condition. Regarding MEF2D, after the NKL54 treatment most peaks (n=6878) appear to be unique and only 885 overlapped WT peaks (Fig.15B).

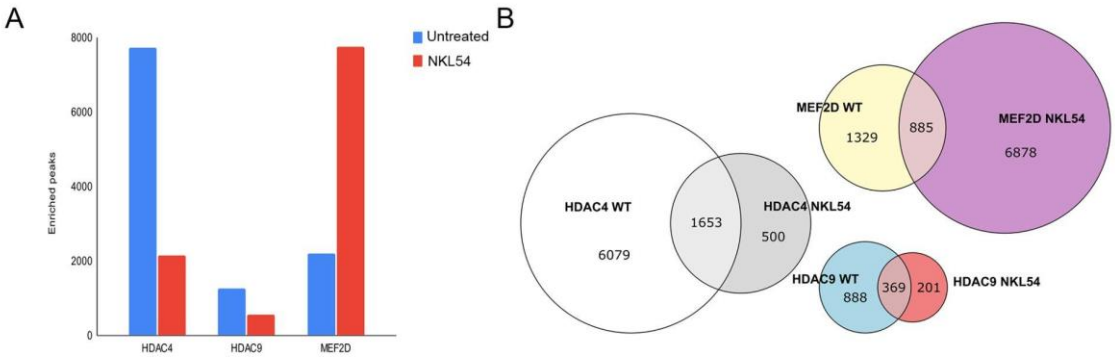


Fig.16 (A) Comparison of the number of HDAC4, HDAC9 and MEF2D enriched peaks between untreated (WT) and NKL54 treatment conditions. (B) Venn diagrams showing, in untreated and NKL54-treated SK-UT-1 cells, the 1 bp overlap of enriched peaks for HDAC4, HDAC9 and MEF2D.

Then, a quantitative genomic annotation of the enriched peaks was performed through the R package 'ChIPseeker'. Proportionally, after NKL54 treatment there is a sharp decrease of distal binding by HDACs IIa, while the proximal binding is very poor and limitedly affected by the treatment, with the only exception of intronic binding for HDAC4 which appears reduced after the treatment (Fig.17A-B). In the case of MEF2D, the binding is more distributed across the genomic compartments and following the NKL54 treatment the binding pattern displays an opposite behavior compared to HDACs IIa, with an increased binding particularly at the promoter level but also distally (Fig.17C); however, this could be simply attributed to the variation of expression levels, particularly for HDACs IIa, as observed *in vitro*. In relation to acetylation, there is a slightly higher prevalence of H3K27ac marks both at the proximal and distal regions (Fig.17D).

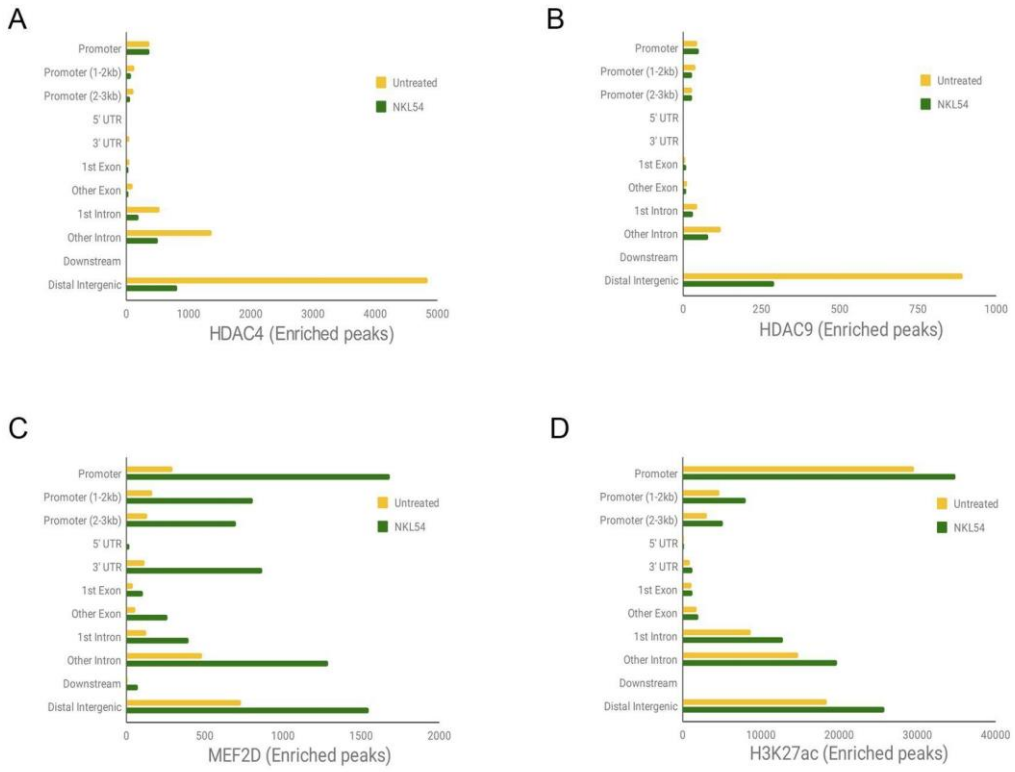


Fig.17 Comparative genomic mapping through ChIPseeker of the HDAC4 (A), HDAC9 (B), MEF2D (C) and H3K27ac (D) enriched peaks in untreated cells and following NKL54 treatment.

In order to assess how the NKL54 treatment functionally affects the binding of class IIa HDACs and MEF2 TFs, the GO analysis of the peak-associated genes revealed an enrichment for histone H4K20 demethylase activity of HDAC4-associated genes after NKL54 treatment, in contrast with HDAC4 WT peaks which are mostly correlated with sequence-specific DNA binding of RNApol II-associated regulatory regions. On the contrary, no enriched terms were found for HDAC9-associated genes. Finally, the analysis of MEF2D peaks-associated genes, either WT or NKL54-treated, defined chromatin binding and, specifically for treated cells, protein kinase binding as enriched functional terms, with no significantly enriched terms related to regulatory sequence binding (Fig.18).

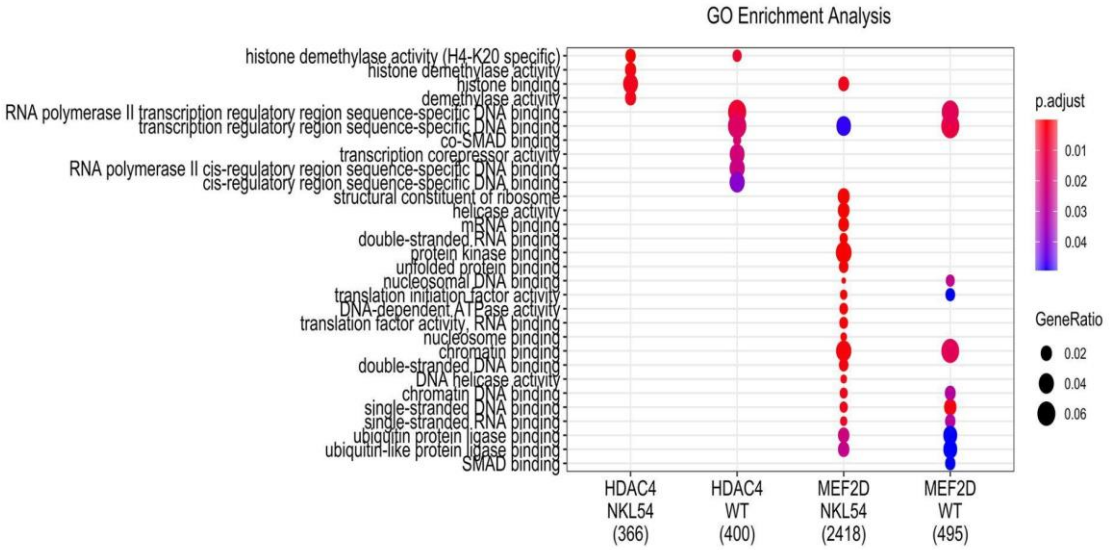


Fig.18 Gene Ontology (GO) enrichment analysis of ChIP-seq enriched peaks. The number of annotated genes is indicated in brackets.

3.3.2 HISTONE ACETYLATION AND HDACs IIA AND MEF2 BINDING PATTERNS AROUND THE TSS

With the aim to further dissect the available ChIP-seq data to discover how the histone acetylation and the binding of class IIa HDACs and MEF2s localize at the TSS, the 'Deeptools' suite was used to integrate and compare the raw ChIP-seq data as well as the enriched peaks, either at a whole-genome scale or in specific genomic regions of interest. To support the previous results regarding the overlap of the enriched peaks, and in order to extend this information to the raw ChIP-seq signal, the Spearman correlation between the reads coverage of H3K27ac marks in the untreated (WT) versus NKL54-treated conditions was calculated, showing a very strong correlation (ρ=0.93) of the two signals, highlighting an overall high similarity in reads coverage that could have been overlooked with an analysis restricted to enriched peaks alone. On the contrary, the MEF2D signal before treatment emerged as the most divergent in terms of a lower correlation with MEF2D after treatment (ρ=0.53) and also with the other conditions; this result may highlight the overall rearrangement of MEF2D binding across the genome as a consequence of NKL54 exposure. Similarly, the HDAC4 signal in both conditions (ρ=0.76) clustered separately, showing an overall lower Spearman correlation with the remaining samples. HDAC9, on the contrary, displayed a strong correlation between

untreated and treated samples (ρ=0.88), as well as with respect to H3K27ac reads coverage (Fig.19).

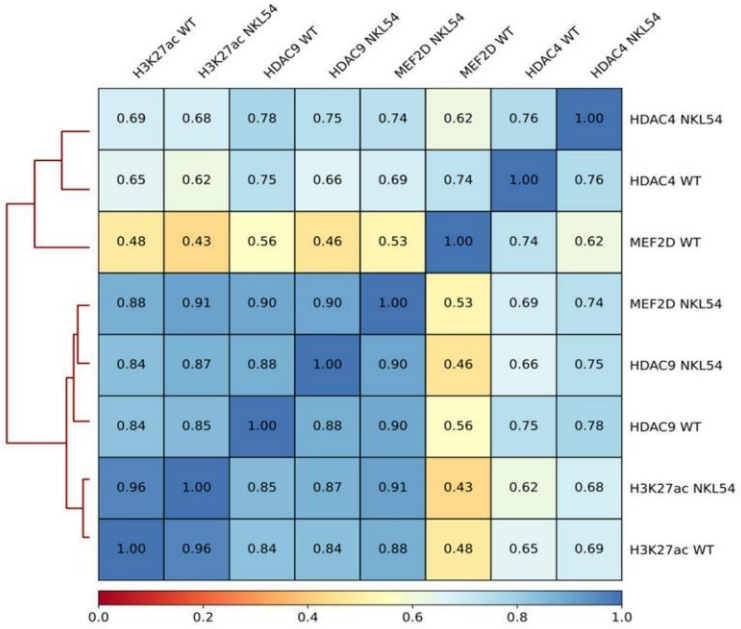


Fig.19 Clustered correlogram representing the Spearman correlation coefficients calculated between the ChIP-seq raw signal of every sample for each condition (untreated (WT) vs NKL54-treated).

Looking at the acetylation signal within 3 kb of the TSS of all the RefSeq curated genes, as well as that of the RNA-seq up- and down-regulated genes (Fig.20), a decrease of this histone mark in proximity of the TSS is quite evident, even if it must be admitted that, in the case of the up-regulated genes, the average signal is about halved compared to the down-regulated genes; nonetheless, for the former group of genes this trend is quite unexpected, due to the up-regulation following the NKL54 treatment.

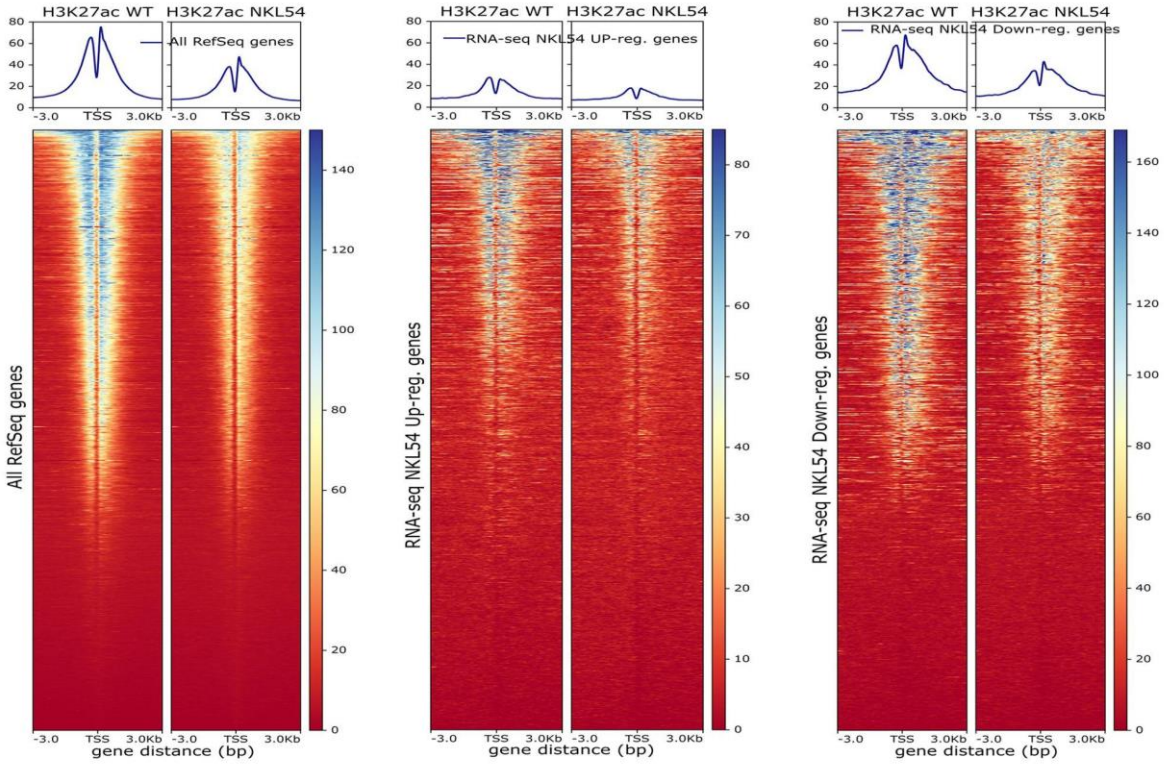


Fig.20 Heatmaps displaying the WT and NKL54-treated ChIP-seq H3K27ac raw signal within 3 kb from the TSS of all the annotated RefSeq genes (left), of the RNA-seq up-regulated genes (mid) and of the down-regulated ones (right).

Moreover, regarding HDAC4 binding proximally to the TSS of these gene sets, this HDAC binds preferentially the down-regulated genes and a quite small percentage of all genes (Fig.21A); however, the proximal binding is not significantly affected by the inhibitor, similarly to HDAC9 (data not shown), evidence that also emerged when evaluating the enriched peaks. Interestingly, NKL54 leads to a strong rise in MEF2D binding at the TSS, in particular increasing the fraction of bound RNA-seq down-regulated genes (Fig.21B). Among this subset, 73 genes show a substantial increase (fold change > 2) in MEF2D signal at the TSS following NKL54 treatment, and these MEF2D binding events are newly formed; on the other hand, only limited class IIa HDACs binding (specifically HDAC4) is found in a window of 3 kb within the TSS. Only two out of these 73 genes displayed new MEF2D TSS binding and HDAC4 proximal positioning following the treatment, namely GYS1 (glycogen synthase I) and SSU72 (RNA polymerase II CTD phosphatase), the latter involved in chromatin regulation and RNA processing.

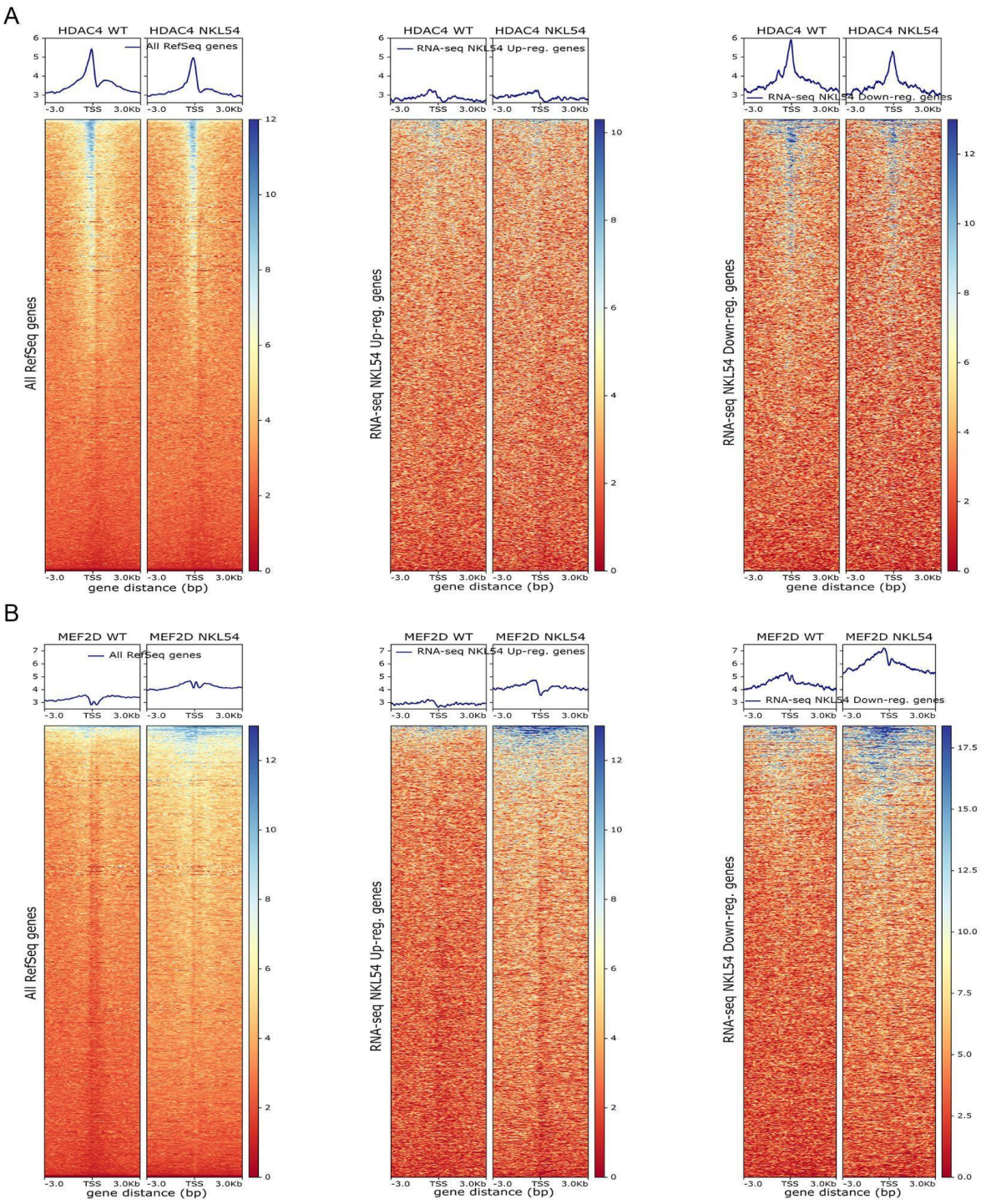


Fig.21 Heatmaps displaying the ChIP-seq HDAC4 (A) and MEF2D (B) raw signals around the TSS of all the annotated RefSeq genes (left), of the NKL54 RNA-seq up-regulated genes (mid) and of the down-regulated ones (right).

Finally, since our *in vitro* data demonstrated that NKL54 is able to trigger a change in the MEF2s and class IIa HDACs protein levels in LMS cells, the loci for these genes were visualized to compare the H3K27ac levels in untreated versus treated conditions. In the case of HDAC4 (Fig.22A), there is an enrichment of the H3K27ac mark 2 Mb upstream of the TSS, which gets reduced after NKL54 treatment; these data could highlight the presence of a distal regulatory mechanism mediated, for instance, by a super-enhancer (SE) that could be responsible for the decreasing levels of HDAC4 in LMS cells seen after NKL54 exposure. In the case of MEF2D (Fig.22B), the plot similarly depicts a slight down-modulation of the acetylation signal upon treatment in a large region mostly upstream to the TSS but extending also beyond the gene, despite its (modest) up-regulation *in vitro* at a later stage. Overall, NKL54 exerts a generalized decrease of the H3K27ac histone mark at promoter regions, in contrast to the evidence of an increase of the corresponding enriched peaks which rise in every genomic district (both proximal and distal). On the other hand, the binding of class IIa HDACs at promoter regions is not strongly affected by the inhibitor, which on the contrary triggers MEF2D binding, particularly of the down-regulated genes, likely as a consequence of the NKL54 inhibitory activity on class IIa HDACs.

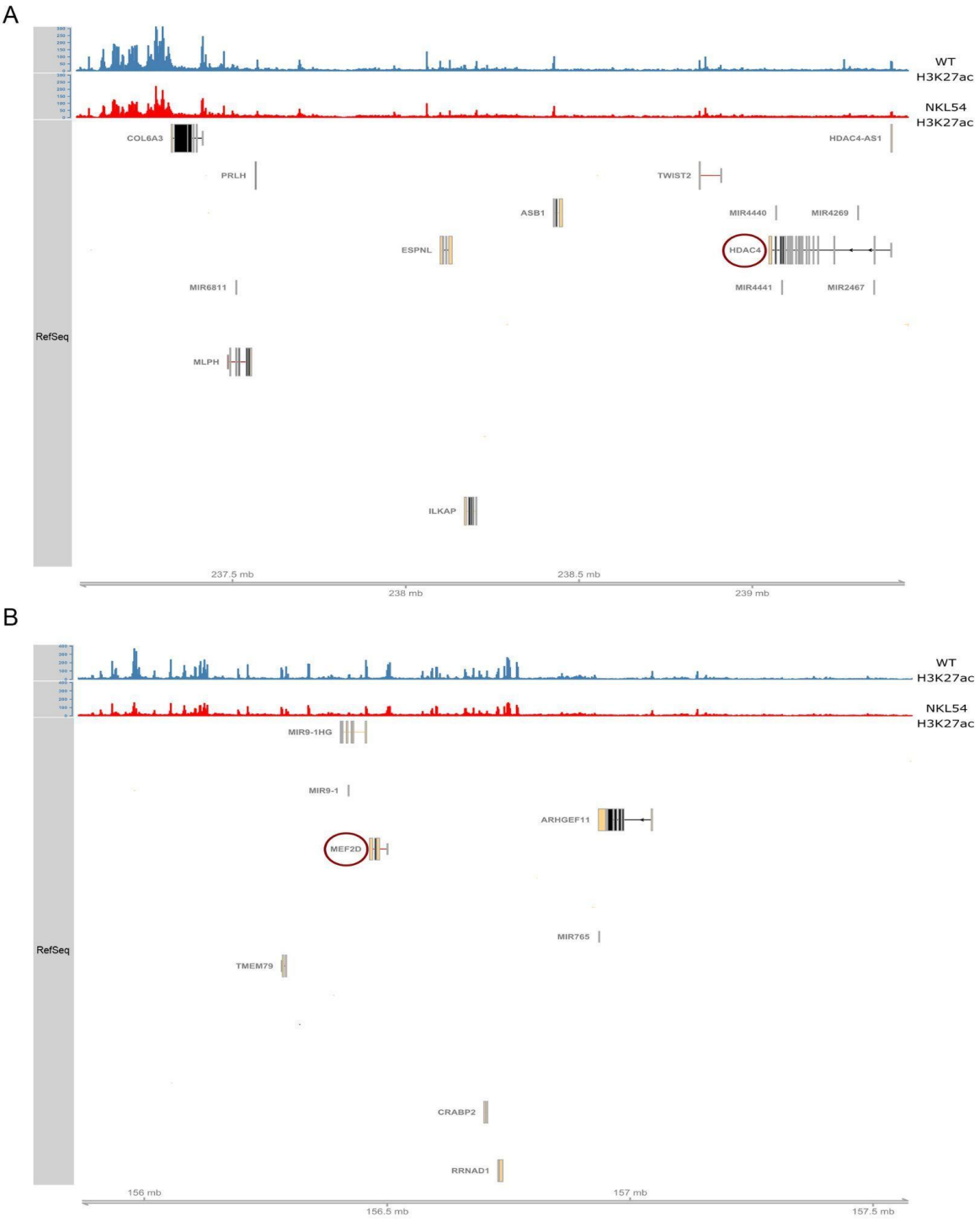


Fig.22 (A) Genomic locus of the HDAC4 gene showing the H3K27ac (WT) and H3K27ac (NKL54) tracks (scale 0-300), extending approximately 2 Mb upstream from its TSS (chr2: 237048168-239451654). (B) Genomic locus of the MEF2D gene showing the H3K27ac (WT) and H3K27ac (NKL54) tracks (scale 0-400) (chr1: 155863727-157580840).

Discussion

The present doctoral research project has been devoted to expanding the knowledge of class IIa HDACs functions in diverse tumoral contexts (CRC and LMS), with an emphasis on the traditional axis with the MEF2 transcriptional factors and, additionally, with the inclusion of the histone mark H3K27ac as a means of considering chromatin accessibility and, thus, the expression changes occurring to several putative target genes. Moreover, the interest in HDAC inhibitors, such as the small molecule NKL54, gained our attention as well: in fact, epigenetic drugs are becoming a promising therapeutic strategy in oncology and, in specific cases, they have been already clinically validated, as is the case for SAHA.

Regarding the investigation of class IIa HDACs and MEF2 TFs in colorectal cancer patients, very limited data are currently available, making this analysis pioneering. Admittedly, we could assess that the alteration of the HDACs IIa and MEF2s genes in the CRC samples from TCGA is rare, nonetheless their expression often rises with tumoral stage; in addition, for all these genes, a significant difference in expression levels exists between the normal colon tissue and matched tumoral tissue, with only HDAC7 displaying an increased expression in CRC, highlighting the fact that at least this HDAC member could exert a prominent role in this context, although not neglecting the fact that the down-regulation of the other HDACs and all MEF2s could also imply a biological effect. This is supported by the negative prognostic value associated with high HDAC7 expression in CRC patients and, furthermore, by not yet published *in vitro* data by our research group obtained from two CRC cell lines obtained from the same patient, in which HDAC7 expression appears thrice higher in the latter than the former and, moreover, its over-expression induces an increased proliferation rate in both cell lines. This evidence, together with a recent paper reporting the MEF2A potential contribution to CRC aggressiveness, points to the HDACs IIa-MEF2 axis as an interesting candidate for additional investigation in CRC pathogenesis. About the contribution of the TME to cancer progression and metastasis, which cannot be neglected anymore, through the analysis of the available gene signatures that are representative of the various cell populations constituting the tumor niche, we could delineate a more detailed picture of the patients' genomic profile, from a relative quantification of the signatures' abundance to the definition of their prognostic value. The fibroblasts signature was particularly interesting: its expression in CRC patients, in fact, indicates a relative enrichment of this stromal component in the CRC samples, likely in the form of cancer-associated fibroblasts (CAFs) and, secondly, it also appears to be a good prognostic biomarker which could serve for further validation. Lastly, together with the

endothelial cells signature, the former displayed a quite strong correlation with MEF2C, which itself appeared associated with lower overall survival. However, a study of the TME cannot be considered complete without the inclusion of the hypoxic phenotype as a crucial factor in determining the metastatic potential. Thus, different signatures were selected to reflect the underlying biological mechanisms contributing to building up a hypoxic environment (for instance the metabolic switch and the angiogenesis), in addition to signatures reflecting the *in vivo* metastatic process and, also, the specificity for the CRC transcriptomic fingerprint. These signatures appeared variably correlated with HIF1α, the master regulator of hypoxia, and noteworthy, VEGF-A in particular was negatively associated with HIF1α in CRC samples, while displaying a predictable increase in expression in patients with metastasis. Moreover, another evidence came from the observation that almost all the evaluated signatures, considering the median expression of their constituting genes, present an opposite trend compared to the fraction of genome altered, a parameter representing the degree of genome instability. All these results will deserve a biological assessment, to corroborate these bioinformatics associations, which are intrinsically correlative but can hardly lead to a causative relationship. On the other hand, the discovery of a prognostic significance for these signatures, as it is the case with VEGF-A for DFS or with the Lendhal signature for CRC patients at early stages, could provide a more direct clinical benefit to these patients, for example by predicting their disease outcome. Taken together, TCGA datasets offer a wide spectrum of analytic opportunities, which can ramp up from few genes to whole signatures composed of several gene panels; supported by *in vitro* data, this analysis has also highlighted a new, unreported role for HDACs IIa, and HDAC7 in particular, in colorectal cancer.

Another fundamental research activity in which I was involved during my PhD consisted in analyzing ChIP-seq datasets generated from the SK-UT-1 cell line, derived from a high grade uterine leiomyosarcoma. As an important and frequently assessed histone mark, H3K27ac was included in the analyses as a focal point to infer the open chromatin status and, thus, the gene expression activation. Nonetheless, we chose to take a divergent approach focusing on those genomic regions associated with a drop of the acetylation levels following the depletion of HDAC4 or HDAC9, the most important HDACs IIa members in the context of LMS. In the examined datasets, these two HDACs differ substantially in the number of genome binding regions (in favor of HDAC4), even if a relatively large subset of these regions is in common. After the HDAC4 and HDAC9 knockout by means of the CRISPR-Cas9 technology, an increase in the number of H3K27ac peaks within 5-10 kb from the HDAC4/9 peaks could be observed; this evidence appears physiologic, as HDACs IIa are mostly found at co-repressor

complexes on the genome. Another characteristic emerged when mapping the peak-associated genes, that is the prevalence of distal binding (over 50 kb) of HDACs IIa from their TSS, either up-stream or down-stream, this confirming previously published data regarding the putative mechanism of action mediated by super-enhancer complexes, which postulates the existence of long-range interactions between the HDACs IIa complexes and more proximal regions. Restricting the analysis solely to the H3K27ac peaks overlapping HDAC4 binding aimed at establishing additional proof of the causative relationship between these two events. Indeed, a substantial proportion of acetylation events were found to localize within 1 bp from the HDAC4 peaks, and the subsequent annotation retrieved a subset of 38 genes characterized by a proximal decrease in the number of TSS-associated H3K27ac peaks. The highest drop in the acetylation signal was observed for DACH1 (Dachshund family transcription factor-1), a TF implicated in organogenesis, in the inhibition of the TGF-beta signaling and, as it seems, also in periosteal osteogenic sarcoma. Given the availability of microarray data generated from the same cell line, harboring either the KO of HDAC4 or the silencing of MEF2, a validation was made possible to assess the *in vitro* expression level of the identified genes. The data from shMEF2A/D were not very consistent, although it highlighted DACH1 as the most repressed gene. On the contrary, out of the 26 identified genes satisfying the filtering conditions, 12 were found repressed at a statistically significant level in HDAC4 KO cells; in particular, GJA1 is an interesting gene, somehow strongly repressed in HDAC4-depleted cells and coding for Connexin-43, one of the most important members of the connexin family, well-known for their structural function in forming gap junctions but also involved in several signaling circuits (e.g. the regulation of cell migration, the transcriptional repression and TGF-beta signaling, by interacting with tubulin). Connexins are either associated with improved or poor prognosis, depending on the tumor type. In this analysis, considering the SARC-TCGA dataset comprising 80 LMS samples among more than 200 STS, we could confirm the bad prognostic potential of this gene, when highly expressed, in terms of DFS: in fact, its expression increases in higher grade STS samples. Another interesting finding is the three-fold up-regulation of GJA1 in LMS compared to the matched normal tissue, although this result would need an increase of the available samples to be definitive. Furthermore, considering the importance of connexins in promoting heterotypic cell interactions in the tumor milieu, we could also observe a higher GJA1 expression in patients with increased levels of leukocyte infiltration, although this result may appear in contrast with the previous ones in terms of prognostic outcome. Finally, aiming at the evaluation of a putative distal regulatory mechanism mediated by class IIa HDACs, the Hi-C data confirmed the presence of a three-dimensional chromatin loop connecting the TSS

of the GJA1 gene with HDAC4 binding, evidence corroborated by the decreasing intensity of the H3K27ac signal in the same region after the KO of this HDACs IIa member.

A major point of interest for our laboratory, furthermore, resides in targeting the class IIa HDAC-MEF2 axis in leiomyosarcoma to hamper its exacerbated proliferation, ultimately bringing to light an effective epigenetic cure for this rare and aggressive tumor. For this purpose, a series of small compounds have been tested, with a particular interest in the PAOA derivative NKL54. Recent data from our research group have demonstrated that this molecule displays a potent inhibitory effect on proliferation by inducing apoptosis of LMS cells; NKL54 could also act as a HDAC inhibitor, although being ineffective in displacing the interaction between HDACs IIa and MEF2s. The ChIP-seq experiments revealed that NKL54 treatment at 14 hours led to a substantial redistribution of the genome occupancy by these transcriptional regulators: the majority of HDAC4 binding was removed while MEF2D binding was greatly enhanced; the mechanism involved remains to be investigated and it is only partially explained by the induced variation in protein levels, which however occurs from a later time point. A distinction, however, must be made between the ChIP-seq continuous signal, represented by the raw reads coverage across the genome, and on the other hand the enriched peaks, derived from the former after the peak calling; this contraposition leads to two distinct analyses and, generally, the continuous signal provides a more realistic picture of the underlying biology. When looking at the genomic mapping of the enriched peaks, it appears that the NKL54 effect on the downsizing of class IIa HDACs peaks mostly happens at distal regions, with respect to the TSS of the annotated genes, while the binding of TSS-proximal regions is not affected but still remains quantitatively limited; by contrast, the inhibitor induces an overall increase of MEF2D positioning at virtually all genomic regions, most importantly at new promoters where it could act as a transcriptional activator thanks to the absence of the repressive influence of class IIa HDACs. Looking at the distribution of the ChIP-seq signal following NKL54 treatment, MEF2D binding particularly increases at the TSS of the RNA-seq down-regulated genes, while a relatively small subset of these genes appears to be bound by class IIa HDACs. With regard to the H3K27ac mark, after the treatment there is a surge in enriched peaks (more than thirty-thousand) which are equally distributed among proximal, intronic and distal intergenic regions, showing however a substantial overlap with the WT condition; nonetheless, in proximity of the TSS the histone mark visibly decreases after the treatment, an unexpected result particularly for the RNA-seq up-regulated genes, although the acetylation signal is visibly lower in this group of genes, implying that the genes up-regulated by the inhibitor are generally weakly expressed in comparison to the down-regulated genes. Overall, NKL54 may represent

a promising small molecule with therapeutic potential in LMS; such a benefit could derive from the inhibition of class IIa HDACs and the overall enhancement of MEF2 transcriptional activity, although the underlying biological mechanism appears still unclear. In conclusion, the data presented in this doctoral thesis may help expanding the horizon about the oncogenic potential of the class IIa HDAC-MEF2 axis in two different tumoral contexts, CRC and LMS, also considering the contribution given by the tumor microenvironment, pointing towards a better understanding of epigenetic-driven carcinogenesis.

Material and Methods

5.1 RSTUDIO AND R PACKAGES

RStudio is an Integrated Development Environment (IDE) for R, a programming language for statistical computing and graphics, widely used in the scientific community. The RStudio version used was 1.4.1106 (RStudio Team, 2020), while different versions of R have been used, from 3.5.1 to 4.1.0 (R Core Team, 2021).

Depending on the task, different R packages were used; for survival analysis, the packages "survival" (v. 3.2.11, Therneau, 2021) and "survminer" (v. 0.4.9, Kassambara et al., 2021) were applied. For ChIP-seq peak annotation, "ChIPseeker" (v. 1.28.3, Yu G, Wang L, He Q, 2015) was used, while for the visualization of ChIP signal "Gviz" (v. 1.36.1, Hahne and Ivanek, 2016). For TCGA data query and download from cBioPortal, "cgdsr" (v. 1.3.0, Jacobsen and Luna, 2019) was applied. For the retrieval of genomic data from Ensembl, biomaRt was used (v. 2.48.0, Durinck et al., 2009). To manipulate data, the main packages used have been "tidyr" (v. 1.1.3, Wickham et al., 2019), "reshape2" (v. 1.4.4, Wickham, 2007) and "magrittr" (v. 2.0.1, Milton Bache et al., 2020). For the generation of most plots, the packages "ggplot2" (v. 3.3.3, Wickham, 2016), "ggpubr" (v. 0.4.0, Kassambara, 2020), "ggcorrplot" (v. 0.1.3, Kassambara, 2019) and "gplots" (v. 3.1.1, Warnes et al., 2015) were exploited.

5.7 STATISTICAL ANALYSIS

The examined sample populations were not normally distributed: for this reason, and in order to reach a greater robustness in the presence of outliers, non-parametric statistical tests were applied. For the survival analysis, the log-rank test was used to assess a statistically significant difference in the outcome between two Kaplan-Meier curves ($P<0.05$). The difference between two populations' means was calculated with the Wilcoxon paired test ($P<0.05$), while with more than two samples the Kruskal-Wallis test was used ($P<0.05$). For correlation analyses, the Spearman coefficient was calculated.

5.4 ACCESSING TCGA DATA THROUGH CBIOPORTAL

cBioPortal is an integrated public platform for accessing, analyzing, and downloading large cancer genomics data, initially developed at the Memorial Sloan Kettering Cancer Center, and currently maintained by a multi-institutional team. This portal includes all the major cancer types grouped by organ/tissue of origin, and for each cancer type multiple datasets are available

in addition to PanCancer studies, providing data about genetic alterations, gene expression data and clinical information. The user can choose the study, the kind of information to take in account and additionally insert gene/s of interest. Data were last accessed in February 2021, downloading data from the following studies: colorectal adenocarcinoma (TCGA, Firehose Legacy) and adult soft tissue sarcomas (TCGA, Cell 2017).

5.2 SURVIVAL ANALYSIS, HAZARD RATIO AND MULTIVARIATE ANALYSIS

Statistics is an essential component of medicine, allowing researchers to reach sound conclusions about their data; survival analysis, which is regarded as part of inferential statistics, has many applications not limited to the biosciences, but most of its value derives from its extensive use in cancer clinical trials, a medical field which has greatly contributed to design and develop survival methods. These are based on few key concepts: the occurrence of an initial event, which can be represented by the diagnosis of a disease or a surgical intervention, followed by a subsequent, often undesired event which can be disease recurrence/progression or death; the time passing between these two events is commonly referred as event-free time. However, not every subject in the study will face the second event: these non-happening events are defined as 'censored' and, correspondingly, the survival time as 'censored survival time'. The Kaplan-Meier (K-M) method is used for computing the survival probability and designing the homonym survival curve; a typical statistics summary is embodied by the median survival time, representing the time at which half of the subjects present a lower survival time and the other half a higher one. The statistical comparison between two or more K-M curves is performed mostly through the log-rank test, a statistical test which has the advantage of considering the entire follow-up period by testing the null hypothesis of the absence of differences between the groups regarding the probability of an event (for instance death) at any time point, basing this calculation on every time an event occurs. However, this approach can't provide an estimate of the size of the difference between the groups or a confidence interval, hence requiring an additional type of analysis, the Hazard ratio, that includes the Cox proportional hazards model (Bland and Altman, 2004). The Hazard ratio (HR) was specifically developed for survival analysis and represents an additional metric which quantifies the relative risk of an event between two groups (Machin et al., 2006). Survival analysis may be performed using several different variables, either categorical or continuous; for example, based on the tumor stage classification, patients display varying clinical outcomes with later stages generally associated with a lower survival time compared to the first stage. In the case of continuous

variables, like the gene expression level of a particular gene, patients can be stratified based on the median or percentile expression cut-offs, allowing the definition of a clinical prognostic predictor. In some cases, the optimal cut-point approach was used, which consists in finding and reporting the optimal cutoff at which there is the most significant (log-rank test) stratification of patient groups. The independence of the predictor can additionally be assessed through a multivariate analysis, like the Cox regression analysis, applied when there are multiple potentially interacting covariates, especially continuous variables (Wiesweg, 2021). A stepwise selection (backward elimination) was applied to the multivariate analysis, which consists in starting from the inclusion of all covariates in the model, then gradually removing the non-significant ones to maintain only the most informative covariates.

5.3 LINUX COMMAND LINE PROGRAMS

The Linux distribution 'Ubuntu' was used, either in a virtual machine (VirtualBox) in Windows 64-bit (Ubuntu 18.04.4) or as a full Operating System (OS) in a desktop workstation (Ubuntu 20.04.2). For Python packages and environment management Conda (v. 4.8.3, https://docs.conda.io/projects/conda/en/latest/, Anon, 2020) was used. For the quality control of raw reads from next-generation sequencing, the FastQC (v 0.11.9, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/, Andrews et al., 2010) tool was applied in Windows. For genome indexing and read alignment to the human genome (GRCh38), bowtie2 (v. 2.2.5, http://bowtie-bio.sourceforge.net/bowtie2/index.shtml, Langmead and Salzberg, 2012) was used, while for BAM conversion, indexing and sorting samtools, respectively with the 'view', 'index' and 'sort' functions (v. 1.9, http://samtools.sourceforge.net/, Li H. et al., 2009) were applied. The peak calling was performed with MACS2 function 'callpeak' and default settings (v. 2.1.4, https://github.com/macs3-project/MACS, Zhang et al., 2008). Genomics analysis and manipulation of BED files was done with bedtools (v. 2.29.2 and 2.30.0,https://bedtools.readthedocs.io/en/latest/, Quinlan and Hall, 2010), in particular using the commands 'intersect', 'window', 'merge', 'getfasta', 'bamtobed'. The integrative analysis and visualization of ChIP-seq data was finally performed using the DeepTools suite (v. 3.5.1, https://deeptools.readthedocs.io/en/develop/, Ramírez et al., 2016), using the 'bamCoverage', 'multiBAMSummary', 'multiBigwigSummary', 'plotCorrelation', 'computeMatrix' (reference-point) and 'plotHeatmap' commands.

5.6 THE 'GREAT' TOOL FOR CHIP-SEQ PEAK ANNOTATION

"Genomic Regions Enrichment of Annotations Tool" (GREAT, v. 4.0.4, http://great.stanford.edu/public/html/, McLean et al., 2010) was developed at the Stanford University (Bejerano lab), initially with the purpose of annotating the biological function of non-coding regions and cis-regulatory regions. More generally, a set of genomic regions, e.g. the binding regions resulting from the peak calling of a ChIP-seq dataset, can be used as input. GREAT associates both proximal and distal genomic regions to the putative closest genes, also attributing a functional annotation derived from diverse ontologies (primarily GO), providing the corresponding statistical enrichment. The plots and the list of associated genes (with the computed distance) can be downloaded for additional analysis. For this thesis, the GRCh38 genome assembly was used with default options: "Basal plus extension" with proximal distances of 5 kb upstream and 1 kb downstream from TSS (basal regulatory domain) and extended to the nearest gene regulatory domain but no more than 1 Mb in both directions for distal association. The assembly used was "Human: GRCh38 (UCSC hg38, Dec. 2013)" and the whole genome as background. The queried functional databases were Gene Ontology (Biological process, Cellular component, Molecular function), Human phenotype and Mouse phenotype.

5.5 FUNCTIONAL ENRICHMENT ANALYSIS WITH CLUEGO, A CYTOSCAPE APP

Cytoscape (v. 3.8.2, Shannon et al., 2003) is an open-source Java-based software for the visualization and annotation of complex networks, initially developed for biological research but later extended to other research fields. The functionality of Cytoscape is greatly expanded by the apps, formerly known as plugins, which provide additional features. For this thesis, the ClueGO app (v. 2.5.8, Mlecnik and Bindea, 2009) was used for the functional annotation of gene sets. In ClueGO, the following databases were queried: Gene Ontology (biological process, cellular component, molecular function, ImmuneSystemProcess), KEGG, REACTOME (Pathways and Reactions), WikiPathways and Clinvar. Only pathways with p-value<0.05 were retained. The option "GO term fusion" was also applied, meaning that if many associated terms occurred for each gene, ClueGO would automatically level up in the hierarchical cluster in order to simplify the analysis results. Regarding the statistical methods, the two-sided hypergeometric test was used with Bonferroni step-down correction for multiple testing.

References

Alwine, J., Kemp, D., & Stark, G. (1977). Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proceedings Of The National Academy Of Sciences*, *74*(12), 5350-5354. https://doi.org/10.1073/pnas.74.12.5350

Armaghany, T., Wilson, J., Chu, Q., & Mills, G. (2012). Genetic alterations in colorectal cancer. *Gastrointest Cancer Res*.

Arnold, M., Kim, Y., Czubryt, M., Phan, D., McAnally, J., & Qi, X. et al. (2007). MEF2C Transcription Factor Controls Chondrocyte Hypertrophy and Bone Development. *Developmental Cell*, *12*(3), 377-389. https://doi.org/10.1016/j.devcel.2007.02.004

Assali, A., Harrington, A., & Cowan, C. (2019). Emerging roles for MEF2 in brain development and mental disorders. *Current Opinion In Neurobiology*, *59*, 49-58. https://doi.org/10.1016/j.conb.2019.04.008

Audia, J., & Campbell, R. (2016). Histone Modifications and Cancer. *Cold Spring Harb Perspect Biol* .. https://doi.org/DOI: 10.1101/cshperspect.a019521

Bainbridge, M., Warren, R., Hirst, M., Romanuik, T., Zeng, T., & Go, A. et al. (2006). Analysis of the prostate cancer cell line LNCaP transcriptome using a sequencing-by-synthesis approach. *BMC Genomics*, *7*(1). https://doi.org/10.1186/1471-2164-7-246

Balmain, A. (2001). Cancer genetics: from Boveri and Mendel to microarrays. *Nature Reviews Cancer*, *1*(1), 77-82. https://doi.org/10.1038/35094086

Bannister, A., & Kouzarides, T. (1996). The CBP co-activator is a histone acetyltransferase. *Nature*, *384*(6610), 641-643. https://doi.org/10.1038/384641a0

Barlin, J., Zhou, Q., Leitao, M., Bisogna, M., Olvera, N., & Shih, K. et al. (2015). Molecular Subtypes of Uterine Leiomyosarcoma and Correlation with Clinical Outcome. *Neoplasia*, *17*(2), 183-189. https://doi.org/10.1016/j.neo.2014.12.007

Bártová, E., Krejčí, J., Harničarová, A., Galiová, G., & Kozubek, S. (2008). Histone Modifications and Nuclear Architecture: A Review. *Journal Of Histochemistry & Cytochemistry*, *56*(8), 711-721. https://doi.org/10.1369/jhc.2008.951251

Baylies, M., & Bate, M. (1996). twist: A Myogenic Switch in Drosophila. *Science*, *272*(5267), 1481-1484. https://doi.org/10.1126/science.272.5267.1481

Becht, E., Giraldo, N., Lacroix, L., Buttard, B., Elarouci, N., & Petitprez, F. et al. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biology*, *17*(1). https://doi.org/10.1186/s13059-016-1070-5

Becker-André, M., & Hahlbrock, K. (1989). Absolute mRNA quantification using the polymerase chain reaction (PCR). A novel approach by aPCR aided transcript titration assay (PATTY). *Nucleic Acids Research*, *17*(22), 9437-9446. https://doi.org/10.1093/nar/17.22.9437

Berns, K., Hijmans, E., Mullenders, J., Brummelkamp, T., Velds, A., & Heimerikx, M. et al. (2004). A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature*, *428*(6981), 431-437. https://doi.org/10.1038/nature02371

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., & Kirilovsky, A. et al. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, *25*(8), 1091-1093. https://doi.org/10.1093/bioinformatics/btp101

Bland, J., & Altman, D. (2004). The logrank test. *BMJ*, *328*(7453), 1412.1. https://doi.org/10.1136/bmj.328.7453.1412

Borges, J., Pandiella, A., & Esparís-Ogando, A. (2007). Erk5 nuclear location is independent on dual phosphorylation, and favours resistance to TRAIL-induced apoptosis. *Cellular Signalling*, *19*(7), 1473-1487. https://doi.org/10.1016/j.cellsig.2007.01.023

Brenner, H., Kloor, M., & Pox, C. (2014). Colorectal cancer. *The Lancet*, *383*(9927), 1490-1502. https://doi.org/10.1016/s0140-6736(13)61649-9

Bristow, R., & Hill, R. (2008). Hypoxia, DNA repair and genetic instability. *Nature Reviews Cancer*, *8*(3), 180-192. https://doi.org/10.1038/nrc2344

Butts, B., Hudson, H., Linseman, D., Le, S., Ryan, K., Bouchard, R., & Heidenreich, K. (2005). Proteasome inhibition elicits a biphasic effect on neuronal apoptosis via differential regulation of pro-survival and pro-apoptotic transcription factors. *Molecular And Cellular Neuroscience*, *30*(2), 279-289. https://doi.org/10.1016/j.mcn.2005.07.011

Cancer Genome Atlas Research Network. (2017). Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. *Cell*. https://doi.org/DOI: 10.1016/j.cell.2017.10.014

Candido, E. (1978). Sodium butyrate inhibits histone deacetylation in cultured cells. *Cell*, *14*(1), 105-113. https://doi.org/10.1016/0092-8674(78)90305-7

Caslini, C., Hong, S., Ban, Y., Chen, X., & Ince, T. (2019). HDAC7 regulates histone 3 lysine 27 acetylation and transcriptional activity at super-enhancer-associated genes in breast cancer stem cells. *Oncogene*, *38*(39), 6599-6614. https://doi.org/10.1038/s41388-019-0897-0

Chan, S., Sances, S., Brill, L., Okamoto, S., Zaidi, R., & McKercher, S. et al. (2014). ATM-Dependent Phosphorylation of MEF2D Promotes Neuronal Survival after DNA Damage. *Journal Of Neuroscience*, *34*(13), 4640-4653. https://doi.org/10.1523/jneurosci.2510-12.2014

Chang, S., Young, B., Li, S., Qi, X., Richardson, J., & Olson, E. (2006). Histone Deacetylase 7 Maintains Vascular Integrity by Repressing Matrix Metalloproteinase 10. *Cell*, *126*(2), 321-334. https://doi.org/10.1016/j.cell.2006.05.040

Chen, G., Ning, B., & Shi, T. (2019). Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. *Frontiers In Genetics*, *10*. https://doi.org/10.3389/fgene.2019.00317

Chen, J., Mandel, E., Thomson, J., Wu, Q., Callis, T., & Hammond, S. et al. (2005). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nature Genetics*, *38*(2), 228-233. https://doi.org/10.1038/ng1725

Chen, X., Gao, B., Ponnusamy, M., Lin, Z., & Liu, J. (2017). MEF2 signaling and human diseases. *Oncotarget*, *8*(67), 112152-112165. https://doi.org/10.18632/oncotarget.22899

Choi, J., Jang, H., Kim, H., Lee, J., Kim, S., Cho, E., & Youn, H. (2013). Modulation of lysine methylation in myocyte enhancer factor 2 during skeletal muscle cell differentiation. *Nucleic Acids Research*, *42*(1), 224-234. https://doi.org/10.1093/nar/gkt873

Choi, M., Ryu, S., Hao, R., Wang, B., Kapur, M., Fan, C., & Yao, T. (2014). HDAC 4 promotes P ax7-dependent satellite cell activation and muscle regeneration. *EMBO Reports*, *15*(11), 1175-1183. https://doi.org/10.15252/embr.201439195

Clarke, N., Arenzana, N., Hai, T., Minden, A., & Prywes, R. (1998). Epidermal Growth Factor Induction of the c- jun Promoter by a Rac Pathway. *Molecular And Cellular Biology*, *18*(2), 1065-1073. https://doi.org/10.1128/mcb.18.2.1065

Clocchiatti, A., Di Giorgio, E., Demarchi, F., & Brancolini, C. (2013). Beside the MEF2 axis: Unconventional functions of HDAC4. *Cellular Signalling*, *25*(1), 269-276. https://doi.org/10.1016/j.cellsig.2012.10.002

Clocchiatti, A., Florean, C., & Brancolini, C. (2011). Class IIa HDACs: from important roles in differentiation to possible implications in tumourigenesis. *Journal Of Cellular And Molecular Medicine*, *15*(9), 1833-1846. https://doi.org/10.1111/j.1582-4934.2011.01321.x

Cormier, J., & Pollock, R. (2004). Soft Tissue Sarcomas. *CA: A Cancer Journal For Clinicians*, *54*(2), 94-109. https://doi.org/10.3322/canjclin.54.2.94

Coso, O., Montaner, S., Fromm, C., Lacal, J., Prywes, R., Teramoto, H., & Gutkind, J. (1997). Signaling from G Protein-coupled Receptors to the c-jun Promoter Involves the MEF2 Transcription Factor. *Journal Of Biological Chemistry*, *272*(33), 20691-20697. https://doi.org/10.1074/jbc.272.33.20691

Creyghton, M., Cheng, A., Welstead, G., Kooistra, T., Carey, B., & Steine, E. et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings Of The National Academy Of Sciences*, *107*(50), 21931-21936. https://doi.org/10.1073/pnas.1016071107

Cripps, R., Black, B., Zhao, B., Lien, C., Schulz, R., & Olson, E. (1998). The myogenic regulatory gene Mef2 is a direct target for transcriptional activation by Twist during Drosophila myogenesis. *Genes & Development*, *12*(3), 422-434. https://doi.org/10.1101/gad.12.3.422

Cserjesi, P., & Olson, E. (1991). Myogenin induces the myocyte-specific enhancer binding factor MEF-2 independently of other muscle-specific gene products. *Molecular And Cellular Biology*, *11*(10), 4854-4862. https://doi.org/10.1128/mcb.11.10.4854

Cutano, V., Di Giorgio, E., Minisini, M., Picco, R., Dalla, E., & Brancolini, C. (2019). HDAC7-mediated control of tumour microenvironment maintains proliferative and stemness competence of human mammary epithelial cells. *Molecular Oncology*, *13*(8), 1651-1668. https://doi.org/10.1002/1878-0261.12503

Danielson, L., Menendez, S., Attolini, C., Guijarro, M., Bisogna, M., & Wei, J. et al. (2010). A Differentiation-Based MicroRNA Signature Identifies Leiomyosarcoma as a Mesenchymal Stem Cell-Related Malignancy. *The American Journal Of Pathology*, *177*(2), 908-917. https://doi.org/10.2353/ajpath.2010.091150

Delehanty, L., Bullock, G., & Goldfarb, A. (2012). Protein kinase D-HDAC5 signaling regulates erythropoiesis and contributes to erythropoietin cross-talk with GATA1. *Blood*, *120*(20), 4219-4228. https://doi.org/10.1182/blood-2011-10-387050

Dequiedt, F., Kasler, H., Fischle, W., Kiermer, V., Weinstein, M., Herndier, B., & Verdin, E. (2003). HDAC7, a Thymus-Specific Class II Histone Deacetylase, Regulates Nur77 Transcription and TCR-Mediated Apoptosis. *Immunity*, *18*(5), 687-698. https://doi.org/10.1016/s1074-7613(03)00109-2

Dhani, N., Fyles, A., Hedley, D., & Milosevic, M. (2015). The Clinical Significance of Hypoxia in Human Cancers. *Seminars In Nuclear Medicine*, *45*(2), 110-121. https://doi.org/10.1053/j.semnuclmed.2014.11.002

Di Giorgio, E., Clocchiatti, A., Piccinin, S., Sgorbissa, A., Viviani, G., & Peruzzo, P. et al. (2013). MEF2 Is a Converging Hub for Histone Deacetylase 4 and Phosphatidylinositol 3-Kinase/Akt-Induced Transformation. *Molecular And Cellular Biology*, *33*(22), 4473-4491. https://doi.org/10.1128/mcb.01050-13

Di Giorgio, E., Franforte, E., Cefalù, S., Rossi, S., Dei Tos, A., & Brenca, M. et al. (2017). The co-existence of transcriptional activator and transcriptional repressor MEF2 complexes influences tumor aggressiveness. *PLOS Genetics*, *13*(4), e1006752. https://doi.org/10.1371/journal.pgen.1006752

Di Giorgio, E., Hancock, W., & Brancolini, C. (2018). MEF2 and the tumorigenic process, hic sunt leones. *Biochimica Et Biophysica Acta (BBA) - Reviews On Cancer*, *1870*(2), 261-273. https://doi.org/10.1016/j.bbcan.2018.05.007

Di Giorgio, E., Dalla, E., Franforte, E., Paluvai, H., Minisini, M., & Trevisanut, M. et al. (2020). Different class IIa HDACs repressive complexes regulate specific epigenetic responses related to cell survival in leiomyosarcoma cells. *Nucleic Acids Research*, *48*(2), 646-664. https://doi.org/10.1093/nar/gkz1120

Dionyssiou, M., Nowacki, N., Hashemi, S., Zhao, J., Kerr, A., Tsushima, R., & McDermott, J. (2013). Cross-talk between glycogen synthase kinase 3β (GSK3β) and p38MAPK regulates myocyte enhancer factor 2 (MEF2) activity in skeletal and cardiac muscle. *Journal Of Molecular And Cellular Cardiology*, *54*, 35-44. https://doi.org/10.1016/j.yjmcc.2012.10.013

Dixon, K., & Kopras, E. (2004). Genetic alterations and DNA repair in human carcinogenesis. *Seminars In Cancer Biology*, *14*(6), 441-448. https://doi.org/10.1016/j.semcancer.2004.06.007

Dow, L., O'Rourke, K., Simon, J., Tschaharganeh, D., van Es, J., Clevers, H., & Lowe, S. (2015). Apc Restoration Promotes Cellular Differentiation and Reestablishes Crypt Homeostasis in Colorectal Cancer. *Cell*, *161*(7), 1539-1552. https://doi.org/10.1016/j.cell.2015.05.033

Dumont, N., Liu, B., DeFilippis, R., Chang, H., Rabban, J., & Karnezis, A. et al. (2013). Breast Fibroblasts Modulate Early Dissemination, Tumorigenesis, and Metastasis through Alteration of Extracellular Matrix Characteristics. *Neoplasia*, *15*(3), 249-IN7. https://doi.org/10.1593/neo.121950

Dunn, S., Näthke, I., & Osborne, J. (2013). Computational Models Reveal a Passive Mechanism for Cell Migration in the Crypt. *Plos ONE*, *8*(11), e80516. https://doi.org/10.1371/journal.pone.0080516

Durinck, S., Spellman, P., Birney, E., & Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature Protocols*, *4*(8), 1184-1191. https://doi.org/10.1038/nprot.2009.97

Esteller, M. (2011). Non-coding RNAs in human disease. *Nature Reviews Genetics*, *12*(12), 861-874. https://doi.org/10.1038/nrg3074

Frey, D., Droeser, R., Viehl, C., Zlobec, I., Lugli, A., & Zingg, U. et al. (2010). High frequency of tumor-infiltrating FOXP3+regulatory T cells predicts improved survival in mismatch repair-proficient colorectal cancer patients. *International Journal Of Cancer*, NA-NA. https://doi.org/10.1002/ijc.24989

Furey, T. (2012). ChIP–seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions. *Nature Reviews Genetics*, *13*(12), 840-852. https://doi.org/10.1038/nrg3306

Gabrilovich, D., Ostrand-Rosenberg, S., & Bronte, V. (2012). Coordinated regulation of myeloid cells by tumours. *Nature Reviews Immunology*, *12*(4), 253-268. https://doi.org/10.1038/nri3175

Gossett, L., Kelvin, D., Sternberg, E., & Olson, E. (1989). A new myocyte-specific enhancer-binding factor that recognizes a conserved element associated with multiple muscle-specific genes. *Molecular And Cellular Biology*, *9*(11), 5022-5033. https://doi.org/10.1128/mcb.9.11.5022-5033.1989

Grégoire, S., & Yang, X. (2005). Association with Class IIa Histone Deacetylases Upregulates the Sumoylation of MEF2 Transcription Factors. *Molecular And Cellular Biology*, *25*(6), 2273-2287. https://doi.org/10.1128/mcb.25.6.2273-2287.2005

Grégoire, S., & Yang, X. (2005). Association with Class IIa Histone Deacetylases Upregulates the Sumoylation of MEF2 Transcription Factors. *Molecular And Cellular Biology*, *25*(6), 2273-2287. https://doi.org/10.1128/mcb.25.6.2273-2287.2005

Grégoire, S., Xiao, L., Nie, J., Zhang, X., Xu, M., & Li, J. et al. (2007). Histone Deacetylase 3 Interacts with and Deacetylates Myocyte Enhancer Factor 2. *Molecular And Cellular Biology*, *27*(4), 1280-1295. https://doi.org/10.1128/mcb.00882-06

Grivennikov, S., Greten, F., & Karin, M. (2010). Immunity, Inflammation, and Cancer. *Cell*, *140*(6), 883-899. https://doi.org/10.1016/j.cell.2010.01.025

Guo, X., Jo, V., Mills, A., Zhu, S., Lee, C., & Espinosa, I. et al. (2015). Clinically Relevant Molecular Subtypes in Leiomyosarcoma. *Clinical Cancer Research*, *21*(15), 3501-3511. https://doi.org/10.1158/1078-0432.ccr-14-3141

Haberland, M., Arnold, M., McAnally, J., Phan, D., Kim, Y., & Olson, E. (2006). Regulation of HDAC9 Gene Expression by MEF2 Establishes a Negative-Feedback Loop in the Transcriptional Circuitry of Muscle Differentiation. *Molecular And Cellular Biology*, *27*(2), 518-525. https://doi.org/10.1128/mcb.01415-06

Haberland, M., Montgomery, R., & Olson, E. (2009). The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nature Reviews Genetics*, *10*(1), 32-42. https://doi.org/10.1038/nrg2485

Hahne, F., & Ivanek, R. (2016). Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol Biol.*. https://doi.org/DOI: 10.1007/978-1-4939-3578-9_16

Han, J., Jiang, Y., Li, Z., Kravchenko, V., & Ulevitch, R. (1997). Activation of the transcription factor MEF2C by the MAP kinase p38 in inflammation. *Nature*, *386*(6622), 296-299. https://doi.org/10.1038/386296a0

Han, T., & Prywes, R. (1995). Regulatory role of MEF2D in serum induction of the c-jun promoter. *Molecular And Cellular Biology*, *15*(6), 2907-2915. https://doi.org/10.1128/mcb.15.6.2907

Harris, A. (2002). Hypoxia — a key regulatory factor in tumour growth. *Nature Reviews Cancer*, *2*(1), 38-47. https://doi.org/10.1038/nrc704

Hayashi, M., Kim, S., Imanaka-Yoshida, K., Yoshida, T., Abel, E., & Eliceiri, B. et al. (2004). Targeted deletion of BMK1/ERK5 in adult mice perturbs vascular integrity and leads to endothelial failure. *Journal Of Clinical Investigation*, *113*(8), 1138-1148. https://doi.org/10.1172/jci200419890

He, T., Huang, J., Chen, L., Han, G., Stanmore, D., & Krebs-Haupenthal, J. et al. (2020). Cyclic AMP represses pathological MEF2 activation by myocyte-specific hypo-phosphorylation of HDAC5. *Journal Of Molecular And Cellular Cardiology*, *145*, 88-98. https://doi.org/10.1016/j.yjmcc.2020.05.018

HOBSON, G., KRAHE, R., GARCIA, E., SICILIANO, M., & FUNANAGE, V. (1995). Regional Chromosomal Assignments for Four Members of the MADS Domain Transcription Enhancer Factor 2 (MEF2) Gene Family to Human Chromosomes 15q26, 19p12, 5q14, and 1q12–q23. *Genomics*, *29*(3), 704-711. https://doi.org/10.1006/geno.1995.9007

Hontecillas-Prieto, L., Flores-Campos, R., Silver, A., de Álava, E., Hajji, N., & García-Domínguez, D. (2020). Synergistic Enhancement of Cancer Therapy Using HDAC Inhibitors: Opportunity for Clinical Trials. *Frontiers In Genetics*, *11*. https://doi.org/10.3389/fgene.2020.578011

Hu, Z., Fan, C., Livasy, C., He, X., Oh, D., & Ewend, M. et al. (2009). A compact VEGF signature associated with distant metastases and poor outcomes. *BMC Medicine*, *7*(1). https://doi.org/10.1186/1741-7015-7-9

Hüttenhofer, A., Schattner, P., & Polacek, N. (2005). Non-coding RNAs: hope or hype?. *Trends In Genetics*, *21*(5), 289-297. https://doi.org/10.1016/j.tig.2005.03.007

Ikeda, S., He, A., Kong, S., Lu, J., Bejar, R., & Bodyak, N. et al. (2009). MicroRNA-1 Negatively Regulates Expression of the Hypertrophy-Associated Calmodulin and Mef2a Genes. *Molecular And Cellular Biology*, *29*(8), 2193-2204. https://doi.org/10.1128/mcb.01222-08

Jenuwein, T., & Allis, C. (2001). Translating the Histone Code. *Science*, *293*(5532), 1074-1080. https://doi.org/10.1126/science.1063127

Jia, Y., Chng, W., & Zhou, J. (2019). Super-enhancers: critical roles and therapeutic targets in hematologic malignancies. *Journal Of Hematology & Oncology*, *12*(1). https://doi.org/10.1186/s13045-019-0757-y

Jiang, S., & Mortazavi, A. (2018). Integrating ChIP-seq with other functional genomics data. *Briefings In Functional Genomics*, *17*(2), 104-115. https://doi.org/10.1093/bfgp/ely002

Kao, G., McKenna, W., Guenther, M., Muschel, R., Lazar, M., & Yen, T. (2003). Histone deacetylase 4 interacts with 53BP1 to mediate the DNA damage response. *Journal Of Cell Biology*, *160*(7), 1017-1027. https://doi.org/10.1083/jcb.200209065

Karamboulas, C., Dakubo, G., Liu, J., De Repentigny, Y., Yutzey, K., & Wallace, V. et al. (2007). Disruption of MEF2 activity in cardiomyoblasts inhibits cardiomyogenesis. *Journal Of Cell Science*, *120*(1), 200-200. https://doi.org/10.1242/jcs.03369

Kather, J., Heij, L., Grabsch, H., Loeffler, C., Echle, A., & Muti, H. et al. (2020). Pan-cancer image-based detection of clinically actionable genetic alterations. *Nature Cancer*, *1*(11), 1129-1129. https://doi.org/10.1038/s43018-020-00149-6

Kato, H., Tamamizu-Kato, S., & Shibasaki, F. (2004). Histone Deacetylase 7 Associates with Hypoxia-inducible Factor 1α and Increases Transcriptional Activity. *Journal Of Biological Chemistry*, *279*(40), 41966-41974. https://doi.org/10.1074/jbc.m406320200

Kato, Y. (1997). BMK1/ERK5 regulates serum-induced early gene expression through transcription factor MEF2C. *The EMBO Journal*, *16*(23), 7054-7066. https://doi.org/10.1093/emboj/16.23.7054

Khurana, E., Fu, Y., Chakravarty, D., Demichelis, F., Rubin, M., & Gerstein, M. (2016). Role of non-coding sequence variants in cancer. *Nature Reviews Genetics*, *17*(2), 93-108. https://doi.org/10.1038/nrg.2015.17

Kimura, Y., Morita, T., Hayashi, K., Miki, T., & Sobue, K. (2010). Myocardin Functions as an Effective Inducer of Growth Arrest and Differentiation in Human Uterine Leiomyosarcoma Cells. *Cancer Research*, *70*(2), 501-511. https://doi.org/10.1158/0008-5472.can-09-1469

Kolodziejczyk, S., Wang, L., Balazsi, K., DeRepentigny, Y., Kothary, R., & Megeney, L. (1999). MEF2 is upregulated during cardiac hypertrophy and is required for normal post-natal growth of the myocardium. *Current Biology*, *9*(20), 1203-1206. https://doi.org/10.1016/s0960-9822(00)80027-5

Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell*, *128*(4), 693-705. https://doi.org/10.1016/j.cell.2007.02.005

Lahm, A., Paolini, C., Pallaoro, M., Nardi, M., Jones, P., & Neddermann, P. et al. (2007). Unraveling the hidden catalytic activity of vertebrate class IIa histone deacetylases. *Proceedings Of The National Academy Of Sciences*, *104*(44), 17335-17340. https://doi.org/10.1073/pnas.0706487104

Langmead, B., & Salzberg, S. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357-359. https://doi.org/10.1038/nmeth.1923

Lapierre, M., Linares, A., Dalvai, M., Duraffourd, C., Bonnet, S., & Boulahtouf, A. et al. (2016). Histone deacetylase 9 regulates breast cancer cell proliferation and the response to histone deacetylase inhibitors. *Oncotarget*, *7*(15), 19693-19708. https://doi.org/10.18632/oncotarget.7564

Latchney, S., Jiang, Y., Petrik, D., Eisch, A., & Hsieh, J. (2015). Inducible knockout of Mef2a, -c, and -d from nestin-expressing stem/progenitor cells and their progeny unexpectedly uncouples neurogenesis and dendritogenesis in vivo. *The FASEB Journal*, *29*(12), 5059-5071. https://doi.org/10.1096/fj.15-275651

Lee, J., Bae, S., Jeong, J., Kim, S., & Kim, K. (2004). Hypoxia-inducible factor (HIF-1)α: its protein stability and biological functions. *Experimental & Molecular Medicine*, *36*(1), 1-12. https://doi.org/10.1038/emm.2004.1

Lendahl, U., Lee, K., Yang, H., & Poellinger, L. (2009). Generating specificity and diversity in the transcriptional response to hypoxia. *Nature Reviews Genetics*, *10*(12), 821-832. https://doi.org/10.1038/nrg2665

Li, B., & Dewey, C. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, *12*(1). https://doi.org/10.1186/1471-2105-12-323

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., & Homer, N. et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078-2079. https://doi.org/10.1093/bioinformatics/btp352

Li, M., Linseman, D., Allen, M., Meintzer, M., Wang, X., & Laessig, T. et al. (2001). Myocyte Enhancer Factor 2A and 2D Undergo Phosphorylation and Caspase-Mediated Degradation during Apoptosis of Rat Cerebellar Granule Neurons. *The Journal Of Neuroscience*, *21*(17), 6544-6552. https://doi.org/10.1523/jneurosci.21-17-06544.2001

Lichtenstein, P., Holm, N., Verkasalo, P., Iliadou, A., Kaprio, J., & Koskenvuo, M. et al. (2000). Environmental and Heritable Factors in the Causation of Cancer — Analyses of Cohorts of Twins from Sweden, Denmark, and Finland. *New England Journal Of Medicine*, *343*(2), 78-85. https://doi.org/10.1056/nejm200007133430201

Lin, Q., Schwarz, J., Bucana, C., & N. Olson, E. (1997). Control of Mouse Cardiac Morphogenesis and Myogenesis by Transcription Factor MEF2C. *Science*, *276*(5317), 1404-1407. https://doi.org/10.1126/science.276.5317.1404

Liu, L., Cavanaugh, J., Wang, Y., Sakagami, H., Mao, Z., & Xia, Z. (2003). ERK5 activation of MEF2-mediated gene expression plays a critical role in BDNF-promoted survival of developing but not mature cortical neurons. *Proceedings Of The National Academy Of Sciences*, *100*(14), 8532-8537. https://doi.org/10.1073/pnas.1332804100

Liu N., Nelson B. R., Bezprozvannaya S., Shelton J. M., Richardson J.A., Bassel-Duby R., Olson E. N. (2014). Requirement of MEF2A, C, and D for skeletal muscle regeneration. *Proceedings Of The National Academy Of Sciences*, *111*(11), 8532-8537. https://doi.org/10.1073/pnas.1401732111

Lobera, M., Madauss, K., Pohlhaus, D., Wright, Q., Trocha, M., & Schmidt, D. et al. (2013). Selective class IIa histone deacetylase inhibition via a nonchelating zinc-binding group. *Nature Chemical Biology*, *9*(5), 319-325. https://doi.org/10.1038/nchembio.1223

Lowe, R., Shirley, N., Bleackley, M., Dolan, S., & Shafee, T. (2017). Transcriptomics technologies. *PLOS Computational Biology*, *13*(5), e1005457. https://doi.org/10.1371/journal.pcbi.1005457

Lu, J., McKinsey, T., Nicol, R., & Olson, E. (2000). Signal-dependent activation of the MEF2 transcription factor by dissociation from histone deacetylases. *Proceedings Of The National Academy Of Sciences*, *97*(8), 4070-4075. https://doi.org/10.1073/pnas.080064097

Lukovic, J., Han, K., Pintilie, M., Chaudary, N., Hill, R., Fyles, A., & Milosevic, M. (2019). Intratumoral heterogeneity and hypoxia gene expression signatures: Is a single biopsy adequate?. *Clinical And Translational Radiation Oncology*, *19*, 110-115. https://doi.org/10.1016/j.ctro.2019.09.006

Luoto, K., Kumareswaran, R., & Bristow, R. (2013). Tumor hypoxia as a driving force in genetic instability. *Genome Integrity*, *4*(1), 5. https://doi.org/10.1186/2041-9414-4-5

Lyons, G., Micales, B., Schwarz, J., Martin, J., & Olson, E. (1995). Expression of mef2 genes in the mouse central nervous system suggests a role in neuronal maturation. *The Journal Of Neuroscience*, *15*(8), 5727-5738. https://doi.org/10.1523/jneurosci.15-08-05727.1995

Ma, K., Chan, J., Zhu, G., & Wu, Z. (2005). Myocyte Enhancer Factor 2 Acetylation by p300 Enhances Its DNA Binding Activity, Transcriptional Activity, and Myogenic Differentiation. *Molecular And Cellular Biology*, *25*(9), 3575-3582. https://doi.org/10.1128/mcb.25.9.3575-3582.2005

Ma, L., Liu, J., Liu, L., Duan, G., Wang, Q., & Xu, Y. et al. (2014). Overexpression of the Transcription Factor MEF2D in Hepatocellular Carcinoma Sustains Malignant Character by Suppressing G2–M Transition Genes. *Cancer Research*, *74*(5), 1452-1462. https://doi.org/10.1158/0008-5472.can-13-2171

Ma, Q., & Telese, F. (2015). Genome-wide epigenetic analysis of MEF2A and MEF2C transcription factors in mouse cortical neurons. *Communicative & Integrative Biology*, *8*(6), e1087624. https://doi.org/10.1080/19420889.2015.1087624

Machin, D., Cheung, Y., & Parmar, M. (2006). *Survival Analysis: A Practical Approach*. Wiley.

Mannaerts, I., Eysackers, N., Onyema, O., Van Beneden, K., Valente, S., & Mai, A. et al. (2013). Class II HDAC Inhibition Hampers Hepatic Stellate Cell Activation by Induction of MicroRNA-29. *Plos ONE*, *8*(1), e55786. https://doi.org/10.1371/journal.pone.0055786

Marco-Puche, G., Lois, S., Benítez, J., & Trivino, J. (2019). RNA-Seq Perspectives to Improve Clinical Diagnosis. *Frontiers In Genetics*, *10*. https://doi.org/10.3389/fgene.2019.01152

Marra, M., Hillier, L., & Waterston, R. (1998). Expressed sequence tags — ESTablishing bridges between genomes. *Trends In Genetics*, *14*(1), 4-7. https://doi.org/10.1016/s0168-9525(97)01355-3

Martin, M., Kettmann, R., & Dequiedt, F. (2007). Class IIa histone deacetylases: regulating the regulators. *Oncogene*, *26*(37), 5450-5467. https://doi.org/10.1038/sj.onc.1210613

McGee, S., van Denderen, B., Howlett, K., Mollica, J., Schertzer, J., Kemp, B., & Hargreaves, M. (2008). AMP-Activated Protein Kinase Regulates GLUT4 Transcription by Phosphorylating Histone Deacetylase 5. *Diabetes*, *57*(4), 860-867. https://doi.org/10.2337/db07-0843

McKinsey, T., Zhang, C., & Olson, E. (2000). Activation of the myocyte enhancer factor-2 transcription factor by calcium/calmodulin-dependent protein kinase-stimulated binding of 14-3-3 to histone deacetylase 5. *Proceedings Of The National Academy Of Sciences*, *97*(26), 14400-14405. https://doi.org/10.1073/pnas.260501497

McLachlan, G., Do, K., & Ambroise, C. (2005). *Analyzing microarray gene expression data*. Wiley-Interscience.

McLean, C., Bristor, D., Hiller, M., Clarke, S., Schaar, B., & Lowe, C. et al. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology*, *28*(5), 495-501. https://doi.org/10.1038/nbt.1630

Mielcarek, M., Landles, C., Weiss, A., Bradaia, A., Seredenina, T., & Inuabasi, L. et al. (2013). HDAC4 Reduction: A Novel Therapeutic Strategy to Target Cytoplasmic Huntingtin and Ameliorate Neurodegeneration. *Plos Biology*, *11*(11), e1001717. https://doi.org/10.1371/journal.pbio.1001717

Milde, T., Oehme, I., Korshunov, A., Kopp-Schneider, A., Remke, M., & Northcott, P. et al. (2010). HDAC5 and HDAC9 in Medulloblastoma: Novel Markers for Risk Stratification and Role in Tumor Cell Growth. *Clinical Cancer Research*, *16*(12), 3240-3252. https://doi.org/10.1158/1078-0432.ccr-10-0395

Miska, E. (1999). HDAC4 deacetylase associates with and represses the MEF2 transcription factor. *The EMBO Journal*, *18*(18), 5099-5107. https://doi.org/10.1093/emboj/18.18.5099

Molkentin, J., Li, L., & Olson, E. (1996). Phosphorylation of the MADS-Box Transcription Factor MEF2C Enhances Its DNA Binding Activity. *Journal Of Biological Chemistry*, *271*(29), 17199-17204. https://doi.org/10.1074/jbc.271.29.17199

Moreno, D., Scrideli, C., Cortez, M., De Paula Queiroz, R., Valera, E., & Da Silva Silveira, V. et al. (2010). Differential expression of HDAC3, HDAC7 and HDAC9 is associated with prognosis and survival in childhood acute lymphoblastic leukaemia. *British Journal Of Haematology*, *150*(6), 665-673. https://doi.org/10.1111/j.1365-2141.2010.08301.x

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

Nakato, R., & Sakata, T. (2021). Methods for ChIP-seq analysis: A practical workflow and advanced applications. *Methods*, *187*, 44-53. https://doi.org/10.1016/j.ymeth.2020.03.005

Naya, F., Black, B., Wu, H., Bassel-Duby, R., Richardson, J., Hill, J., & Olson, E. (2002). Mitochondrial deficiency and cardiac sudden death in mice lacking the MEF2A transcription factor. *Nature Medicine*, *8*(11), 1303-1309. https://doi.org/10.1038/nm789

Ohue, Y., & Nishikawa, H. (2019). Regulatory T (Treg) cells in cancer: Can Treg cells be a new therapeutic target?. *Cancer Science*, *110*(7), 2080-2089. https://doi.org/10.1111/cas.14069

Okamoto, S., Krainc, D., Sherman, K., & Lipton, S. (2000). Antiapoptotic role of the p38 mitogen-activated protein kinase-myocyte enhancer factor 2 transcription factor pathway during neuronal differentiation. *Proceedings Of The National Academy Of Sciences*, *97*(13), 7561-7566. https://doi.org/10.1073/pnas.130502697

Okamoto, S., Li, Z., Ju, C., Schölzke, M., Mathews, E., & Cui, J. et al. (2002). Dominant-interfering forms of MEF2 generated by caspase cleavage contribute to NMDA-induced neuronal apoptosis. *Proceedings Of The National Academy Of Sciences*, *99*(6), 3974-3979. https://doi.org/10.1073/pnas.022036399

Olson, E. (1990). MyoD family: a paradigm for development?. *Genes & Development*, *4*(9), 1454-1461. https://doi.org/10.1101/gad.4.9.1454

Olson, E., & Klein, W. (1994). bHLH factors in muscle development: dead lines and commitments, what to leave in and what to leave out. *Genes & Development*, *8*(1), 1-8. https://doi.org/10.1101/gad.8.1.1

Ozsolak, F., & Milos, P. (2010). RNA sequencing: advances, challenges and opportunities. *Nature Reviews Genetics*, *12*(2), 87-98. https://doi.org/10.1038/nrg2934

Park, P. (2009). ChIP–seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics*, *10*(10), 669-680. https://doi.org/10.1038/nrg2641

Paroni, G., Cernotta, N., Dello Russo, C., Gallinari, P., Pallaoro, M., & Foti, C. et al. (2007). PP2A Regulates HDAC4 Nuclear Import. *Molecular Biology Of The Cell*, *19*(2), 655-667. https://doi.org/10.1091/mbc.e07-06-0623

Paroni, G., Mizzau, M., Henderson, C., Del Sal, G., Schneider, C., & Brancolini, C. (2004). Caspase-dependent Regulation of Histone Deacetylase 4 Nuclear-Cytoplasmic Shuttling

Promotes Apoptosis. *Molecular Biology Of The Cell*, *15*(6), 2804-2818. https://doi.org/10.1091/mbc.e03-08-0624

Passier, R., Zeng, H., Frey, N., Naya, F., Nicol, R., & McKinsey, T. et al. (2000). CaM kinase signaling induces cardiac hypertrophy and activates the MEF2 transcription factor in vivo. *Journal Of Clinical Investigation*, *105*(10), 1395-1406. https://doi.org/10.1172/jci8551

Pennacchio, L., Bickmore, W., Dean, A., Nobrega, M., & Bejerano, G. (2013). Enhancers: five essential questions. *Nature Reviews Genetics*, *14*(4), 288-295. https://doi.org/10.1038/nrg3458

Phan, L., Yeung, S., & Lee, M. (2014). Cancer metabolic reprogramming: importance, main features, and potentials for precise targeted anti-cancer therapies. *Cancer Biol Med*. https://doi.org/doi: 10.7497/j.issn.2095-3941.2014.01.001

Pollock, R., Lang, A., Ge, T., Sun, D., Tan, M., & Yu, D. (1998). Wild-type p53 and a p53 temperature-sensitive mutant suppress human soft tissue sarcoma by enhancing cell cycle control. *Clin Cancer Res*.

Pon, J., & Marra, M. (2015). MEF2 transcription factors: developmental regulators and emerging cancer genes. *Oncotarget*, *7*(3), 2297-2312. https://doi.org/10.18632/oncotarget.6223

Pott, S., & Lieb, J. (2015). What are super-enhancers? *Nature Genetics*, *47*(1), 8-12. https://doi.org/10.1038/ng.3167

Puri, P., Wu, Z., Zhang, P., Wood, L., Bhakta, K., & Han, J. et al. (2000). Induction of terminal differentiation by constitutive activation of p38 MAP kinase in human rhabdomyosarcoma cells. *Genes & Development*, *14*(5), 574-584. https://doi.org/10.1101/gad.14.5.574

Qian, B., & Pollard, J. (2010). Macrophage Diversity Enhances Tumor Progression and Metastasis. *Cell*, *141*(1), 39-51. https://doi.org/10.1016/j.cell.2010.03.014

Quail, D., & Joyce, J. (2013). Microenvironmental regulation of tumor progression and metastasis. *Nature Medicine*, *19*(11), 1423-1437. https://doi.org/10.1038/nm.3394

Quinlan, A., & Hall, I. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841-842. https://doi.org/10.1093/bioinformatics/btq033

R Core Team. (2021). R: A language and environment for statistical computing. https://www.R-project.org/

Ramachandran, B., Yu, G., Li, S., Zhu, B., & Gulick, T. (2008). Myocyte Enhancer Factor 2A Is Transcriptionally Autoregulated. *Journal Of Biological Chemistry*, *283*(16), 10318-10329. https://doi.org/10.1074/jbc.m707623200

Ramírez, F., Ryan, D., Grüning, B., Bhardwaj, V., Kilpert, F., & Richter, A. et al. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, *44*(W1), W160-W165. https://doi.org/10.1093/nar/gkw257

Rampalli, S., Li, L., Mak, E., Ge, K., Brand, M., Tapscott, S., & Dilworth, F. (2007). p38 MAPK signaling regulates recruitment of Ash2L-containing methyltransferase complexes to specific genes during differentiation. *Nature Structural & Molecular Biology*, *14*(12), 1150-1156. https://doi.org/10.1038/nsmb1316

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*, *447*(7143), 425-432. https://doi.org/10.1038/nature05918

Rivera, C., & Ren, B. (2013). Mapping Human Epigenomes. *Cell*, *155*(1), 39-55. https://doi.org/10.1016/j.cell.2013.09.011

Ropero, S., & Esteller, M. (2007). The role of histone deacetylases (HDACs) in human cancer. *Molecular Oncology*, *1*(1), 19-25. https://doi.org/10.1016/j.molonc.2007.01.001

RStudio Team. (2020). RStudio: Integrated Development for R.

Rubio, R., Gutierrez-Aranda, I., Sáez-Castillo, A., Labarga, A., Rosu-Myles, M., & Gonzalez-Garcia, S. et al. (2012). The differentiation stage of p53-Rb-deficient bone marrow mesenchymal stem cells imposes the phenotype of in vivo sarcoma development. *Oncogene*, *32*(41), 4970-4980. https://doi.org/10.1038/onc.2012.507

Rudnicki, M., & Jaenisch, R. (1995). The MyoD family of transcription factors and skeletal myogenesis. *Bioessays*, *17*(3), 203-209. https://doi.org/10.1002/bies.950170306

Sandmann, T., Jensen, L., Jakobsen, J., Karzynski, M., Eichenlaub, M., Bork, P., & Furlong, E. (2006). A Temporal Map of Transcription Factor Activity: Mef2 Directly Regulates Target Genes at All Stages of Muscle Development. *Developmental Cell*, *10*(6), 797-807. https://doi.org/10.1016/j.devcel.2006.04.009

Savickiene, J., Borutinskaite, V., Treigyte, G., Magnusson, K., & Navakauskiene, R. (2006). The novel histone deacetylase inhibitor BML-210 exerts growth inhibitory, proapoptotic and differentiation stimulating effects on the human leukemia cell lines. *European Journal Of Pharmacology*, *549*(1-3), 9-18. https://doi.org/10.1016/j.ejphar.2006.08.010

Schena, M., Shalon, D., Davis, R., & Brown, P. (1995). Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. *Science*, *270*(5235), 467-470. https://doi.org/10.1126/science.270.5235.467

Sebastian, S., Faralli, H., Yao, Z., Rakopoulos, P., Palii, C., & Cao, Y. et al. (2013). Tissue-specific splicing of a ubiquitously expressed transcription factor is essential for muscle differentiation. *Genes & Development*, *27*(11), 1247-1259. https://doi.org/10.1101/gad.215400.113

Semenza, G. (2003). Targeting HIF-1 for cancer therapy. *Nature Reviews Cancer*, *3*(10), 721-732. https://doi.org/10.1038/nrc1187

Shalizi, A., & Bonni, A. (2005). Brawn for Brains: The Role of MEF2 Proteins in the Developing Nervous System. *Neural Development*, *69*, 239-266. https://doi.org/https://doi.org/10.1016/S0070-2153(05)69009-6

Shalizi, A., Gaudillière, B., Yuan, Z., Stegmüller, J., Shirogane, T., & Ge, Q. et al. (2006). A Calcium-Regulated MEF2 Sumoylation Switch Controls Postsynaptic Differentiation. *Science*, *311*(5763), 1012-1017. https://doi.org/10.1126/science.1122513

Shannon, P., Markiel, A., Ozier, O., Baliga, N., Wang, J., & Ramage, D. et al. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, *13*(11), 2498-2504. https://doi.org/10.1101/gr.1239303

Sparrow, D. (1999). MEF-2 function is modified by a novel co-repressor, MITR. *The EMBO Journal*, *18*(18), 5085-5098. https://doi.org/10.1093/emboj/18.18.5085

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: the teenage years. *Nature Reviews Genetics*, *20*(11), 631-656. https://doi.org/10.1038/s41576-019-0150-2

Sugo, N., Oshiro, H., Takemura, M., Kobayashi, T., Kohno, Y., & Uesaka, N. et al. (2010). Nucleocytoplasmic translocation of HDAC9 regulates gene expression and dendritic growth in developing cortical neurons. *European Journal Of Neuroscience*, no-no. https://doi.org/10.1111/j.1460-9568.2010.07218.x

Suzuki, E., Satonaka, H., Nishimatsu, H., Oba, S., Takeda, R., & Omata, M. et al. (2004). Myocyte Enhancer Factor 2 Mediates Vascular Inflammation via the p38-Dependent Pathway. *Circulation Research*, *95*(1), 42-49. https://doi.org/10.1161/01.res.0000134631.75684.4a

Tang, X. (2005). Cyclin-Dependent Kinase 5 Mediates Neurotoxin-Induced Degradation of the Transcription Factor Myocyte Enhancer Factor 2. *Journal Of Neuroscience*, *25*(19), 4823-4834. https://doi.org/10.1523/jneurosci.1331-05.2005

Tang, X. (2005). Cyclin-Dependent Kinase 5 Mediates Neurotoxin-Induced Degradation of the Transcription Factor Myocyte Enhancer Factor 2. *Journal Of Neuroscience*, *25*(19), 4823-4834. https://doi.org/10.1523/jneurosci.1331-05.2005

Tao, R., de Zoeten, E., Özkaynak, E., Chen, C., Wang, L., & Porrett, P. et al. (2007). Deacetylase inhibition promotes the generation and function of regulatory T cells. *Nature Medicine*, *13*(11), 1299-1307. https://doi.org/10.1038/nm1652

Taylor, M., & Hughes, S. (2017). Mef2 and the skeletal muscle differentiation program. *Seminars In Cell & Developmental Biology*, *72*, 33-44. https://doi.org/10.1016/j.semcdb.2017.11.020

The Cancer Genome Atlas Network. (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, *487*(7407), 330-337. https://doi.org/10.1038/nature11252

Therneau, T. (2021). A Package for Survival Analysis in R. Retrieved 30 December 2021, from.

Tobin, S., Hashemi, S., Dadson, K., Turdi, S., Ebrahimian, K., & Zhao, J. et al. (2017). Heart Failure and MEF2 Transcriptome Dynamics in Response to β-Blockers. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-04762-x

Tomczak, K., Czerwińska, P., & Wiznerowicz, M. (2015). The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Współczesna Onkologia*, *1A*, 68-77. https://doi.org/10.5114/wo.2014.47136

Tóth, A., Schell, R., Lévay, M., Vettel, C., Theis, P., & Haslinger, C. et al. (2018). Inflammation leads through PGE / EP 3 signaling to HDAC 5/ MEF 2-dependent transcription in cardiac myocytes. *EMBO Molecular Medicine*, *10*(7). https://doi.org/10.15252/emmm.201708536

Townley-Tilson, W., Callis, T., & Wang, D. (2010). MicroRNAs 1, 133, and 206: Critical factors of skeletal and cardiac muscle development, function, and disease. *The International Journal Of Biochemistry & Cell Biology*, *42*(8), 1252-1255. https://doi.org/10.1016/j.biocel.2009.03.002

Van Verk, M., Hickman, R., Pieterse, C., & Van Wees, S. (2013). RNA-Seq: revelation of the messengers. *Trends In Plant Science*, *18*(4), 175-179. https://doi.org/10.1016/j.tplants.2013.02.001

Vega, R., Matsuda, K., Oh, J., Barbosa, A., Yang, X., & Meadows, E. et al. (2004). Histone Deacetylase 4 Controls Chondrocyte Hypertrophy during Skeletogenesis. *Cell*, *119*(4), 555-566. https://doi.org/10.1016/j.cell.2004.10.024

Velculescu, V., Zhang, L., Vogelstein, B., & Kinzler, K. (1995). Serial Analysis of Gene Expression. *Science*, *270*(5235), 484-487. https://doi.org/10.1126/science.270.5235.484

Verzi, M., Agarwal, P., Brown, C., McCulley, D., Schwarz, J., & Black, B. (2007). The Transcription Factor MEF2C Is Required for Craniofacial Development. *Developmental Cell*, *12*(4), 645-652. https://doi.org/10.1016/j.devcel.2007.03.007

Wang, A., Kruhlak, M., Wu, J., Bertos, N., Vezmar, M., & Posner, B. et al. (2000). Regulation of Histone Deacetylase 4 by Binding of 14-3-3 Proteins. *Molecular And Cellular Biology*, *20*(18), 6904-6912. https://doi.org/10.1128/mcb.20.18.6904-6912.2000

Wang, B., Zhao, Q., Zhang, Y., Liu, Z., Zheng, Z., & Liu, S. et al. (2021). Targeting hypoxia in the tumor microenvironment: a potential strategy to improve cancer immunotherapy. *Journal Of Experimental & Clinical Cancer Research*, *40*(1). https://doi.org/10.1186/s13046-020-01820-7

Wang, X., Tang, X., Gong, X., Albanis, E., Friedman, S., & Mao, Z. (2004). Regulation of hepatic stellate cell activation and growth by transcription factor myocyte enhancer factor 2. *Gastroenterology*, *127*(4), 1174-1188. https://doi.org/10.1053/j.gastro.2004.07.007

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), 57-63. https://doi.org/10.1038/nrg2484

Warnes, G. (2015). gplots: Various R programming tools for plotting data. *R Package Vers*.

West, A., & Johnstone, R. (2014). New and emerging HDAC inhibitors for cancer treatment. *Journal Of Clinical Investigation*, *124*(1), 30-39. https://doi.org/10.1172/jci69738

Westermann, A., Förstner, K., Amman, F., Barquist, L., Chao, Y., & Schulte, L. et al. (2016). Dual RNA-seq unveils noncoding RNA functions in host–pathogen interactions. *Nature*, *529*(7587), 496-501. https://doi.org/10.1038/nature16547

Whiteside, T., Schuler, P., & Schilling, B. (2012). Induced and natural regulatory T cells in human cancer. *Expert Opinion On Biological Therapy*, *12*(10), 1383-1397. https://doi.org/10.1517/14712598.2012.707184

Wickham, H. (2007). Reshaping Data with thereshape Package. *Journal Of Statistical Software*, *21*(12). https://doi.org/10.18637/jss.v021.i12

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., & François, R. et al. (2019). Welcome to the Tidyverse. *Journal Of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wiesweg, M. (2021). *Univariate Survival Analysis*. Cran.r-project.org. Retrieved 30 December 2021, from https://cran.r-project.org/web/packages/survivalAnalysis/vignettes/univariate.html.

Wilkinson, L. (2016). ggplot2: Elegant Graphics for Data Analysis. *Biometrics*, *67*(2), 678-679. https://doi.org/10.1111/j.1541-0420.2011.01616.x

Wilson, W., & Hay, M. (2011). Targeting hypoxia in cancer therapy. *Nature Reviews Cancer*, *11*(6), 393-410. https://doi.org/10.1038/nrc3064

Wu, H., Yang, K., Zhang, Z., Leisten, E., Li, Z., & Xie, H. et al. (2019). Development of Multifunctional Histone Deacetylase 6 Degraders with Potent Antimyeloma Activity. *Journal Of Medicinal Chemistry*, *62*(15), 7042-7057. https://doi.org/10.1021/acs.jmedchem.9b00516

Wu, W., Folter, S., Shen, X., Zhang, W., & Tao, S. (2011). Vertebrate Paralogous MEF2 Genes: Origin, Conservation, and Evolution. *Plos ONE*, *6*(3), e17334. https://doi.org/10.1371/journal.pone.0017334

Wu, Y., Dey, R., Han, A., Jayathilaka, N., Philips, M., Ye, J., & Chen, L. (2010). Structure of the MADS-box/MEF2 Domain of MEF2A Bound to DNA and Its Implication for Myocardin Recruitment. *Journal Of Molecular Biology*, *397*(2), 520-533. https://doi.org/10.1016/j.jmb.2010.01.067

Xu, W., Parmigiani, R., & Marks, P. (2007). Histone deacetylase inhibitors: molecular mechanisms of action. *Oncogene*, *26*(37), 5541-5552. https://doi.org/10.1038/sj.onc.1210620

Yamashita, S., Kishino, T., Takahashi, T., Shimazu, T., Charvat, H., & Kakugawa, Y. et al. (2018). Genetic and epigenetic alterations in normal tissues have differential impacts on cancer risk among tissues. *Proceedings Of The National Academy Of Sciences*, *115*(6), 1328-1333. https://doi.org/10.1073/pnas.1717340115

Yang, J., Du, X., Chen, K., Ylipää, A., Lazar, A., & Trent, J. et al. (2009). Genetic aberrations in soft tissue leiomyosarcoma. *Cancer Letters*, *275*(1), 1-8. https://doi.org/10.1016/j.canlet.2008.06.013

Yang, Q., She, H., Gearing, M., Colla, E., Lee, M., Shacka, J., & Mao, Z. (2009). Regulation of Neuronal Survival Factor MEF2D by Chaperone-Mediated Autophagy. *Science*, *323*(5910), 124-127. https://doi.org/10.1126/science.1166088

Yang, Q., She, H., Gearing, M., Colla, E., Lee, M., Shacka, J., & Mao, Z. (2009). Regulation of Neuronal Survival Factor MEF2D by Chaperone-Mediated Autophagy. *Science*, *323*(5910), 124-127. https://doi.org/10.1126/science.1166088

Yin, Y., She, H., Li, W., Yang, Q., Guo, S., & Mao, Z. (2012). Modulation of Neuronal Survival Factor MEF2 by Kinases in Parkinson's Disease. *Frontiers In Physiology*, *3*. https://doi.org/10.3389/fphys.2012.00171

Yoshida, M., Horinouchi, S., & Beppu, T. (1995). Trichostatin A and trapoxin: Novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays*, *17*(5), 423-430. https://doi.org/10.1002/bies.950170510

Youn, H., Grozinger, C., & Liu, J. (2000). Calcium Regulates Transcriptional Repression of Myocyte Enhancer Factor 2 by Histone Deacetylase 4. *Journal Of Biological Chemistry*, *275*(29), 22563-22567. https://doi.org/10.1074/jbc.c000304200

Yu, G., Wang, L., & He, Q. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, *31*(14), 2382-2383. https://doi.org/10.1093/bioinformatics/btv145

Yu, W., Huang, C., Wang, Q., Huang, T., Ding, Y., & Ma, C. et al. (2014). MEF2 transcription factors promotes EMT and invasiveness of hepatocellular carcinoma through TGF-β1 autoregulation circuitry. *Tumor Biology*, *35*(11), 10943-10951. https://doi.org/10.1007/s13277-014-2403-1

Yu, Y. (1996). Distinct Domains of Myocyte Enhancer Binding Factor-2A Determining Nuclear Localization and Cell Type-specific Transcriptional Activity. *Journal Of Biological Chemistry*, *271*(40), 24675-24683. https://doi.org/10.1016/s0021-9258(18)40058-0

Yu, Y., Breitbart, R., Smoot, L., Lee, Y., Mahdavi, V., & Nadal-Ginard, B. (1992). Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors. *Genes & Development*, *6*(9), 1783-1798. https://doi.org/10.1101/gad.6.9.1783

Zahr, A., Alcaide, P., Yang, J., Jones, A., Gregory, M., & dela Paz, N. et al. (2016). Endomucin prevents leukocyte–endothelial cell adhesion and has a critical role under resting and inflammatory conditions. Nature Communications, 7(1). https://doi.org/10.1038/ncomms10363

Zhang, C., McKinsey, T., Chang, S., Antos, C., Hill, J., & Olson, E. (2002). Class II Histone Deacetylases Act as Signal-Responsive Repressors of Cardiac Hypertrophy. *Cell*, *110*(4), 479-488. https://doi.org/10.1016/s0092-8674(02)00861-9

Zhang, M., Truscott, J., & Davie, J. (2013). Loss of MEF2D expression inhibits differentiation and contributes to oncogenesis in rhabdomyosarcoma cells. *Molecular Cancer*, *12*(1). https://doi.org/10.1186/1476-4598-12-150

Zhang, M., Zhu, B., & Davie, J. (2015). Alternative Splicing of MEF2C pre-mRNA Controls Its Activity in Normal Myogenesis and Promotes Tumorigenicity in Rhabdomyosarcoma Cells. *Journal Of Biological Chemistry*, *290*(1), 310-324. https://doi.org/10.1074/jbc.m114.606277

Zhang, Y., Liu, T., Meyer, C., Eeckhoute, J., Johnson, D., & Bernstein, B. et al. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*, *9*(9). https://doi.org/10.1186/gb-2008-9-9-r137

Zou, Y., Rong, Y., Tan, Y., Xiao, J., Yu, Z., & Chen, Y. et al. (2019). A signature of hypoxia-related factors reveals functional dysregulation and robustly predicts clinical outcomes in stage I/II colorectal cancer patients. *Cancer Cell International*, *19*(1). https://doi.org/10.1186/s12935-019-0964-1

Published Articles

"Transcriptomic and genomic studies classify NKL54 as a histone deacetylase inhibitor with indirect influence on MEF2-dependent transcription" Minisini, Martina; Di Giorgio, Eros; Kerschbamer, Emanuela; Dalla, Emiliano; Faggiani, Massimo; Franforte, Elisa; Meyer-Almes, Franz-Josef; Ragno, Rino; Antonini, Lorenzo; Mai, Antonello; Fiorentino, Francesco; Rotili, Dante; Chinellato, Monica; Cendron, Laura; Perin, Stefano; Weichenberger, Christian; Angelini, Alessandro; Brancolini, Claudio. "Nucleic Acids Research", 2022.

Acknowledgments

I am thankful for having received the opportunity of doing this journey, made of constant learning and slow improvements, both at the professional and the personal level. I think that pursuing a PhD represents a great lesson of humility but at the same time of trust in one's own intellectual potential. For these reasons, I would like firstly to thank Professor Claudio Brancolini for believing in me once again, and my co-supervisor Dr. Emiliano Dalla for his precious advice and scientific rigor. I also need to say thanks to Raffaella, Eros and all the other group components for their meaningful questions and help.

Thank you to my entire family too (including Kira and Assia) for the unconditioned support and patience, after a thousand hurdles I could finally find my way.

2566–2586 *Nucleic Acids Research, 2022, Vol. 50, No. 5* *Published online 12 February 2022*
https://doi.org/10.1093/nar/gkac081

Transcriptomic and genomic studies classify NKL54 as a histone deacetylase inhibitor with indirect influence on MEF2-dependent transcription

Martina Minisini[1,†], Eros Di Giorgio[1,†], Emanuela Kerschbamer[2], Emiliano Dalla[1], Massimo Faggiani[1], Elisa Franforte[1], Franz-Josef Meyer-Almes[3], Rino Ragno[4], Lorenzo Antonini[4], Antonello Mai[5], Francesco Fiorentino[5], Dante Rotili[5], Monica Chinellato[6], Stefano Perin[7,8], Laura Cendron[6], Christian X. Weichenberger[2], Alessandro Angelini[7,8] and Claudio Brancolini[1,*]

[1]Department of Medicine, Università degli Studi di Udine. P.le Kolbe 4, 33100 Udine Italy, [2]Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck. Via Galvani 31, 39100 Bolzano, Italy, [3]Department of Chemical Engineering and Biotechnology, University of Applied Science, Haardtring 100, 64295 Darmstadt, Germany, [4]Rome Center for Molecular Design, Department of Chemistry and Technology of Drugs, "Sapienza" University of Rome, Piazzale Aldo Moro 5, Rome 00185, Italy, [5]Department of Chemistry and Technology of Drugs, "Sapienza" University of Rome, Piazzale Aldo Moro 5, Rome 00185, Italy, [6]Department of Biology, University of Padova, Via U. Bassi, 58/B, 35121 Padova, Italy, [7]Department of Molecular Sciences and Nanosystems, Ca' Foscari University of Venice, Via Torino 155, 30172 Mestre, Italy and [8]European Centre for Living Technology (ECLT), Dorsoduro 3911, Calle Crosera, 30123 Venice, Italy

Received January 08, 2022; Editorial Decision January 25, 2022; Accepted January 25, 2022

ABSTRACT

In leiomyosarcoma class IIa HDACs (histone deacetylases) bind MEF2 and convert these transcription factors into repressors to sustain proliferation. Disruption of this complex with small molecules should antagonize cancer growth. NKL54, a PAOA (pimeloylanilide o-aminoanilide) derivative, binds a hydrophobic groove of MEF2, which is used as a docking site by class IIa HDACs. However, NKL54 could also act as HDAC inhibitor (HDACI). Therefore, it is unclear which activity is predominant. Here, we show that NKL54 and similar derivatives are unable to release MEF2 from binding to class IIa HDACs. Comparative transcriptomic analysis classifies these molecules as HDACIs strongly related to SAHA/vorinostat. Low expressed genes are upregulated by HDACIs, while abundant genes are repressed. This transcriptional resetting correlates with a reorganization of H3K27 acetylation around the transcription start site (TSS). Among the upregulated genes there are several BH3-only family members, thus explaining the induction of apoptosis. Moreover, NKL54 triggers the upregulation of MEF2 and the downregulation of class IIa HDACs. NKL54 also increases the binding of MEF2D to promoters of genes that are upregulated after treatment. In summary, although NKL54 cannot outcompete MEF2 from binding to class IIa HDACs, it supports MEF2-dependent transcription through several actions, including potentiation of chromatin binding.

INTRODUCTION

The MEF2 family of transcription factors (TFs) includes four paralogues MEF2A, B, C and D that regulate differentiation and important adaptive responses. They coordinate the expression of a rather large number of genes in a context- and partner-dependent manner (1). Dysregulations of these TFs have been documented in various diseases (1–5). The involvement of MEF2s in various pathological contexts makes them attractive candidates for novel therapeutic approaches aimed at restarting a dysregulated transcriptional program. MEF2 proteins are characterized by the presence of the highly conserved MADS and MEF2 domains in the N-terminal region. These domains are essential for DNA binding, dimerization and interaction with other partners. In contrast, the C-terminal region is much less conserved and is involved in transcriptional activation

*To whom correspondence should be addressed. Tel: +39 0432 494382; Fax: +39 0432 494301; Email: claudio.brancolini@uniud.it
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

(1). A hydrophobic groove within the MADS/MEF2 domain contains the binding site for amphipathic α-helices present in some MEF2 partners. Transcriptional repressors, such as class IIa HDACs or Cabin1, as well as activators, for example the histone acetyltransferase p300, bind MEF2 via this mechanism (6,7). A β-sheet organizes the floor of this deep hydrophobic groove, while two helices form the rim (6). In HDAC9, the hydrophobic side of the amphipathic α-helix consisting of Val143, Leu147, Phe150 and Leu151 fits precisely into the hydrophobic groove of MEF2B (8). Similarly, the co-repressor Cabin1 adopts an amphipathic α-helix to bind this hydrophobic groove, forming a triple-helical interaction (6).

The possibility of affecting the transcriptional activity of MEF2 by small molecules that can bind this hydrophobic groove has been exploited in the past (9). A virtual screen identified a series of small molecules belonging to the class of PAOA (pimeloylanilide-o aminoanilides). Starting from the original compound BML-210 ((*N*-(2-aminophenyl)-*N*′ phenyloctanol diamine), which was initially identified as a pan-HDACs inhibitor (10), several analogous compounds with improved solubility were characterized (11).

The MEF2–HDAC axis is frequently circuited in leiomyosarcoma (LMS) a rare group of soft tissue sarcomas (STS), highly aggressive and with few therapeutic options (11–14). In this manuscript, we have investigated the possibility to block LMS proliferation, by targeting the interaction of MEF2 with class IIa HDACs, using PAOA derivatives. We found that PAOA derivatives are potent inhibitors of LMS cell proliferation, however, they are unable to disrupt the binding between MEF2 and class IIa HDACs. Conversely, PAOA derivatives appear to act mainly as inhibitors of zinc-dependent class I HDACs.

MATERIALS AND METHODS

Antibodies and chemicals

The primary antibodies used were anti: MEF2D (BD Bioscience); MEF2A (C-21), (Santa Cruz Biotechnology); MEF2C (15); Actin (Sigma-Aldrich); HDAC4 (16); HDAC5 (17); HDAC7 (18); HDAC9 (19); H3K27ac (ab4729) and H3K9ac (ab4441) (Abcam); Histone H3 (H0164, Sigma-Aldrich) HDAC3 (PA5-29026, Invitrogen). The following chemicals were used: SAHA (Cayman Chemicals); TMP195 (MedChemExpress), BML-210 (Sigma-Aldrich). The PAOA derivatives MC2983, MC2984, MC2985 and MC2991 were synthetized**.** All new compounds had spectral (^{1}H NMR, ESI-MS) data in agreement with the structure**.** Full details of the syntheses will be reported elsewhere. NKL54 (*N*-(2-aminophenyl)-*N*'-[3-(trifluoromethyl)phenyl]heptanediamide) was synthesized by SIA Chemspace (Riga, Latvia).

Cell cultures and cytofluorimetric analysis

Leiomyosarcomas cells (LMS), SK-UT-1, SK-LMS-1 and DMR were grown as previously described (19). For PI staining, cells were collected and resuspended in 0.1 ml of 10 μg/ml propidium iodide (PI) (Sigma-Aldrich), in PBS and incubated for 10 min at RT. After washes, cells were fixed with 1% formaldehyde (Sigma-Aldrich) and treated with 10 μg/ml RNase A. Fluorescence was determined with a FACScan™ (Beckman Dickinson) and with Countess II FL automated cell counter (Invitrogen).

Immunoblotting

Cell lysates, after SDS-PAGE and immunoblotting on nitrocellulose (Whatman), were incubated with primary antibodies. HPR-conjugated secondary antibodies were obtained from Sigma-Aldrich and blots were developed with Super Signal West Dura (Thermo Scientific). For antibodies stripping, blots were incubated for 10min with Restore PLUS Western Blot Stripping Buffer (Thermo Scientific).

Caspase and resazurin reduction assays

The caspase activity was evaluated using the Apo-ONE caspase-3/7 homogeneous assay (Promega). Cells grown in 96-well plates were treated with the different insults and tested for caspase activity as recommended by the vendor. Resazurin assay was done as already described (20). Briefly, cells were incubated for 150 min. at 37°C with resazurin solution (0.15 mg/ml) (Sigma-Aldrich). The product of reduction was quantified by using the PerkinElmer EnSpire 2300 Multilabel Reader.

Molecular modelling

Three-dimensional atomic coordinates of crystallized MEF2A in complex with DNA and BML-210 were retrieved from the protein data bank using accession code 3MU6. The structure was subjected to a cleaning procedure eliminating all water and non-protein atoms/molecules. During the complex cleaning the hydrogen positions were optimized by means of a single point minimization using the default settings available in UCSF Chimera (version 1.14) (21) using the AmberF14SB force field (22). The cleaned minimized complex was then separated into lock (protein + DNA) and key (BML-210) for the subsequent docking assessment procedure. Smina and Plants programs (23,24) implemented in the Py-Docking web app of 3d-qsar.com portal (2019) were assessed for docking suitability. Experimental conformation re-docking (ECRD) and random conformation re-docking (RCRD) procedures (25–27) indicated the Plants/PLP combination as the best performing. For all docking the program default setting were maintained with an extended docking space of 4 Å (extended grid for Smina and extended radius for Plants). Only the lowest energy conformation was considered for the RMSD evaluation. In the case of Smina the docked conformations were also re-scored by the internal minimization available features.

HDAC assay

Lysine deacetylase assay was carry-out using the HDAC-Glo I/II assay kit (Promega), following manufacturer specifications. Briefly, native lysates were generated from 1.0×10^5 SK-UT-1 cells, previously incubated for 4 h with HDAC inhibitors. The luminescence was quantified by using the Modulus II microplate multimode reader (Turner Biosystem).

For the *in vitro* enzyme activity assays, recombinant HDAC4 and 8 were produced as described previously (28). The other HDAC isoenzymes were purchased from BPS Bioscience. To determine the inhibitory effect of compounds on HDACs, 1 nM of the respective HDAC isozyme was incubated with a serial dilution of the compounds for 30 min at 30°C in the assay buffer (25 mM Tris–HCl, pH 8.0, 75 mM KCl, 0.001% v/v Pluronic F-127). The catalytic reaction was carried out by the addition of 50 μM of the substrate Boc-Lys(Ac)-AMC for HDACs 1, 2, 3, and 6 or 20 μM of the substrate Boc-Lys(trifluoracetyl)-AMC for HDACs 4 and 8 followed by an incubation for 60 min at 30°C. The reaction was stopped by adding 40 μM SAHA for HDACs 1, 2, 3 and 6 or 20 μM SATFMK for HDACs 4 and 8. The deacetylated substrate was converted into the fluorescent product AMC by the addition of 0.5 mg/ml trypsin. The release of the AMC was followed in a microplate reader (excitation: 360 nm, emission: 460 nm; PHERAstar FS, BMG LABTECH) and then correlated to enzyme activity. All obtained dose–response curves were fitted to a four parameter fit model provided by Prism 6 yielding the IC_{50}-value

Protein expression and purification

MEF2A (1–92) and MEF2D (1–95) were cloned into pRham and pETite vectors (Lucigen), respectively. Both proteins were expressed using *Escherichia coli* T7 SHuffle cells (NEB) and growth in Terrific Broth (TB) media. Expression of MEF2A and MEF2D was induced at OD_{600} = 0.8 by adding 0.2% (w/v) rhamnose and 1 mM isopropyl-β-D-1thiogalactopyranoside (IPTG), respectively. Induced cells were maintained at 28°C overnight. Cell pellets were resuspended in lysis buffer (10 mM HEPES pH 7.7, 30 mM NaCl, 0.5 mM EDTA, 0.5 mM DTT) and processed by French press. Both MEF2A and MEF2D proteins were purified via ion exchange chromatography using 20 ml of SP-Sepharose resin (GE Healthcare) equilibrated with 10 mM HEPES, pH 7.7. Elution was achieved by applying a 0–2 M $(NH_4)_2SO_4$ linear gradient. Eluted fractions were collected and further purified using a HiPrep Butyl FF 16/10 column (GE Healthcare) equilibrated with 10 mM HEPES pH 7.7, 2 M $(NH_4)_2SO_4$, 0.5 mM EDTA, 0.5 mM DTT. Elution was performed by applying a 2–0 M salt gradient. Purest fractions were collected and loaded on a HiLoad Superdex 75 16/60 column (GE Healthcare), equilibrated with storage buffer 10 mM HEPES pH 7.7, 200 mM NaCl, 0.5 mM EDTA, 0.5 mM DTT, 10% v/v glycerol). Both proteins were concentrated to 66 μM, aliquoted and stored at −80°C.

Chemical synthesis of peptides

Peptides pHDAC4 (aa 170–183; AcNH-GSGEVK MKLQEFVLNKK-$CONH_2$) and F-pHDAC4 (aa 170–183, fluorescein-GSGEVKMKLQEFVLNKK-$CONH_2$) were synthesized by standard Fmoc (9-fluorenylmethoxycarbonyl) solid-phase peptide synthesis (SPPS). Fmoc-protected amino acids, PyBOP, 5(6)-carboxyfluorescein, acetic anhydride, anisole, dichloromethane (DCM) and *N,N*-dimethylformamide (DMF) and Rink Amide MBHA resin (100–200 mesh, loading 0.4–0.9 mmol g^{-1} resin, 0.01 mmol scale) were purchased from Novabiochem. Acetonitrile (ACN), formic acid, *N*-methylmorpholine (NMM), octanedithiol (ODT), piperidine, trifluoroacetic acid (TFA) and thioanisole were purchased from Sigma-Aldrich. *N*-methylpirrolidone (NMP) was purchased from VWR. All chemicals were used as received without further purification. Peptides were prepared using a MultiPep RSi peptide synthesiser (Intavis). Fmoc groups were removed using a 20% v/v solution of piperidine in DMF (180 μl × 2). Amino acid coupling was carried out twice for each Fmoc-amino acid (7.5 eq., 0.5 M solution in DMF). Final acetylation capping was performed using a 5% (v/v) solution of acetic anhydride in DMF. DCM washes (0.3 ml × 5) were performed at the end of synthetic process. NMP was used as cosolvent in the peptide synthesis. A 4 M NMM solution in DMF was added as weak base for Fmoc deprotection. The final peptides were deprotected (side-chain protected groups) and cleaved from the resin using a TFA/thioanisole/H_2O/anisole/ODT mixture (90/2.5/2.5/2.5/2.5% v/v) for 3 h at room temperature. The resin was removed by filtration under vacuum and the peptides were precipitated with cold diethyl ether (50 ml). The precipitated peptides were resuspended in diethyl ether (30 ml × 2) and centrifuged (3 times). Finally, the peptides were dissolved in H_2O:ACN (1:1), freeze-dried and lyophilized. Crude peptides were dissolved in DMSO and purified by preparative reversed-phase high performance liquid chromatography (RP-HPLC) on a Waters Delta Prep LC 4000 System equipped with Waters 2489 dual λ absorbance detector and with both Waters 600 pump, PrepLC Controller (Waters) and a C18 SymmetryPrep (Waters) functionalized silica column (7 μm, 19 mm × 150 mm). At a flow rate of 20 ml min^{-1}, a linear gradient (10% to 50% in 35 min) was applied with a mobile phase composed of eluant A (99.9% v/v H_2O, 0.1% v/v TFA) and eluant B (99.9% v/v ACN and 0.1% v/v TFA). The purified peptides were freeze-dried. The purity and molecular mass of the peptides was assessed by LC-ESI as described below. Concentrations of peptides were determined by UV spectrophotometry.

Mass spectrometric analysis

The molecular mass of each peptide was determined by electrospray ionisation mass spectrometry (ESI–MS) performed on a single quadrupole liquid chromatograph InfinityLab LC/MSD mass spectrometer (InfinityLab LC/MSD, Agilent) coupled to a 1260 Infinity II LC system, Agilent). The reversed-phase HPLC column was a Nucleosil 100-5 C18 Macherey-Nagel (5 μm, 125 mm × 4 mm). The system operated with the standard ESI source and in the positive ionisation mode. For mass spectrometric analysis, the samples were mixed with 50% (v/v) ACN, 50% (v/v) H_2O. Peptides were run at a flow rate of 1 ml min^{-1} with a linear gradient of solvent B over 15 min (A: 99.9% v/v H_2O and 0.1% v/v formic acid; B: 99.9% v/v ACN and 0.1% v/v formic acid). Data were acquired, processed and analysed using Agilent OpenLAB CDS (Agilent Technologies) and MestReNova (Mestrelab Research S.L.) software.

Fluorescence polarization binding assay

Fluorescence polarization (*FP*) values were determined using the Equation (1), where *S* is the fluorescence intensity of emitted light parallel to excitation, *P* is the fluorescence intensity of emitted light perpendicular to excitation, and *G* is the correction factor that correct for instrument bias.

$$FP = 1000 \quad (1)$$

The *G* factor was experimentally determined using the probe alone.

Fluorescence polarization assays, direct binding

Proteins were diluted in 10 mM HEPES, 200 mM NaCl, 1 mM DTT, 1 mM EDTA, pH 7.7 at final concentrations ranging from 0.5 to 300 μM. Titration assays were performed using fluorescently labelled peptide F-pHDAC4 at a final concentration of 0.26 μM. Each mixture (100μl) was transferred into black 96-well microplates (Optiplate, PerkinElmer) and incubated at room temperature for 1 h. Polarization signals were recorded at 25°C using an EnVision Multlabel Plate Reader (PerkinElmer) with an excitation filter at 480 nm, an emission filter at 535 nm and a 505 nm dichroic mirror. Orbital shaking (200 rpm for 0.1 s) was applied. The average fluorescence polarization values of at least three independent experiments were plotted as a function of MEF2A or MEF2D concentration. Equilibrium dissociation constants (K_D) were determined by non-linear regression analyses of polarization (*FP*) versus the total protein concentration (P_T) using Equation (2):

$$F = F_L + \left(\frac{F_{LP} - F_L}{2L_T}\right)(L_T + P_T + K_D - \sqrt{(L_T + P_T + K_D)^2 - 4L_T P_T} \quad (2)$$

where *F* is the measured average fluorescence polarization, F_L is the fluorescence polarization of free labelled peptide, F_{LP} is the maximum fluorescence polarization of the peptide–protein complex and L_T represents the total labelled peptide concentration.

Fluorescence polarization assays, competition

Protein binding by different compounds was measured by incubating different concentrations of each ligand (5-fold dilutions, ranging from 10 to 250 μM) with 0.26 μM F-pHDAC4 and 20 μM of MEF2A or MEF2D. Unlabelled pHDAC4 peptide (3-fold dilutions ranging from 1 to 100 μM) was used as positive control. Each mixture (100 μl) containing 10 mM HEPES, 200 mM NaCl, 1 mM DTT, 1 mM EDTA, pH 7.7, 0.26 μM F-pHDAC4, 20 μM of protein and the ligand of interest was transferred into black 96-well microplates (Optiplate, PerkinElmer) and incubated at room temperature for 1 and 24 h. Controls samples without proteins and without ligands were also prepared to estimate fluorescence of displaced and bound probe, respectively. Polarization signals were recorded at 25°C using an EnVision Multlabel Plate Reader (PerkinElmer) as described above. The average fluorescence polarization values of at least three independent experiments were plotted as a function of ligand concentration. Equation (3) was applied to determine the half-maximal inhibitory concentration (IC_{50}) of each molecule:

$$F = F_L + \frac{(F_{LP} - F_L)}{1 + \frac{I}{IC50}} \quad (3)$$

where *F* is the measured average fluorescence polarization, F_L is the fluorescence polarization of free labelled peptide, F_{LP} is the maximum fluorescence polarization of the peptide–protein complex and *I* is the concentration of the compound. Finally, the inhibition constants (K_i) were calculated using Equation (4):

$$K_i = \frac{IC50}{1 + \frac{L_T}{K_D}} \quad (4)$$

where L_T is the concentration of the labelled peptide and K_D is the dissociation constant of the labelled peptide F-pHDAC4 for MEF2A or MEF2D. All fluorescence polarization data were analysed using GraphPad Prism software.

Glutathione *S*-transferase (GST) pulldown and co-immunoprecipitations

GST-MEF2D (1–190) was produced in BL21-DE3 competent cells as previously explained (15) and used as a bait. NIH-3T3 HDAC4/TM cells were lysed with a hypertonic buffer (25 mM Tris–HCl pH 7.5, 100 mM NaCl, 50 mM KCl, 5 mM $MgCl_2$, 0.5% NP40, glycerol 10%, PIC 100×, PMSF 100×) Lysates were next incubated with 2 μg of GST-MEF2D and with HDAC inhibitors for 3 h at 4°C. For co-immunoprecipitation experiments, cells were lysed in a hypotonic buffer (20 mM Tris–HCl, pH 7.5; 2 mM EDTA; 10 mM $MgCl_2$; 10 mM KCl; and 1% Triton X-100) supplemented with protease inhibitors. For each immunoprecipitation 1.5 μg of HDAC4 antibody or IgGs were used.

RNA extraction and quantitative qRT-PCR

Cells were lysed using Tri-Reagent (Molecular Research Center). 1.0μg of total RNA was retro-transcribed by using 100 units of M-MLV Reverse transcriptase (Life Technologies) in the presence of 1.6 μM oligo(dT) and 4μM Random hexamers (Euroclone). qRT-PCRs were performed using SYBR green technology (KAPA Biosystems). Data were analyzed by comparative threshold cycle (delta delta Ct ΔΔCt) using *HPRT* and *GAPDH* as normalizer. The list of the primers used for qRT-PCR and ChIP-qPCR was previously published (19).

RNA-seq analysis

SK-UT-1 cells were lysed using Tri Reagent (Molecular Research Center). Total RNA was treated with DNAse I (NEB) and purified with RNA Clean & Concentrator (Zymo Research). RNA-seq library preparation and sequencing were performed at BMR-Genomics (Padua, Italy) following Illumina specifications. Quality control for raw sequencing reads was performed with programs FastQC (v0.11.9) (www.bioinformatics.babraham.

ac.uk/projects/fastqc/) and MultiQC (v1.09) (29). Transcript quantification was conducted with Salmon (v1.4.0) (30) on human transcriptome GRCh38 Ensembl version 100 (gene set patch level 13). Transcript quantifications were imported into R (v4.0.3) running Bioconductor (v3.11) for downstream analysis with tximeta (v1.6.3) (31) and summarized at the gene level. Principal component analysis was carried out with the plotPCA function from the DESeq2 package (v1.28.1) (32). Genes with a raw counts mean <64 between each condition replicates were removed from the analysis. Differential expression analysis was performed using DESeq2 with Wald test for significance. We adjusted for multiple hypothesis testing by employing Benjamini-Hochberg correction at a false discovery rate (FDR) of 0.05. Genes reported significantly by DESeq2 with an absolute fold change >2 were considered as differentially expressed. Genes were annotated with package AnnotationHub (v2.20.2) utilizing Ensembl annotation 100 data.

Normalization within differential expression analysis was run on the full dataset, including three samples treated with a derivative of NKL54 lacking the trifluoromethyl group (NKL22) (9). As only one time point was available for this inhibitor, NKL22 treatment was excluded from the downstream analyses. Plots were generated with ggplot2 (v3.3.3). Venn diagrams were created with VennDiagram (v1.6.20) or with the Venn diagram tool by the bioinformatics and evolutionary genomics group at VIB/Ghent University (http://bioinformatics.psb.ugent.be/webtools/Venn/). Functional annotation was performed on KEGG, Reactome slimGO and Gene Ontology databases with ClusterProfiler (v3.16.1) and ReactomePA (v1.32.0), respectively (19,33). GO term and pathway analysis results are reported at an FDR of 0.05.

ChIP, library construction and ChIP-seq data analysis

Chromatin was obtained from SK-UT-1 cells, 14 h after DMSO or NKL54 (5μM) treatment and immunoprecipitated with 2 μg of anti-H3K27ac, 3 μg of anti-MEF2D antibody, 4 μg of anti-HDAC4 or anti-HDAC9 antibodies or control IgG, as previously described (34). Three independent biological replicates were pulled according to BLUEPRINT requirements and 5 ng of total DNA were used to prepare ChIP-seq libraries, according to TruSeq ChIP Sample Preparation guide (Illumina). Libraries were sequenced on the Illumina HiSeq 2000 sequencer. The ShortRead R/Bioconductor package was used to evaluate the quality of sequencing reads and Bowtie 2 was used to align them to NCBI GRCh38 human genome reference. Peak calling and gene annotations were performed as previously described (19). ChIP-seq replicates were compared using the Irreproducibility Discovery Rate (IDR) framework (35), with the MACS2 narrow peaks as input and applying the following settings: –input-file-type narrowPeak, –rank signal.value, –output-file-type narrowPeak. The IDR reported peaks were used for further processing. Gplots, BiomaRt, and Gviz R/Bioconductor packages and the DeepTools suite were used to generate peak heatmaps and for the visualization of genomic loci. DeepTools was also used for generating the correlogram showing the genome-wide Spearman correlation between the ChIP-seq replicates (average scores per genomic bin (10kb)).

Statistics

For experimental data Student t-test was employed. Mann–Whitney test was applied when normality could not be assumed. We chose $P < 0.05$ as the statistical limit of significance. For comparisons between more than 2 samples, the Anova test was applied coupled to Kruskal–Wallis and Dunn's Multiple Comparison Test. We marked with * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Unless otherwise indicated, all the data in the figures were represented as arithmetic means ± the standard deviations from at least three independent experiments.

RESULTS

NKL54 induces cell death by apoptosis in LMS cells

LMS is a rare and aggressive tumor that has some smooth muscle features and accounts for ∼10% of adult STS (12,13). The therapeutic outlook for advanced LMS has not improved over the past decades, thus new approaches are urgently needed. The MEF2/Class IIa HDACs axis is dysregulated in a constant percentage of LMS. Furthermore, genetic ablation of this axis results in impaired cell growth with induction of cell death (19,36). These properties make this axis a promising therapeutic target for the treatment of LMS. For these reasons, we investigated previously characterized small molecules inhibitors of MEF2 class IIa HDACs interaction (BML-210 and NKL54) (9), for their ability to affect proliferation of LMS cells. BML-210 and NKL54 are pimelic diphenylamides belonging to the benzamide group. These compounds were initially identified and characterized as selective class I HDACs inhibitors (37).

BML-210 and NKL54 were compared for their ability to suppress LMS cell proliferation with the class IIa-selective HDAC inhibitor TMP195 and the pan-HDAC inhibitor SAHA (suberoylanilide hydroxamic acid), a zinc-chelator (Supplementary Figure S1A) (38,39). Two different LMS cell lines DMR and SK-UT-1 were used. SAHA showed the strongest antiproliferative activity with an IC_{50} of 2.7 μM in both cell lines (Supplementary Figure S1B/C). TMP195 only slightly affected the proliferation of LMS cells with an IC_{50} of 50 μM. NKL54 was more potent than BML-210 in both cell lines with an IC_{50} of 7.9 and 10.4 μM compared to the values of 20.2 and 18.3 μM of BML-210 (Supplementary Figure 1B/C). Next, we examined the induction of cell death. We chose NKL54 for its stronger antiproliferative activity compared to BML-210, and cell death was assessed by propidium iodide positivity and caspase activity using the DEVDase assay. Since the knock-out of HDAC9 in SK-UT-1 cells increases FAS expression and susceptibility to cell death, we evaluated the contribution of the extrinsic apoptotic pathway in NKL54-induced cell death. LMS cells expressing the inhibitor of DISC-activation FLIPs (the short isoform of CFLAR/FLIP) were used (40). Dose-dependent studies showed that NKL54 induces cell death in LMS cells, in part via the extrinsic pathway (Supplementary Figure 1D). Caspase activation demonstrates induction of apoptosis and dependence on the extrinsic pathway (Supplemen-

tary Figure 1E). Finally, NKL54 treatment increased FAS mRNA levels in LMS cells (Supplementary Figure 1F). In conclusion, NKL54 can induce apoptosis in LMS cells. This cell death response is characterized by the upregulation of FAS, as was also observed in cells null for HDAC9 (19).

Identification and characterization of new PAOA derivatives

To identify new compounds with improved ability to disrupt the interaction between MEF2 and class IIa HDACs, we synthesized a series of small molecules resembling NKL54 structure (MC2983, MC2984, MC2985 and MC2991), as shown in Figure 1A. MC2983 and MC2991, which present the amide group in an inverted position, should be less efficient as zinc chelators, but still able to interact with MEF2s similarly to the other compounds.

To confirm this hypothesis, a virtual screening was performed using the structure of MEF2A. The binding affinity of the different compounds to the hydrophobic groove of MEF2s was compared. For this purpose, the program PLANTS with the PLP scoring function was used (41). Docking energy analysis indicates that the new compounds should be able to interact with the hydrophobic groove of MEF2 with similar potency to NKL54 (Figure 1A and Supplementary Table S1). Indeed, NKL54 is estimated to be the most potent with a docking energy of −81.43 Kcal/mol. Among the four new compounds, MC2983 is predicted to be the most active, with a docking energy (−80.25 kcal/mol) slightly lower than that of NKL54. Interestingly, MC2991, the compound most divergent to the reference BML-210, is estimated to be the least potent of the series. In fact, MC2991 is the only compound that shows a docked conformation that does not fit the hydrophobic groove of MEF2 (Figure 1B and Supplementary Figure S2A).

Next, the different compounds were tested for cell death induction. Three different LMS cell lines were used: SK-UT-1, DMR and SK-LMS-1. SAHA is the reference for a pan-HDAC inhibitor and zinc chelator. The results were comparable in the three cell lines, with SK-LMS-1 cells showing some resistance to cell death, as observed previously (14). Compounds MC2983 and MC2991, which are structurally related to NKL54 but should not act as zinc chelators, were significantly much less effective in inducing cell death (Figure 1C–E).

The different ability of the tested compounds to inhibit class I and IIb HDACs was confirmed *in vivo*. Apart from MC2985, which inhibits lysine deacetylase activities only at high concentrations, all compounds capable of inducing cell death also inhibited KDACs (Figure 1F). As expected, MC2983 and MC2991 were inactive in this assay. SAHA and TMP195 were used as positive and negative controls, respectively.

To confirm the ability of NKL54, MC2984 and MC2985 to act as epigenetic drugs, we examined the levels of histone H3 lysine 9 acetylation (H3K9ac) by immunoblot. SAHA triggered an increase in H3K9 acetylation within 15 min of treatment (Figure 1G). NKL54 and MC2984 also increased H3K9 acetylation, albeit with lower efficiency, compared with SAHA. The increase in H3K9ac in response to MC2985 treatment was very small, almost undetectable. All these compounds and BML-210 also increased H3K27ac levels with similar kinetics and potency as for H3K9ac (Supplementary Figure S2B). Overall, the modulation of H3K9ac and of H3K27ac levels in response to the different compounds confirms the inhibitory potency observed with the KDAC assay *in vivo*. Finally, we compared the *in vitro* inhibitory activity of SAHA and of PAOA derivatives against different purified HDACs (NKL54 was chosen as an example). The inhibitory activity of NKL54 was specific to HDAC1/2/3 (Figure 1H). SAHA confirmed the broader effect by inhibiting HDAC1/2/3/6 and with lower potency also HDAC8 (Figure 1I). IC50s are shown in Supplementary Table S2. Curiously, NKL54 can inhibit HDAC4 at high concentrations (IC_{50} 35μM), possibly due to the trifluoro group (42).

The PAOA derivatives do not unleash the interaction between MEF2 and class IIa HDACs

We have shown that the PAOA derivatives, that induce cell death in LMS cells, can act as HDAC inhibitors. However, the same derivatives could also affect the action of MEF2 by binding its hydrophobic groove and abolishing the inhibitory influence of class IIa HDACs. To verify the possible dual action of these compounds, we set up an *in vitro* fluorescence polarization assay (FP). By using a fluorescent labelled HDAC4-derived peptide (F-pHDAC4), capable of binding the hydrophobic groove of MEF2, we could assess the ability of the different compounds to unleash the interaction between MEF2 and class IIa HDACs. Recombinant MEF2A and MEF2D were produced and purified to homogeneity. First, the ability of the fluorescent labelled HDAC4-derived peptide (F-pHDAC4) to bind the MEF2 proteins was tested using a direct FP binding assay. MEF2A and MEF2D show similar affinity for pHDAC4 with a K_D of 16–17 μM (Figure 2A and B). Next, a FP competition assay was used to assess the ability of the different compounds to disrupt the interaction between MEF2A with F-pHDAC4 (Figure 2C). The same assay was performed with MEF2D (Figure 2D). We used an unlabelled HDAC4 peptide (pHDAC4) as a positive control. Although the affinity of the F-pHDAC4 probe does not allow to accurately assign compounds with binding affinity below 16μM, none of the tested PAOA derivatives was able to displace F-pHDAC4 from MEF2A or MEF2D, even when tested at concentrations ten times higher (250μM) the affinity of the probe. In contrast, pHDAC4 efficiently competed with F-pHDAC4 for MEF2A or MEF2D binding, and pHDAC4 released F-pHDAC4 within 1 h, indicating that binding was dynamic. We also investigated whether the PAOA derivatives could compete with F-pHADC4 during prolonged incubation times. However, longer incubation (24 h) did not disrupt the interaction between MEF2A or MEF2D with F-pHADC4 (Figure 2E and F). The inability of these compounds to compete for the binding of MEF2 with class IIa HDACs was also verified by a GST pull-down assay with recombinant MEF2D and HDAC4-GFP. Recombinant GST-MEF2D (2 μg) was incubated with cell lysates from NIH3T3 cells overexpressing HDAC4 with mutations in the 14–3–3 binding sites. This mutant cannot be phosphorylated and exported from the nucleus, increasing the pool of HDAC4 available for MEF2D binding (Figure 3A

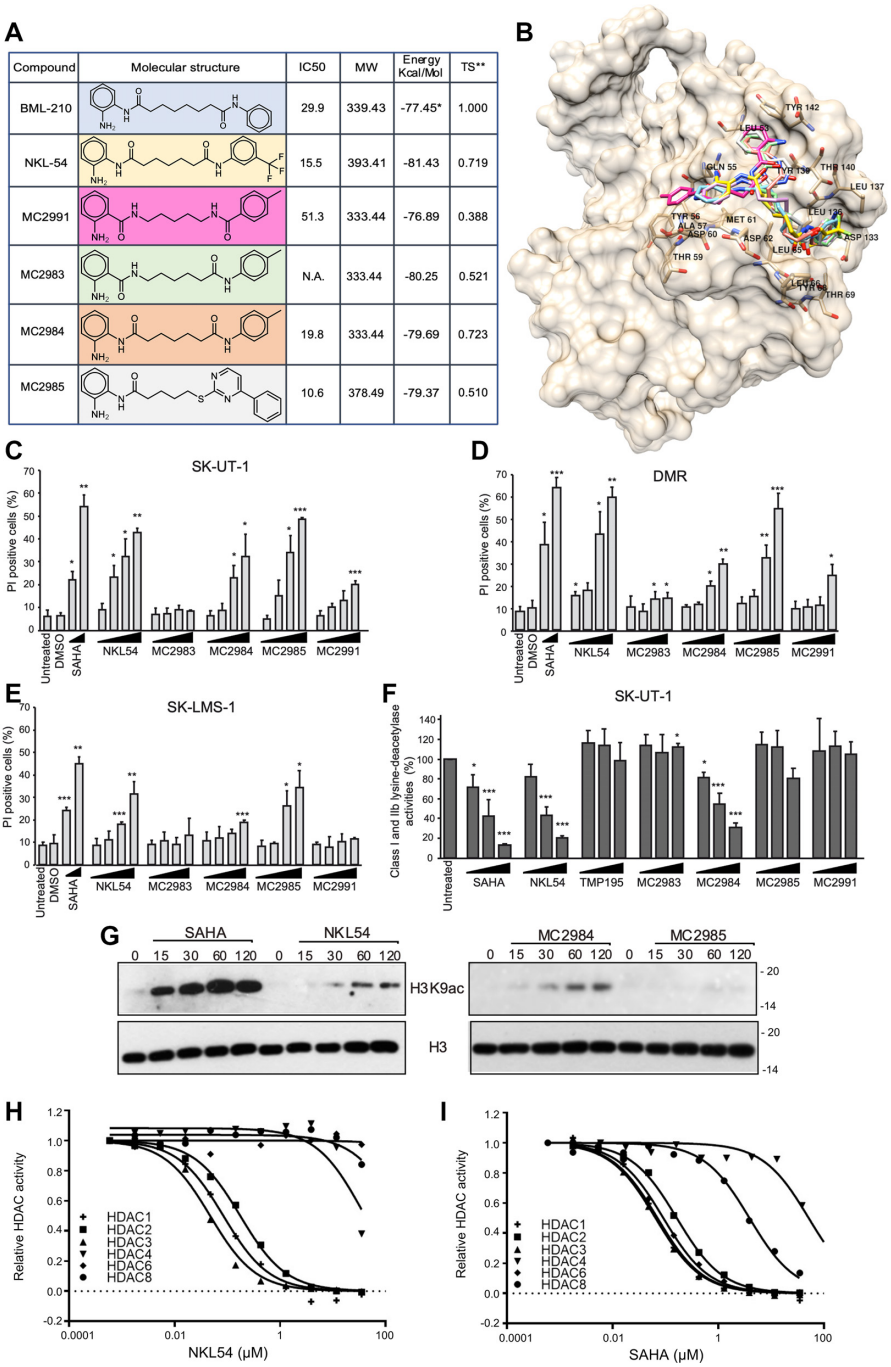


Figure 1. Characterization of new PAOA derivatives for mechanisms of action and anti-proliferative activities in LMS cells. (**A**) Characteristic and chemical structures of different PAOA derivatives used in this study in comparison with BML-210 and NKL54. The IC50 in SK-UT-1 cells is show as well as the Plants/PLP docking energies for the compound assayed in this study. The Tanimoto coefficient is reported for all the compounds versus the reference BML-210. *Energy calculated on the re-docked ligand extracted from the 3MU6 complex available from PDB. **Tanimoto similarity index calculated with radius 2 circular Morgan fingerprints (73) in a python script (74) using the RDKit library (75). (**B**) BML-210, NKL54, (MC2983, MC2984, MC2985 and MC2991 docked conformations in the MEF2A (pdb entry code 3MU6) hydrophobic groove. Plants/PLP combination as implemented in 3d-qsar.com was used to dock the compounds. Lowest energy docked conformations were imported in UCSF Chimera along with the cleaned minimized MEF2A protein and co-crystallized BML-210, for binding mode inspection and comparison. (**C**) Analysis of cell death in SK-UT-1 cells as percentage of PI positive cells treated with the indicated compounds. Cell death was scored after 48 h from treatments. Data are from three independent experiments, + S.D. (**D**) Analysis of cell death in DMR cells as percentage of PI positive cells treated with the indicated compounds. Cell death was scored after 48 h from treatments. Data are from three independent experiments, + S.D. (**E**) Analysis of cell death in SK-LMS-1 cells as percentage of PI positive cells treated with the indicated compounds. Cell death was scored after 48 h from treatments. Data are from 3 independent experiments, + S.D. (**F**) Lys-deacetylase activity as measured from SK-UT-1 cells treated with increasing concentrations of the indicated compounds [1, 5, 10 μM]. Data are from three independent experiments, + S.D. (**G**) Immunoblotting analysis of H3K9 acetylation levels in SK-UT-1 cells treated with SAHA [2.5 μM], NKL54, MC2984 and MC2985 [5μM] for the indicated minutes. (**H**) Dose response curves of NKL54 against a panel of indicated HDAC isozymes using standard enzyme activity assays. (**I**) Dose response curves of SAHA against a panel of indicated HDAC isozymes using standard enzyme activity assays.

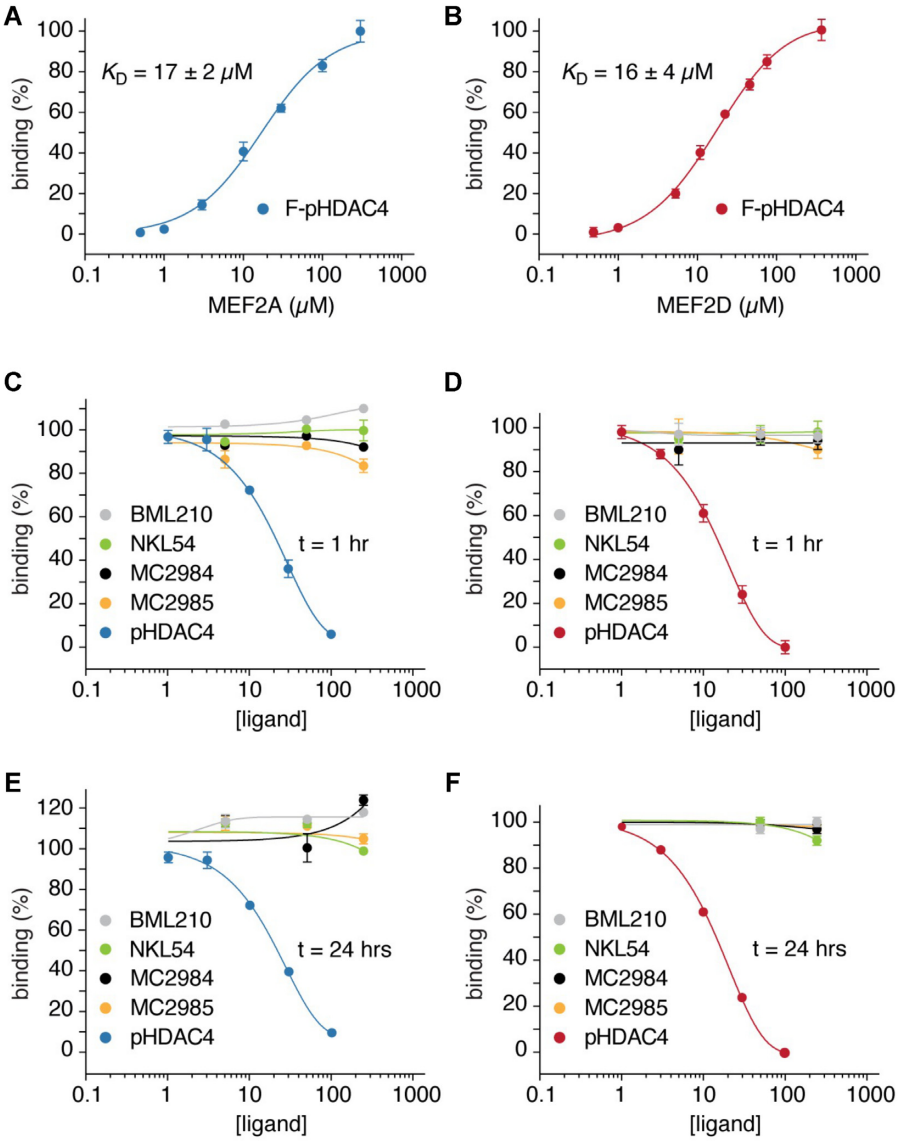


Figure 2. NKL54 does not compete for the binding between MEF2 and class IIa HDACs. (**A**) MEF2A binding curve (0.5–300 μM) titrated to fluorescently labelled peptide F-pHDAC4 (0.26 μM). The average fluorescence polarization values of at least three independent experiments ± SD were plotted as a function of ligand concentration. (**B**) MEF2D binding curve. Experiments were performed as in panel A. (**C**) Protein binding by the indicated compounds was measured by incubating different concentrations of each ligand (5-fold dilutions, ranging from 10 to 250 μM) with 0.26 μM F-pHDAC4 and 20 μM of MEF2A. Unlabelled pHDAC4 peptide (3-fold dilutions ranging from 1 to 100 μM) was used as positive control. Titrations were performed at room temperature for 1 h as indicated. The average fluorescence polarization values of at least 3 independent experiments ± SD were plotted as a function of ligand concentration. (**D**) MEF2D binding by the indicated compounds as performed in panel C. (**E**) MEF2A binding by the indicated compounds as performed in panel C, with titrations measured at 24 h. (**F**) MEF2A binding by the indicated compounds as performed in panel E.

and Supplementary Figure S2C). Finally, immunofluorescence analysis confirmed that PAOA derivatives cannot interfere with the ability of MEF2D to cause nuclear accumulation of HDAC4 (Supplementary Figure S3).

NKL54 and SAHA influence the expression levels of MEF2D and HDAC7

Pan-HDAC inhibitors such as hydroxamates (TSA or SAHA) and benzamides (entinostat and mocetinostat) affect the stability of HDAC7 (43). Consequently, these compounds might indirectly upregulate MEF2-dependent transcription by reducing the expression of class IIa HDACs. To verify this hypothesis, we analyzed the expression levels of the different members of the class IIa HDACs family in SK-UT-1 cells in response to treatment with SAHA or NKL54. NKL54 was chosen as a prototype for the different PAOA derivatives. We also analyzed the levels of MEF2A and MEF2D, which are the two major MEF2 family members expressed in this cell line. Figure 3B shows that HDAC7 is downregulated starting 12 h after addition of SAHA. Similarly, NKL54 triggers the downregulation of HDAC7. HDAC4 and HDAC9 are also downregulated at later time points (starting from 24 h), but only in NKL54-treated cells. In contrast, MEF2D expression levels were increased after treatment with the two HDACIs. These results were con-

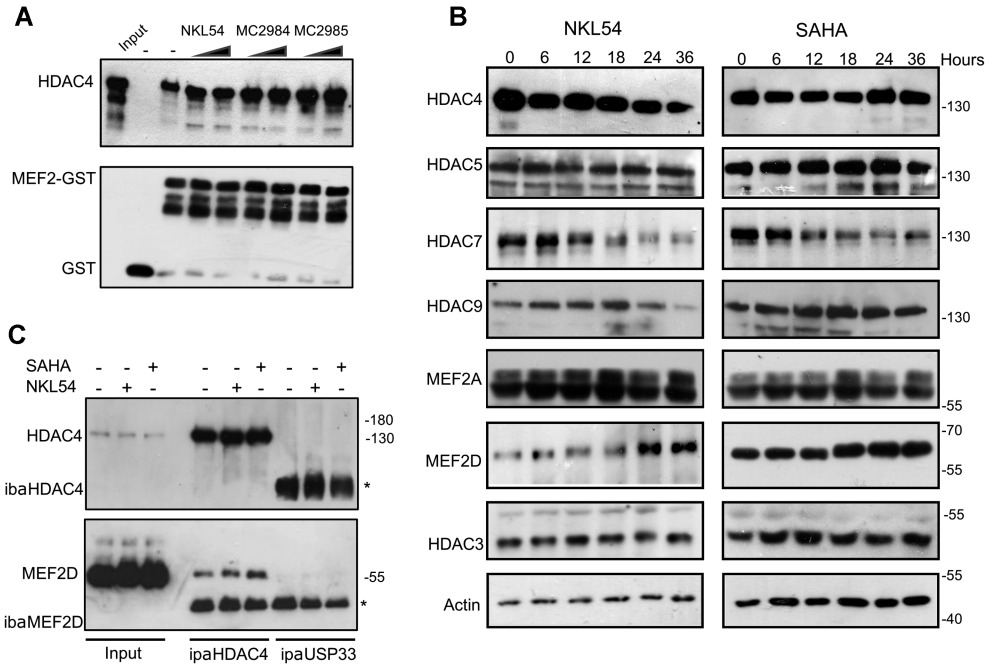


Figure 3. Regulation of MEF2-class IIa HDACs axis by PAOA-derivatives. (**A**) GST pull-down assay, using recombinant MEF2D (1–190) or GST as control. Purified GST or GST-MEF2D recombinant proteins (2 μg) were incubated with cellular lysates obtained from NIH3T3 cells overexpressing HDAC4 mutated in 14–3–3 binding sites. Two different concentrations [14 and 42 μM] were used. Immunoblots were performed to visualize HDAC4 or recombinant GST. (**B**) SK-UT-1 cells were treated for the indicated times with NKL54 [5 μM] or SAHA [2.5 μM]. Cellular lysates were generated and immunoblot performed using the indicated antibodies. Actin and Histone 3 (H3) were used as loading control. (**C**) Lysates from SK-UT-1 cells treated for 12 h with DMSO or with SAHA [2.5 μM] or NKL54 [5μM] MEF2D-HDAC4 complexes were immunoprecipitated with antibodies against HDAC4 or USP33, as a control. Immunoblotting using an anti-MEF2D antibody was next used for the detection of the MEF2-HDAC4 complexes. Asterisks point to IGs.

firmed by a second immunoblot analysis and relative densitometric evaluations (Supplementary Figure S4A/B). When used at high concentrations, MC2984 and MC2985 also upregulated MEF2D expression and reduced HDAC7 levels (Supplementary Figure S2D).

This detailed characterization allowed us to select the best time point to verify the ineffectiveness of NKL54 in releasing the binding between MEF2D and HDAC4 *in vivo* as well. SK-UT-1 cells were treated with NKL54 or SAHA, as control, and 12 h later cellular extracts were generated and subjected to co-immunoprecipitation with an anti-HDAC4 antibody (Figure 3C Supplementary Figure S2C). This time point was chosen to preclude excessive variation in protein concentrations of the two targets. Consistent with the *in vitro* studies, NKL54 was ineffective in releasing MEF2D from HDAC4 binding *in vivo*.

Transcriptome remodelling as elicited by different HDAC inhibitors

To better classify the different HDACIs, the induced transcriptional changes were mapped by RNA-seq. A single concentration of the different compounds (2,5μM for SAHA and 5μM for all PAOA-derivatives), triggering moderate and comparable percentages of cell death at 36 h was used. Percentages of cell death were: 23.2%± 1.92 for SAHA, 21.4% ± 3.44 for MC2984, 27.0% ± 4.0 for MC2985 and 39.00% ± 2,44 for NKL54. In untreated cells the percentage of cell death was 7.00%± 2.24. We also used TMP195 (5 μM) a poor inducer of cell death (13.0% ± 1.0). RNA was isolated at 6 h, to map early changes, and after 24 h from treatments, to investigate delayed modulations in gene expression. The PCA (principal component analysis) shows the high reproducibility of the three biological replicates analyzed (Figure 4A). TMP195 induced few modifications in the transcriptome. Twenty four hours of treatment with MC2984 elicited modifications of the transcriptome like the early responses to SAHA and NKL54 treatments, thus suggesting a similar but delayed effect of this PAOA-derivatives. Finally, later changes in gene expression more manifestly separate NKL54 from SAHA (Figure 4A). The top upregulated and downregulated genes for each treatment are listed in Supplementary Tables S3 and S4.

To further validate these observations, we compared genes upregulated and downregulated in response to the different compounds using SAHA as a reference. Venn diagrams show that 66.8% of genes regulated at 6 h by NKL54 are shared with SAHA (Supplementary Figure S5). After 24 h 82.5% of genes modulated by SAHA were similarly modulated by NKL54. In parallel, the number of NKL54 specific genes increased up to 41.3%. Only 19.4% of genes ($n = 184$) were specific for MC2984 at 6 h, a percentage that decreased to 13.7% after 24 h of treatment ($n = 246$). MC2985 modulated very few genes at both 6 and 24 h ($n = 218$ and 317, respectively), of which only 4.6% (6 h) and 9.5% (24 h) are specific. TMP195 showed the most divergent profile respect to SAHA with only 48.4% of genes in common at 6 h. Interestingly, this percentage increased up to 60,5% after 24

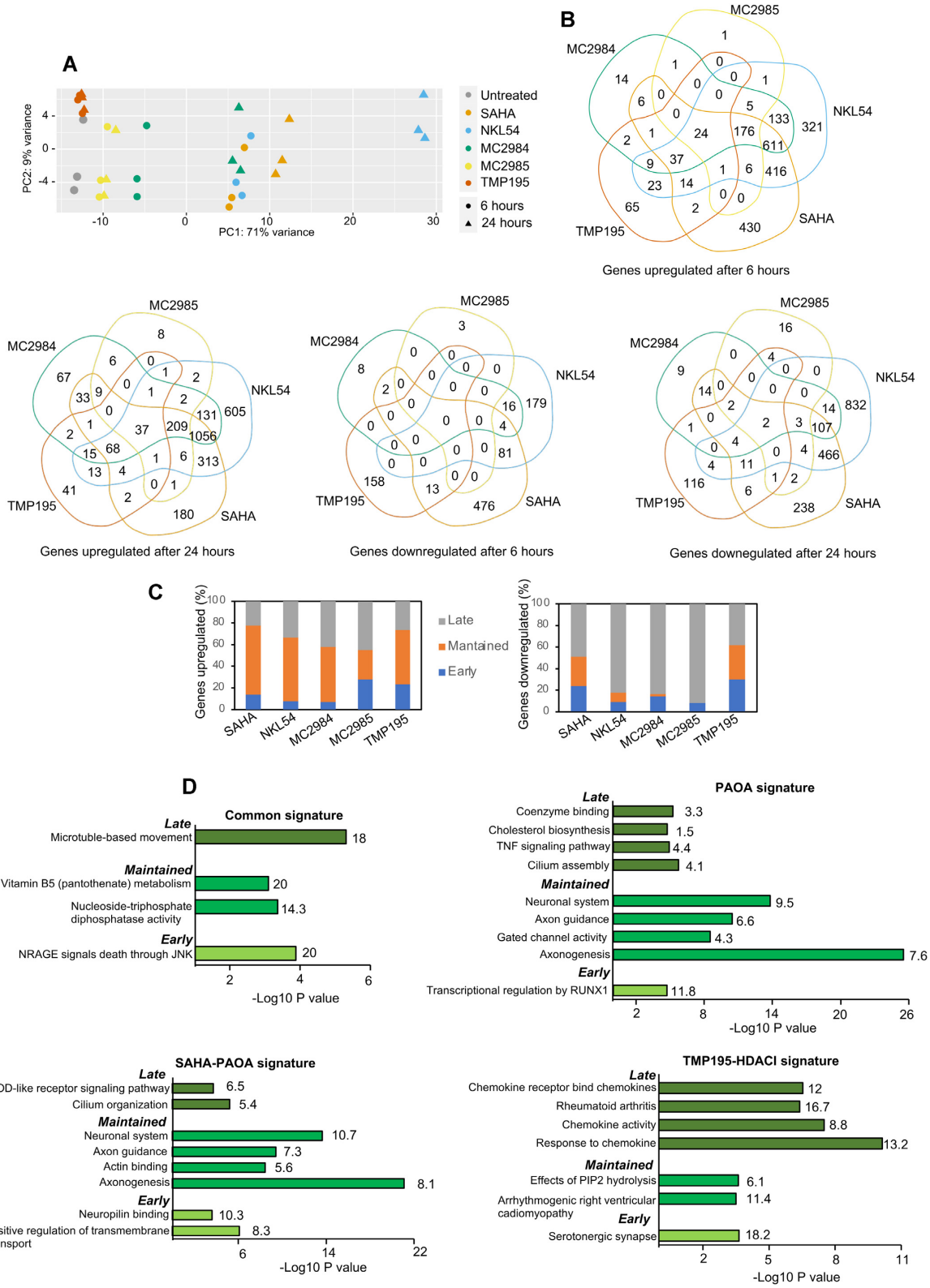


Figure 4. Comparative transcriptomic analysis in response to the different HDACIs. (**A**) PCA analysis performed on the expression profiles of the indicated treatments at shown times in SK-UT-1 cells. (**B**) Venn diagrams showing the number of transcripts commonly and differentially upregulated or downregulated between the different HDACIs at the indicated hours. (**C**) Percentage of genes upregulated or downregulated by the different HDACIs at 6 h (early genes) at both 6 and 24 h (maintained genes) and at 24 h (late genes). (**D**) Bar plots of the ClusterProfiler-ReactomePA most significantly enriched functional terms according to the GO: Biological Process, GO: Molecular Function, Reactome or KEGG databases. The analysis was performed for the indicated groups of signatures, retaining the top terms for each functional database. Numbers to the right of the bars represent the percentage of significantly enriched genes found within each category.

h of treatment, thus indicating common adaptive responses (Supplementary Figure S5).

SAHA, NKL54 and MC2984 upregulated a larger number of genes compared to TMP195 and MC2985 (Figure 4B and Supplementary Figures S5 and S6A). For SAHA and NKL54 the effect on gene transcription was rather stable through the time, with respectively 64, and 61% of genes upregulated at both time points (Figure 4C). There are also genes which expression was upregulated at 24 h but not at 6 h. The percentage of these genes was lower for SAHA (22.6%) and TMP195 (26.4%), intermediate for NKL54 (33.6%) and greater for MC2984 and MC2985 (42 and 45%, respectively). These different delayed responses could reflect different kinetics/modifications of the compounds or different adaptive responses engaged. Among the early DEGs (differentially expressed genes), the downregulated genes were a minor fraction (Supplementary Figure S6): 13.6% (NKL54), 2.8% (MC2984), 1.5% (MC2985) and 25% (SAHA).

Again, the response to TMP195 was highly divergent, with 45.9% of DEGs that were downregulated. Interestingly, in all treatments, the down-regulated genes were modulated in higher percentages at 24 h (Figure 4C and Supplementary Figure S6B). This feature is particularly evident for the PAOA derivatives. Here, the few overlaps observed between 6 and 24 h suggest that many genes downregulated at later time points may be indirect targets of these compounds.

Common biological responses modulated by the different HDACIs

The above analysis has shown that the various inhibitors trigger both common and specific transcriptional modulations. To define the respective cellular responses, we selected both compound-specific DEGs and DEGs common among different compounds. Common genes were grouped into three distinct categories: (i) the SAHA-PAOA signature, which includes genes regulated by SAHA and also by at least one PAOA derivative; (ii) the PAOA signature, which includes genes regulated by at least two different PAOA derivatives; and finally, (iii) the TMP195-HDACIs signature, which includes genes regulated by TMP195 and also by at least SAHA or a PAOA derivative. Gene signatures were also analyzed in terms of temporal regulation (Supplementary Figure S7A and B). The early response includes genes regulated at 6 but not at 24 h of treatments. The maintained response includes genes regulated at both 6 and 24 h Finally, the late response includes genes that are regulated at 24 h, but not at 6 h.

We used ClusterProfiler (33) and ReactomePA (44) to understand the functions of genes that are regulated under the different treatments. First, we evaluated upregulated genes. The top category of genes that are upregulated by all inhibitors is related to microtubules-based movements and represents a late response (Figure 4D and Supplementary Table S3). For the SAHA-PAOA signature, neuronal system, axon guidance, and axonogenesis are the most significant results. Since these genes belong to the maintained category, this response might reflect the stable release of a cell-lineage specific inhibitory influence of HDACs. These gene categories are also dominant for the PAOA signature. For both the SAHA-PAOA and PAOA signatures, early and late responses are less clearly defined and include both metabolic and differentiative responses. The TMP195-HDACIs signature is characterized by the chemokine activity/rheumatoid arthritis categories which included the late-responding genes (Figure 4D and Supplementary Table S5).

We also examined cellular responses that were switched off after HDACIs treatment. As expected from the general trends of downregulation (Supplementary Figure S7B), the most significant enrichments were found within the late responses. Interestingly, in the SAHA-PAOA and PAOA signature the most enriched categories were found in the context of chromatin organization (HAT acetylate histones, DNA packaging, Nucleosome assembly, DNA conformation change) possibly reflecting a compensatory response triggered by changes in chromatin dynamics, because of HDACs inhibition (Supplementary Figure S7C and Supplementary Table S6). Another, highly enriched downregulated category is M phase, a plausible consequence of cell-cycle arrest. As observed above for the upregulated genes, the TMP195-HDACIs categories show again specific features with an impact on the microenvironment. Indeed, the ECM genes are among the most highly enriched (Supplementary Figure S7C and Supplementary Table S6).

Specific biological responses engaged by the different HDACIs

Compound-specific DEGs represent a small percentage of regulated genes (Figure 4B). As described above for common DEGs, these signatures were analyzed in relation to the timing of regulation by dividing the DEGs into early, maintained, and late. Among the upregulated genes, significant categories were found for NKL54, SAHA, and TMP195 (Supplementary Figure S8A, Supplementary Table S7). The biological processes diverge greatly among the three compounds. SAHA derepresses genes related to cilia/flagellar organization and dynamics (cilium assembly organization, intraflagellar motility) as maintained and late response. NKL54 affects extracellular matrix dynamics, differentiation and metabolism as an early, maintained and late response. TMP195 shows early activation of chemokine expression and inflammatory responses that are later induced by the other HDACIs (Figure 4D). These selective influences on gene expression are indicative of the existence of distinct complexes containing class I or class IIa HDACs.

Among the down-regulated genes, again NKL54, SAHA and TMP195 achieve the most significant enrichments (Supplementary Figure S8B, Supplementary Table S8). Genes involved in the maintenance of chromatin homeostasis are strongly repressed by both SAHA and NKL54, as observed above for some common signatures. In addition, NKL54 significantly represses cell cycle-related genes. TMP195 shows a differential effect on the ECM, the microenvironment as an early response and, both as

an early and late response, on some genes controlling differentiating responses.

SAHA and PAOA derivatives boost the expression of BH3-only BCL2 family members

Induction of cell death by apoptosis characterizes the response to SAHA and PAOA-derivatives. BCL2 family members are master regulators of apoptosis, through control of mitochondrial outer membrane permeabilization (45). Frequently cell death signalling pathways control the expression of BCL-2 family members. All of the different compounds that trigger cell death in LMS affect expression of a group of BH3-only members. mRNA levels of *BMF*, *BIK* and *HRK* are dramatically increased in response to SAHA and PAOA (although MC2985 was less potent) (Figure 5A). *BBC3/PUMA* and *BCL2L11/BIM*, other BH3-only members, are less upregulated. TMP195, which is a very weak inducer of cell death only modestly and transiently increases *HRK* and *PMAP1/NOXA* mRNA levels.

SAHA and PAOA derivatives modulate the expression of members of the MEF2-HDAC axis

NKL54 and SAHA could indirectly affect the MEF2-HDAC axis by modulating the levels of MEF2 and HDAC family members (Figure 3B). Analysis of RNA-seq data showed that HDAC5 mRNA is slightly upregulated, whereas HDAC7 and HDAC9 are down-regulated (Figure 5B). Importantly, NKL54 exerts a persistent effect on HDAC9, whereas SAHA only transiently downregulates it. Accordingly, HDAC9 protein levels are significantly downregulated only in cells treated with NKL54 (Figure 3B). All other compounds only moderately affect the expression of class IIa HDACs.

When class IIa HDACs are generally downregulated in response to SAHA and NKL54, MEF2 family members are upregulated, particularly MEF2B and MEF2C (Figure 5B and C). The expression of other HDACs is only slightly affected by the investigated compounds. SAHA and the PAOA derivatives (NKL54 and MC2984) upregulate the mRNA levels of HDAC11, HDAC3, and to some extent HDAC1 (Figure 5D). To understand the impact of these regulations on the transcriptional activity of the MEF2-HDAC axis, we evaluated the TPM (transcript per million) measure for each family member. HDAC9 and HDAC7 yield the most highly expressed family members in SK-UT-1 cells (Figure 5E). Therefore, their down-regulation could strongly affect the total pool of class IIa HDACs available for suppression of MEF2 transcription. On the other hand, although MEF2C is strongly upregulated by SAHA, NKL54 and MC2984, because it is expressed at very low levels in SK-UT-1 cells, its impact on the overall MEF2 transcriptional output might be minimal (Figure 5C and F). In contrast, MEF2D is the most expressed paralog in this cell line (Figure 5C), and although it is much less upregulated compared with MEF2C, it could strongly influence the transcriptional landscape, as confirmed by the consistent increased protein levels (Figure 3B).

The expression of several lysine-acetyltransferases (KATs) is downregulated after inhibition of HDACs

A prominent response to HDACIs is the downregulation of genes encoding components of the KAT complexes, an evolutionary adaptive mechanism for buffering HDACs inhibition (46). Therefore, we examined the expression levels of various KATs in response to the treatments (47). With few exceptions, SAHA/NKL54/MC2984 trigger downregulation of several KATs. MC2985 and TMP195 are much less effective and, curiously, both promote downregulation of KAT2B (Figure 5G). In general, there is a strong correlation between the ability of the various compounds to increase histone acetylation and the repression of KAT expression.

NKL54 and the regulation of the MEF2-HDAC axis

NKL54 is not efficient in impeding the binding between MEF2 and class IIa HDACs, but could indirectly sustain MEF2-dependent transcription, through its influence on class IIa HDACs and MEF2D levels. MEF2A and MEF2D are the two major paralogs expressed in SK-UT-1 cells. Hence, we compared genes that were up and downregulated after silencing of these two MEF2 family members, that we defined previously (35), with genes that were modulated by the HDACIs under study. This comparison aimed to provide insight into the ability of NKL54 to affect the MEF2-HDAC axis. It is important to emphasize that in SK-UT-1 cells MEF2 act as transcriptional repressors at some loci and as transcriptional activators at others (35). Taking this into account, genes upregulated after treatment with the different compounds should be compared with genes upregulated after MEF2A/D silencing (abrogation of repression), and with genes repressed after MEF2A/D silencing (Figure 6A). In this last condition, the increased expression in response to HDACIs may depend on the augmented MEF2D levels. Certainly, the contribution of other TFs whose activities may be modulated by the different HDACIs cannot be excluded.

Venn analysis shows that MEF2-regulated genes have the best overlap with NKL54-regulated genes (Figure 6B and Supplementary Figure S9A). 29,1% of genes upregulated, and 19,3% of genes downregulated after MEF2A/D silencing are upregulated in response to NKL54. The other PAOA derivative (MC2984) is the second-best compound. 22,2% of up and 17,9% of downregulated genes after MEF2A/D silencing are upregulated by MC2984. Next, we examined the percentage of overlap between the genes regulated by the different compounds and the MEF2A/D signatures (Figure 6C and Supplementary Figure S9B). In this case, the best result was obtained for TMP195 with 17.4% identity among the up-regulated genes and only 2.5% among the down-regulated genes. This result is another confirmation of the existence of repressive MEF2 complexes in SK -UT-1 cells (35). In summary, this analysis shows that a small percentage of genes upregulated by NKL54, which has the broadest effect among the HDACIs tested, are under the regulation of MEF2A/D. In contrast, TMP195 has limited effects on overall gene expression, but a certain percentage of genes upregulated by this class IIa inhibitor are targets of the MEF2-HDAC axis. GSEA was applied to confirm these

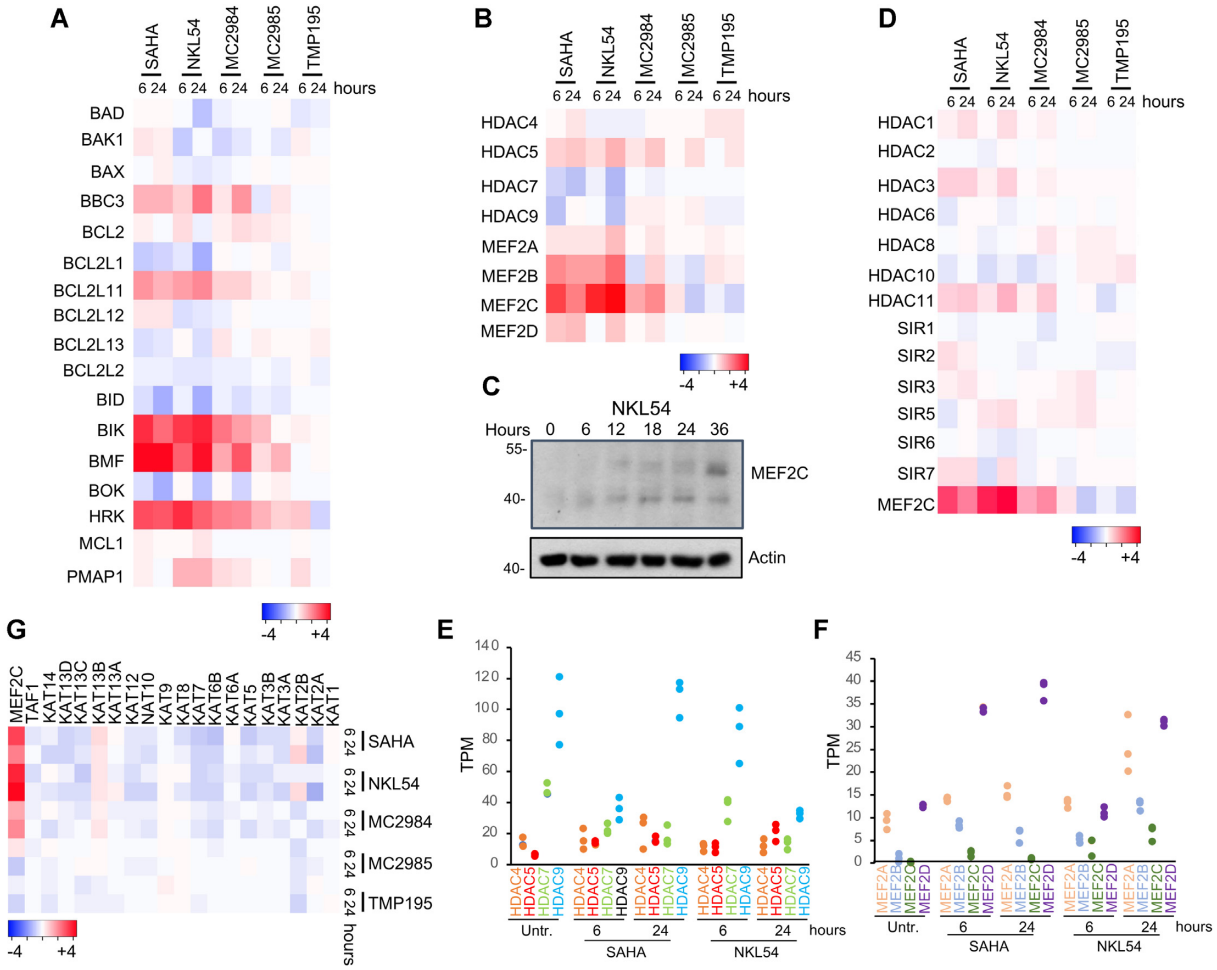


Figure 5. Influence of the different HDACIs on the expression levels of BCL2 family members, of the MEF2-HDAC axis and on KATs. (**A**) Heat-map reporting the expression levels (log_2 fold change relative to untreated cells) of the indicated BCL2 family members in response to the different HDACIs as indicated. (**B**) Like panel A, instead showing MEF2-HDAC axis components. (**C**) SK-UT-1 cells were treated for the indicated times with NKL54 [5μM]. Cellular lysates were generated and immunoblot performed using the anti-MEF2C antibody. Actin was used as loading control. (**D**) Heat-map reporting the expression levels (log2 fold change relative to untreated cells) of class I/IIb/III/IV HDACs family members in response to the different HDACIs as indicated. (**E**) Heat-map reporting the expression levels (log_2 fold change relative to untreated cells) of different KATs in response to the different HDACIs as indicated. (**F**) Class IIa HDAC family members expression in SK-UT-1 cells treated or not with SAHA and NKL54 for the indicated times. Expression values are shown in TPM (transcripts per million) calculated from a gene model where isoforms were collapsed to a single gene. (**G**) Like panel F, instead showing MEF2 family members.

results. Again, the PAOA derivatives (NKL54 and MC2984) achieved the best enrichments compared with MEF2A/D signatures (Figure 6D).

From a therapeutic perspective, it is important to understand whether reactivation of the MEF2A/D signature could be beneficial for LMS patients. Therefore, we investigated whether genes co-regulated by MEF2A/D and NKL54 might be important for LMS aggressiveness. We selected genes modulated by MEF2A/D and upregulated in LMS cells after NKL54 treatment. From this signature, which includes 123 genes, we selected genes that were upregulated in at least 10% of patients (Figure 6E, $n = 10$). We defined this group of genes as the MEF2-NKL54 signature. Interestingly, the BH3-only member *PMAIP1/NOXA* is included in this signature. Figure 6F shows that patients characterized by high expression of genes of the MEF2-NKL54 signature have a better prognosis.

NKL54-induced changes in H3K27ac genomic distribution

To better characterize the effect of NKL54 at the genomic level and its influence on gene expression, we examined the genomic distribution of H3K27ac by performing ChIP-seq. Two distinct sequencing experiments were done, each pooling at least two distinct biological replicates. We selected 14 h after treatment to limit the effect of NKL54 on the levels of class IIa HDACs. A Pearson's correlation coefficient (PCC) test on the two sequencing experiments was performed to assess the reproducibility of each ChIP-seq (48). The results were represented as a heatmap and show a very high reproducibility (Supplementary Figure S10). Treatment with NKL54 resulted in a higher number of IDR-defined H3K27ac peaks ($n = 72\,765$) compared to control ($n = 64\,232$). These peaks were particularly abundant in promoter regions near TSS and at distal intergenic regions (Figure 7A).

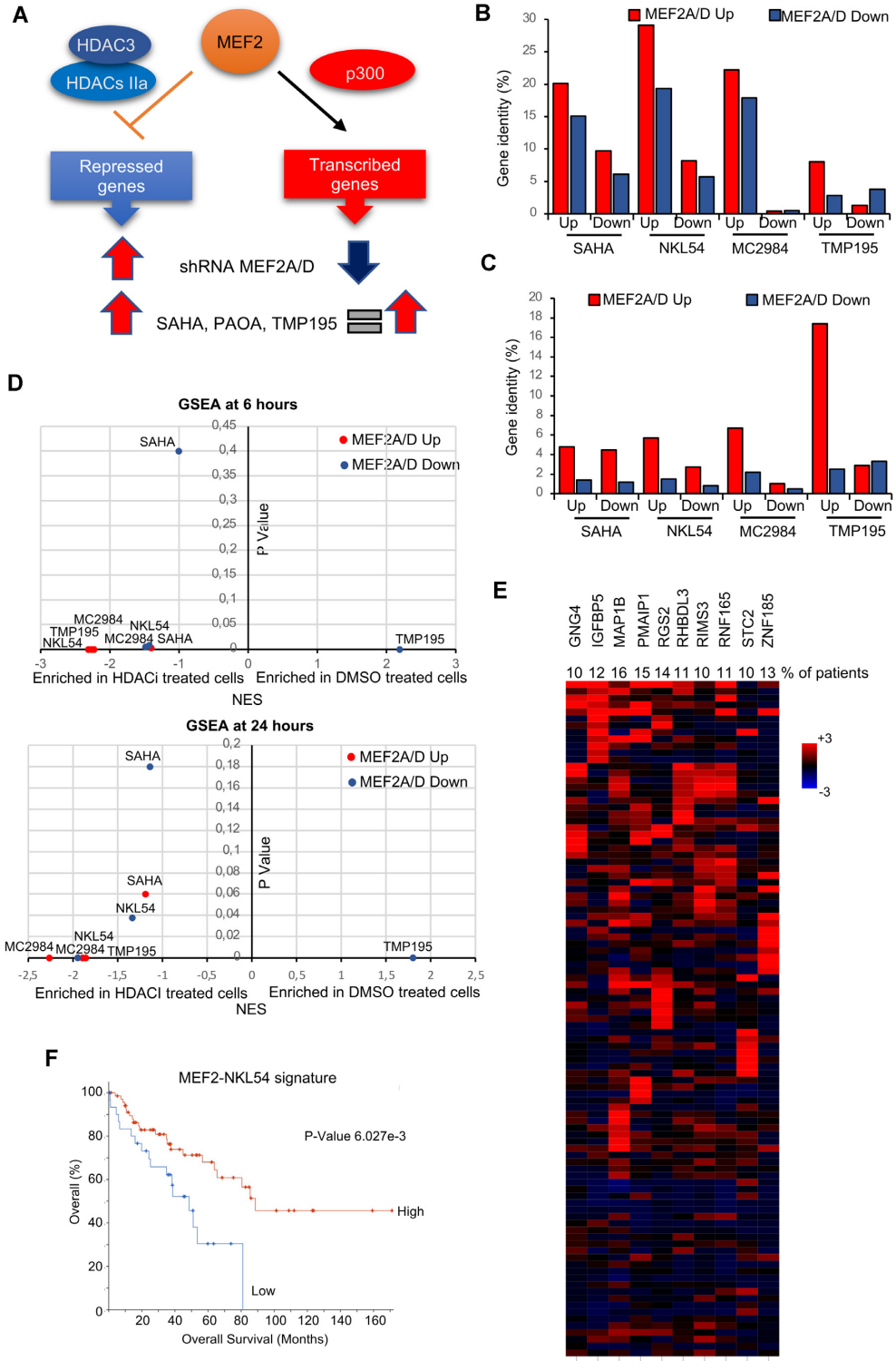


Figure 6. The impact of the different HDACIs on the MEF2 transcriptional activity. (**A**) Scheme comparing the effect of MEF2A/D silencing and HDACIs treatments on genes regulated by MEF2. (**B**) Percentage of identity among genes up or downregulated after MEF2A or MEF2D silencing and genes up or downregulated after treatments with the different HDACIs. (**C**) Percentage of identity among genes up or downregulated after treatments with the different HDACIs and genes up or downregulated after MEF2A or MEF2D silencing. (**D**) GSEA results displayed as the NES and the p value obtained by interrogating the transcriptome of MEF2A/D knocked-down SK-UT-1 cells with the transcriptomes of the same cells treated with the different HDACIs. (**E**) Oncoprint of mRNA expression variations for the indicated genes defined as MEF2-NKL54 signature. Data were obtained from the TCGA database and include RNAseq data of 100 patients with LMS. The heatmap shows the expression levels (z-score normalized log2 (FPKM) values) relative to diploid samples and was generated through cBioPortal (http://www.cbioportal.org). The percentages refer to the number of patients with *z*-score >2. (**F**) Kaplan–Meier plot showing the survival in percent of patients having expression of the 'MEF2-NKL54' gene signature (panel E) with at least one member with *z*-score >2 with respect to diploid samples, as represented by Kaplan–Meier plot. The graph was generated through cBioPortal (http://www.cbioportal.org).

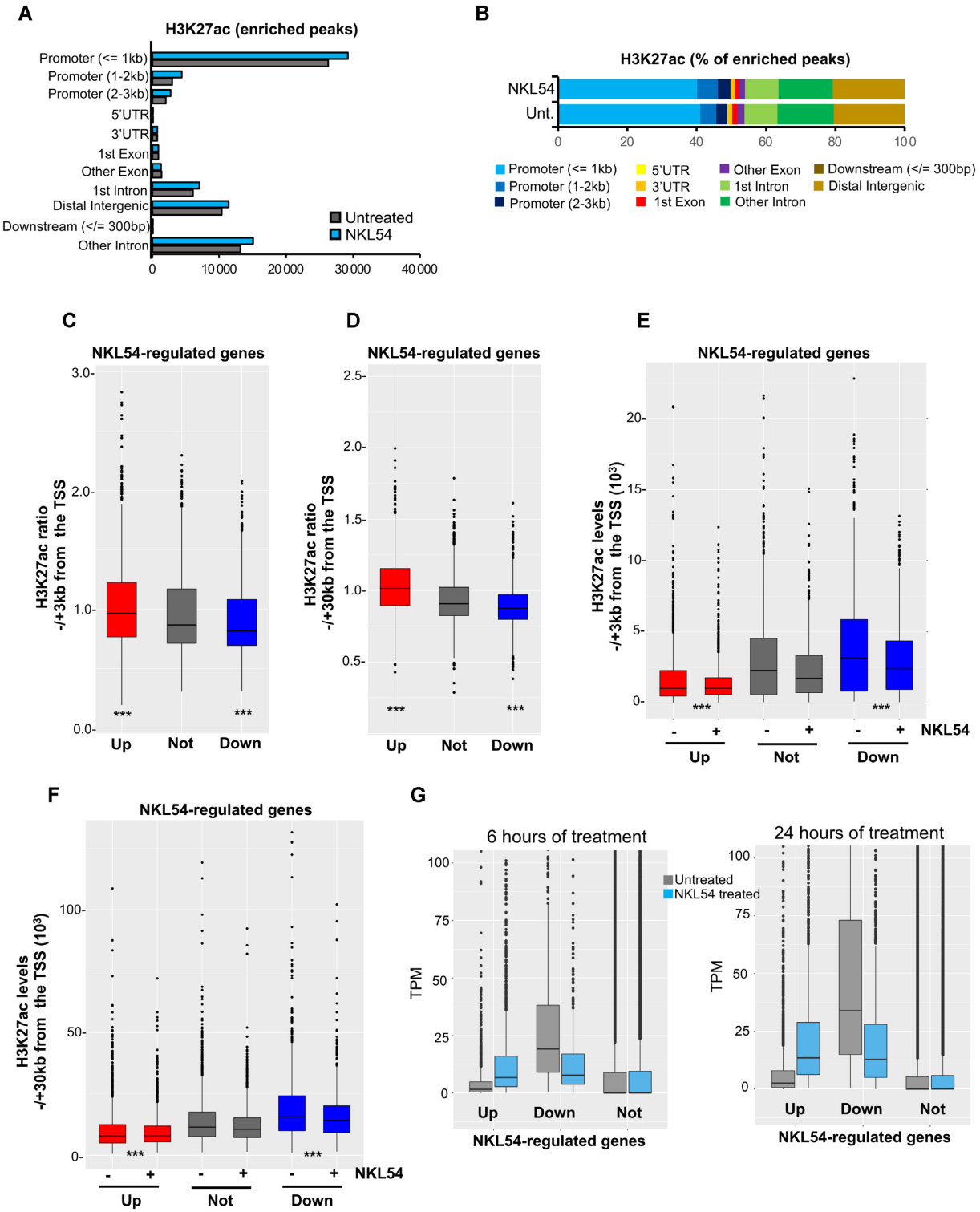


Figure 7. Variation of H3K27ac distribution at the genomic level. (**A**) Genomic distribution of the H3K27ac-enriched IDR-defined peaks identified by MACS2 in SK-UT-1 cells treated (n = 72 765) or untreated (n = 64 232) for 14 h with 5 μM NKL54. (**B**) As in panel A, with values represented as percentages. (**C**) H3K27ac ratio between NKL54 treated and untreated cells within ±3 kb from TSS. ChIP-seq data are from experiment 1. Genes not-regulated by NKL54 were selected based on having the lowest combined gene expression variations at 6 and 24 h from treatment. The boxes indicate the interquartile range with the center line representing the median value. The outliers are plotted as dots. (**D**) As in panel C, with H3K27ac ratio between NKL54 treated and untreated cells calculated within ±30 kb from TSS. (**E**) Overall acetylation level in the ± 3kb region centered on gene TSS of the indicated gene categories in presence or absence of NKL54. Boxes plotted as in panel C. (**F**) As in panel E with overall acetylation level measured within a ±30 kb region centered on gene TSS of the indicated gene categories in presence or not of NKL54. (**G**) TPM values are shown after treatment or not with NKL54 for the respective times. TPM measure was calculated from a gene model where isoforms were collapsed into a single gene. Significances were tested using the Mann–Whitney U test.

Experiments were done 14 h after treatment to limit the effect of NKL54 on the levels of class IIa HDACs. Treatment with NKL54 resulted in a higher number of H3K27ac peaks ($n = 110\,900$) compared to control ($n = 83\,419$). These peaks were particularly abundant in promoter regions near TSS and at distal intergenic regions (Figure 7A). HDACIs increased the number of enriched peaks in all genomic regions, with a slightly higher percentage in promoter regions within 1–3 kb from TSS (Figure 7B). Next, we focused the analysis on genes regulated by NKL54. For reference, we selected 2000 genes whose expression was not regulated by NKL54. Genes upregulated in response to NKL54 show an increase in H3K27ac around TSS (±3 kb) compared to those not regulated. In contrast, NKL54-downregulated genes show a reduction in H3K27ac levels (Figure 7C and Supplementary Figure S11A). When the analysis was performed on a larger region from TSS (±30 kb), the correlations between variations in H3K27ac levels and transcriptional response to NKL54 treatment were even more pronounced (Figure 7D and Supplementary Figure S11B). Analysis of absolute H3K27ac levels shows that genes that are upregulated in response to NKL54 are characterized by lower H3K27ac signals, whereas genes that are downregulated after NKL54 treatment, exhibit higher acetylation levels within a 3kb as well as 30kb around TSS (Figures 7E, F and Supplementary Figure S11C, D).

To confirm this observation, we compared the TPM of genes up and downregulated in response to NKL54. Genes that were not regulated by the inhibitors served as reference. Figure 7G demonstrates that genes upregulated by NKL54 are expressed at low levels, whereas downregulated genes are highly expressed genes. Similar behaviors can be observed for SAHA, for MC2984 after 24 h of treatment and in MC2985, but only after 6 h. TMP195 does not show these differential effects, which further corroborates its distance from the 'classical' HDACIs (Supplementary Figure S12).

In summary, NKL54 promotes the transcription of weakly expressed genes and antagonizes the expression of highly expressed genes, and these effects correlate with sculpting of the H3K27ac epigenome around the TSS.

NKL54 modifies the genomic distribution of MEF2D, HDAC4 and HDAC9

Having shown that NKL54 and other PAOA derivatives can to some extent affect the MEF2-dependent genetic program in LMS, we investigated the influence of NKL54 on the genomic activities of MEF2. We performed MEF2D ChIP-seq using chromatin isolated from SK-UT-1 cells treated with NKL54 for 14 h (Supplementary Figure S10). Under both conditions, a high percentage of MEF2D peaks co-localize with H3K27ac peaks in the presence of NKL54 (Supplementary Figure S13A and B).

To also evaluate changes in chromatin binding of class IIa HDACs, ChIP-seq experiments were performed for HDAC4 and HDAC9 under the same conditions (Supplementary Figure S10). We observed that HDAC4 and HDAC9 show similar behavior in terms of genome binding. NKL54 causes a reduction in the binding of HDAC4 and HDAC9, especially in the intergenic regions. Instead, their binding to promoters is conserved (Figure 8A and B). In contrast, NKL54 causes a dramatic increase in MEF2D binding to several genomic regions and especially to promoters (Figure 8A and B). These new NKL54-enhanced MEF2D-chromatin interactions are often located at new genomic regions (nearly 90% of the enriched peaks are new) (Figure 8C 'new'), whereas 40% of MEF2D peaks in untreated cells are conserved in NKL54-treated cells (Figure 8C 'conserved'). Importantly, in these new regions, MEF2D is frequently recruited in the absence of HDAC4 or of HDAC9 (Figure 8D). In contrast, in 33,4% of MEF2D peaks conserved between untreated and treated cells, HDAC4 is present and HDAC9 is found in approximately 20% of these regions. Therefore, NKL54 promotes the binding of MEF2D to numerous and novel genomic regions and within these regions it should act as a transcriptional activator as HDAC4 and HDAC9 are not recruited.

Interestingly, enriched MEF2D peaks are found in the promoter of 90 genes upregulated by NKL54. To further support this observation, we compared the effect of NKL54 on the genomic binding of MEF2D to the promoters of these 90 genes with respect to the 2000 genes that show no variation in their expression (Figure 8E and Supplementary Figure S13C). MEF2D binding is strongly increased by NKL54 at the promoter regions of the 90 upregulated genes, whereas the effect is much smaller/absent at the promoters of the not-regulated genes. Two different loci (*ERRFI1* and *IER3*) encoding genes regulated by MEF2D and by NKL54 are a good example of these changes. In *ERRFI1*, NKL54 triggers MEF2D to bind to a large genomic region in the absence of HDAC4 and of HDAC9. Instead, at the *IER3* locus, NKL54 promotes binding of MEF2D as well as of HDAC4 and, to a smaller extent, of HDAC9 (Figure 8F and Supplementary Figure S14A). This observation further suggests that NKL54 cannot affect binding between MEF2 and class IIa HDACs *in vivo*. Interestingly, in another region near the locus, binding of both HDAC4 and HDAC9 can be detected, and this binding is independent from MEF2D and NKL54 (Figure 8F and Supplementary Figure S14A, highlighted in green). Two other examples of loci whose expression is upregulated by NKL54 treatment and characterized by the appearance of MEF2D binding after NKL54 treatment are *CXCL1* and *CXCL2* (Supplementary Figure S14B).

DISCUSSION

The availability of small compounds that induce chromatin remodelling in neoplastic cells is a promising anticancer strategy. The development of small molecules that alter protein-protein interactions is a challenging but also a new and growing area of drug discovery. Here, we have investigated and characterized the possibility of disrupting the interaction between MEF2 and class IIa HDACs. The original idea was to target the surface of the interaction between MEF2 and these epigenetic repressors in LMS. However, because the MEF2-HDAC axis is also perturbed in other cancers, our study may have much broader implications (1,17).

The prototype molecule was the PAOA derivative NKL54 (9). Molecular modelling confirmed that NKL54 should be able to fit into the hydrophobic groove of

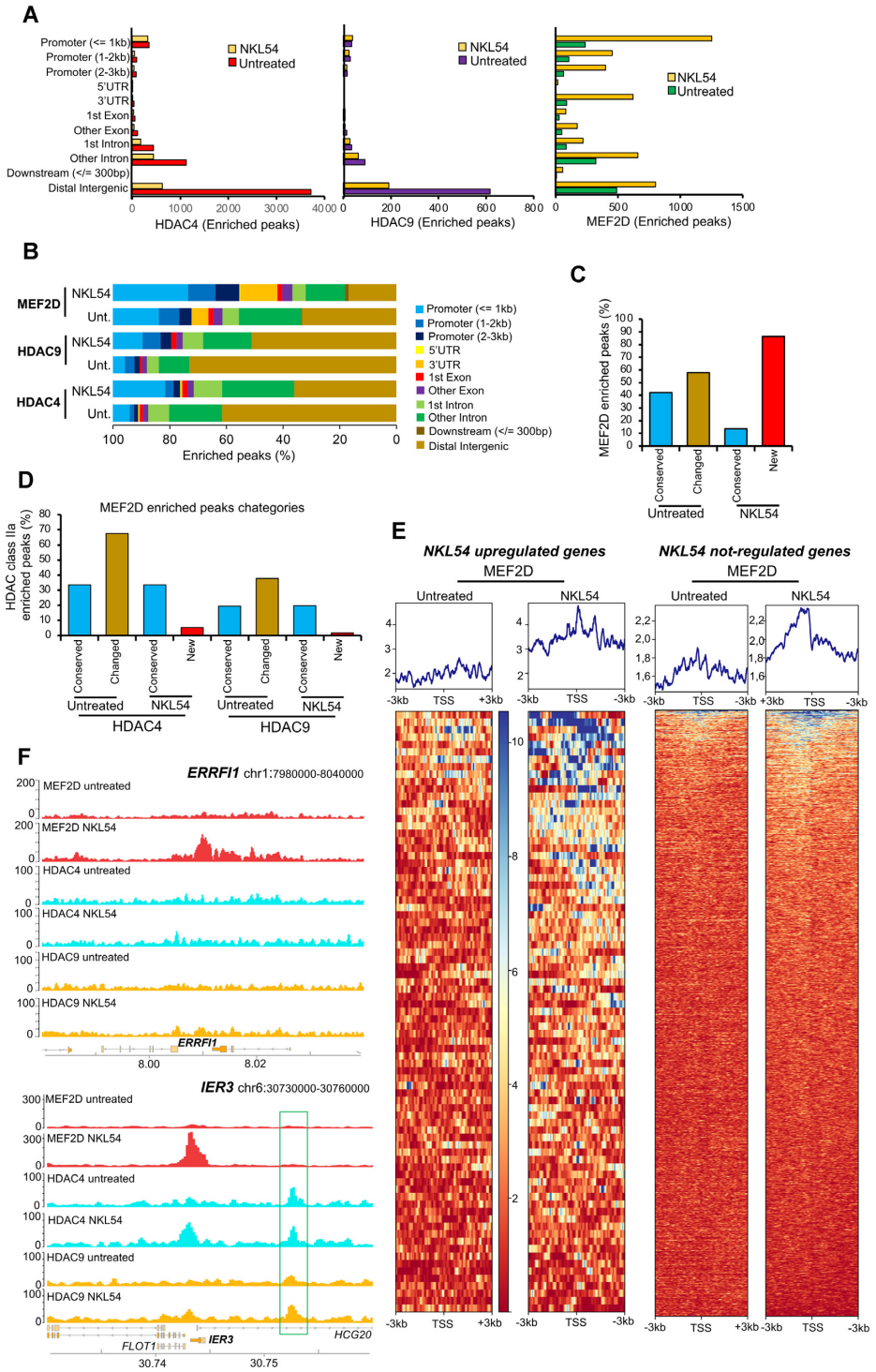


Figure 8. NKL54 exerts a profound influence on the genomic binding of MEF2D, HDAC4 and HDAC9. (**A**) Genomic distribution of HDAC4, HDAC9 and MEF2D-enriched IDR-defined peaks identified by MACS2 in SK-UT-1 cells treated (n = 1770, 393 and 4723) or untreated ($n = 6109$, 891 and 1496) for 14 h with 5 μM NKL54. (**B**) As in panel A, with values represented as percentage. (**C**) Proximity, expressed as percentage of the overall ChIP-seq enriched peaks, between the MEF2D enriched peaks identified in SK-UT-1 cells treated with NKL54 and those found in untreated cells. We defined as 'Conserved' the overlapping peaks between cells treated and untreated with NKL54, whereas 'Changed' are peaks found only in untreated cells and 'New' are peaks found only in NKL54-treated cells. (**D**) Proximity to MEF2D peaks, expressed as percentage, of HDAC4 and HDAC9 ChIP-seq enriched peaks in SK-UT-1 cells treated or untreated with NKL54. MEF2D peaks are defined by the categories showed in panel C. The maximum distance to define overlapping peaks (Conserved) is 1 kb. (**E**) Heat-maps of the MEF2D signal distribution in (left) a region of ±3 kb around the TSS of 90 genes upregulated by NKL54 treatment and showing the appearance of NKL54 *de novo* MEF2D peaks, and (right) around the TSS of 2000 genes not regulated by NKL54 treatment, as indicated. MEF2D signals are compared between untreated and NKL54 treated cells. ChIP-seq data are from experiment 1. (**F**) Detailed view of the MEF2D, HDAC4 and HDAC9 tracks at two representative loci (*ERRFI1* and *IER3*), upregulated by NKL54 and showing *de novo* MEF2D peaks. Gene structure and chromosomal location are shown.

MEF2. Unfortunately, our *in vitro* and *in vivo* studies show that NKL54 could not compete with the binding between HDAC4 and MEF2A/D, although the binding, at least *in vitro*, is dynamic. Several hypotheses can be formulated. Structural deficiencies in mimicking α-helix distribution, limited contact sites within the hydrophobic groove, or failure to release the hot-spots of protein-protein interactions may explain the inability of NKL54 to act as an orthosteric inhibitor of MEF2-HDACs interactions (49,50). MC2984 and MC2985, predicted by molecular modelling to interact with MEF2 similarly to NKL54, are also unable to compete with HDAC4 peptide binding. Importantly, only compounds that inhibit KDACs and increase histone acetylation can induce LMS cell death.

The persistence of HDAC inhibitory activity in these PAOA derivatives was confirmed by comparative transcriptomic analysis. Although these compounds, unlike SAHA, do not inhibit HDAC6 and HDAC8, the DEGs are largely overlapping with those found for SAHA. Thus, in SK-UT-1 cells, inhibition of HDAC1/HDAC2/HDAC3 causes most transcriptional adjustments and is sufficient to trigger cell death. A group of pro-apoptotic BCL2 members belonging to the BH3-only subfamily are upregulated by SAHA and PAOA derivatives, providing a link between HDAC1/2/3 inhibition and their upregulation. The expression of BIM, BMF and HRK is strongly upregulated as an early response and high levels are maintained throughout. Similarly, BCL2L11/BIM and BBC3/PUMA are upregulated, although less strongly. All these BH3-only members are upregulated by different HDACIs in different cancer models (51–55).

In general, upregulated genes are expressed at low levels in untreated cells, whereas downregulated genes are abundantly expressed. Moreover, the downregulated genes characterize the late response to HDACIs. This repressive wave may represent an adaptation to the unscheduled transcriptional reprogramming. The downregulation of several KATs can also be seen in this context (46,56,57).

The repressive effect of HDACIs on highly transcribed genes may be due to different mechanisms. HDACs may limit acetylation in the gene body and intergenic regions (58,59). This action optimizes recruitment of BRD4, a key elongation factor at promoters and enhancers (59). HDACIs can block elongation of RNA polymerase II and increase pausing of RNAPII at enhancers and super-enhancers (60). At super-enhancers, HDACIs can also cause excessive H3K27ac spreading, an effect that alters normal chromosomal looping (61). Erosion of super-enhancers boundaries, because of H3K27ac spreading, may also be responsible for downregulation of highly expressed genes.

Among PAOA derivatives, NKL54 and MC2984 show few differences. In general, NKL54 is more potent and modulates more genes, especially those that are downregulated. The mechanism through which, the trifluoro group can cause such differences deserves further investigation. Curiously, but expected (39,62), high concentrations of NKL54 can inhibit HDAC4 *in vitro*.

Neural differentiation represents the most enriched DEGs category in response to SAHA and PAOA. Interestingly, gene programs related to neural differentiation have also been activated in other cell lines: synovial sarcoma cells, human embryonic stem cells, and malignant rhabdoid tumor cells, in response to structurally unrelated HDACIs (54,63). They may represent a genetic program silenced by HDACs in non-neuronal cells and reactivated in the presence of the inhibitors.

Only few genes are modulated by MC2985. This compound is a very weak HDAC inhibitor but has strong pro-death activity. Therefore, it is plausible that MC2985 has additional targets, possibly through the action of its 2(alkylthio)-4-phenyl-pirimidine group.

TMP195 is a class IIa specific inhibitor. In LMS cells, it shows a weak anti-proliferative effect. In contrast, deletion of HDAC4 and of HDAC9 strongly affects cell survival and proliferation (19,36). The role of these epigenetic regulators, as scaffolds for the assembly of multiprotein complexes, may explain this discrepancy (64,65). Catalytic domain targeting may not be sufficient to knock down all class IIa activities. Approximately 50% of the genes modulated by TMP195 are also modulated by class I HDAC inhibitors. This overlap is not surprising because class IIa enzymes coordinate the activity of the NCOR1-NCORII-HDAC3 complex via the deacetylase domain (66).

We have shown that HDACIs and NKL54 particularly, can affect MEF2 transcriptional activity. First, HDAC7 (at earlier times) and, HDAC4 and HDAC9 (later) are downregulated. Second, MEF2C and MEF2D are upregulated. These changes could contribute to convert MEF2 complexes dedicated to repression into transcriptional activators. Indeed, approximately 30% of genes under MEF2 regulation are also upregulated by NKL54. Consistent with our observations, BML-210 can promote the activation of MEF2-dependent memory-related genes and the increase of synaptic markers in the hippocampus of a mouse model of Huntington's disease (67).

ChIP-seq experiments have revealed a global increase in MEF2D genome occupancy in response to NKL54. Increased recruitment of TFs to regulatory regions in response to HDACIs has been reported for PU.I (68). In the case of MEF2D, further studies will be necessary to clarify the effect of NKL54 on MEF2D genome occupancy. The creation of new and more accessible chromatin regions could be evoked (69), but a direct effect on MEF2D acetylation status and potentiation of its DNA-binding activity cannot be excluded. Indeed, it has been reported that HDAC3 can bind and deacetylate MEF2D (70–72).

In conclusion, upregulation of the MEF2 transcriptional program may be beneficial for LMS patients, as evidenced by the better prognosis when the MEF2-NKL54 signature is expressed at higher levels. Targeting the interaction between MEF2 and class IIa HDACs is still an open challenge. Our results further stimulate the search for new compounds capable of reactivating MEF2-dependent transcription.

DATA AVAILABILITY

The transcriptomic raw data are available as GEO accession GSE180804: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE180804.

Raw data corresponding to ChIP-seq experiments are uploaded with GEO accession GSE180681: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE180681.

The link to a UCSC genome browser session displaying the uploaded sequence tracks is https://genome.ucsc.edu/s/DameBioinfo/NAR_2021_rep.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Fabio Benedetti (University of Trieste, Italy) and Danilo Licastro (ARGO Open Lab Platform for Genome Sequencing, AREA Science Park, Trieste, Italy) for the helpful discussion. E.D.G. thanks AIRC for the recent financial support MFAG 2020 ID25000.

FUNDING

PRIN [2017JL8SRX 'Class IIa HDACs as therapeutic targets in human diseases: new roles and new selective inhibitors' to C.B., A.A., L.C., R.G.]; Interreg Italia-Osterreich [ITAT1054 EPIC to C.B., E.K., C.X.W.]. Funding for open access charge: PRIN 2017JL8SRX and ITAT1054 EPIC.

Conflict of interest statement. None declared.

REFERENCES

1. Di Giorgio,E., Hancock,W.W. and Brancolini,C. (2018) MEF2 and the tumorigenic process, hic sunt leones. *BBA - Rev. Cancer.*, **1870**, 261–273.
2. He,T., Huang,J., Chen,L., Han,G., Stanmore,D., Krebs-Haupenthal,J., Avkiran,M., Hagenmüller,M. and Backs,J. (2020) Cyclic AMP represses pathological MEF2 activation by myocyte-specific hypo-phosphorylation of HDAC5. *J. Mol. Cell Cardiol.*, **145**, 88–98.
3. Tobin,S.W., Hashemi,S., Dadson,K., Turdi,S., Ebrahimian,K., Zhao,J., Sweeney,G., Grigull,J. and McDermott,J.C. (2017) Heart failure and MEF2 transcriptome dynamics in response to β-Blockers. *Sci Rep.*, **7**, 4476.
4. Assali,A., Harrington,A.J. and Cowan,C.W. (2019) Emerging roles for MEF2 in brain development and mental disorders. *Curr. Opin. Neurobiol.*, **59**, 49–58.
5. Pon,J.R. and Marra,M.A. (2016) MEF2 transcription factors: developmental regulators and emerging cancer genes. *Oncotarget*, **7**, 2297–2312.
6. Han,A., Pan,F., Stroud,J.C., Youn,H.D., Liu,J.O. and Chen,L. (2003) Sequence-specific recruitment of transcriptional co-repressor cabin1 by myocyte enhancer factor-2. *Nature*, **422**, 730–734.
7. He,J., Ye,J., Cai,Y., Riquelme,C., Liu,J.O., Liu,X., Han,A. and Chen,L. (2011) Structure of p300 bound to MEF2 on DNA reveals a mechanism of enhanceosome assembly. *Nucleic Acids Res.*, **39**, 4464–4474.
8. Han,A., He,J., Wu,Y., Liu,J.O. and Chen,L. (2005) Mechanism of recruitment of class II histone deacetylases by myocyte enhancer factor-2. *J. Mol. Biol.*, **345**, 91–102.
9. Jayathilaka,N., Han,A., Gaffney,K.J., Dey,R., Jarusiewicz,J.A., Noridomi,K., Philips,M.A., Lei,X., He,J., Ye,J. *et al.* (2012) Inhibition of the function of class IIa HDACs by blocking their interaction with MEF2. *Nucleic Acids Res.*, **40**, 5378–5388.
10. Savickiene,J., Borutinskaite,V.V., Treigyte,G., Magnusson,K.E. and Navakauskiene,R. (2006) The novel histone deacetylase inhibitor BML-210 exerts growth inhibitory, proapoptotic and differentiation stimulating effects on the human leukemia cell lines. *Eur. J. Pharmacol.*, **549**, 9–18.
11. Wei,J., Joshi,S., Speransky,S., Crowley,C., Jayathilaka,N., Lei,X., Wu,Y., Gai,D., Jain,S., Hoosien,M. *et al.* (2017) Reversal of pathological cardiac hypertrophy via the MEF2-coregulator interface. *JCI Insight.*, **2**, e91068.
12. Chibon,F., Darbo,E. and Pérot,G. (2019) Leiomyosarcomas: whole genome sequencing for a whole biology characterization. *Curr. Opin. Oncol.*, **31**, 317–321.
13. Montella,L., Altucci,L., Sarno,F., Buonerba,C., De Simone,S., Facchini,B.A., Franzese,E., De Vita,F., Tafuto,S., Berretta,M. *et al.* (2021) Toward a personalized therapy in soft-tissue sarcomas: state of the art and future directions. *Cancers (Basel)*, **13**, 2359.
14. Iuliano,L., Drioli,S., Pignochino,Y., Cafiero,C.M., Minisini,M., D'Este,F., Picco,R., Dalla,E., Giordano,G., Grignani,G. *et al.* (2021) Enhancing proteotoxic stress in leiomyosarcoma cells triggers mitochondrial dysfunctions, cell death, and antitumor activity in vivo. *Mol. Cancer Ther.*, **20**, 1039–1051.
15. Di Giorgio,E., Gagliostro,E., Clocchiatti,A. and Brancolini,C. (2015) The control operated by the cell cycle machinery on MEF2 stability contributes to the downregulation of CDKN1A and entry into S phase. *Mol. Cell. Biol.*, **35**, 1633–1647.
16. Paroni,G., Mizzau,M., Henderson,C., Del Sal,G., Schneider,C. and Brancolini,C. (2004) Caspase-dependent regulation of histone deacetylase 4 nuclear-cytoplasmic shuttling promotes apoptosis. *Mol. Biol. Cell*, **15**, 2804–2818.
17. Clocchiatti,A., Di Giorgio,E., Ingrao,S., Meyer-Almes,F.J., Tripodo,C. and Brancolini,C. (2013) Class IIa HDACs repressive activities on MEF2-depedent transcription are associated with poor prognosis of ER^+ breast tumors. *FASEB J.*, **27**, 942–954.
18. Cutano,V., Di Giorgio,E., Minisini,M., Picco,R., Dalla,E. and Brancolini,C. (2019) HDAC7-mediated control of tumor microenvironment maintains proliferative and stemness competence of human mammary epithelial cells. *Mol Oncol.*, **13**, 1651–1668.
19. Di Giorgio,E., Dalla,E., Franforte,E., Paluvai,H., Minisini,M., Trevisanut,M., Picco,R. and Brancolini,C. (2020) Different class IIa HDACs repressive complexes regulate specific epigenetic responses related to cell survival in leiomyosarcoma cells. *Nucleic Acids Res.*, **48**, 646–664.
20. Cersosimo,U., Sgorbissa,A., Foti,C., Drioli,S., Angelica,R., Tomasella,A., Picco,R., Semrau,M.S., Storici,P., Benedetti,F. *et al.* (2015) Synthesis, characterization, and optimization for in vivo delivery of a nonselective isopeptidase inhibitor as new antineoplastic agent. *J. Med. Chem.*, **58**, 1691–1704.
21. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
22. Maier,J.A., Martinez,C., Kasavajhala,K., Wickstrom,L., Hauser,K.E. and Simmerling,C. (2013) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.*, **11**, 3696–3713.
23. Quiroga,R. and Villarreal,M.A. (2016) Vinardo: a scoring function based on autodock vina improves scoring, docking, and virtual screening. *PLoS One*, **11**, e0155183.
24 Korb,O., Stutzle,T. and Exner,T.E. (2006) Plants: application of ant colony optimization to structure-based drug design. Ant colony optimization and swarm intelligence, *Proc. AMIA Annu. Fall Symp.*, **4150**, 247–258.
25. Ragno,R. (2019) www.3d-qsar.com: a web portal that brings 3-D QSAR to all electronic devices-the Py-CoMFA web application as tool to build models from pre-aligned datasets. *J. Comput. Aided Mol. Des.*, **33**, 855–864.
26. Ragno,R., Frasca,S., Manetti,F., Brizzi,A. and Massa,S. (2005) HIV-reverse transcriptase inhibition: inclusion of ligand-induced fit by cross-docking studies. *J. Med. Chem.*, **48**, 200–212.
27. Musmuca,I., Caroli,A., Mai,A., Kaushik-Basu,N., Arora,P. and Ragno,R. (2010) Combining 3-D quantitative structure-activity relationship with ligand based and structure-based alignment procedures for in silico screening of new hepatitis c virus NS5B polymerase inhibitors. *J. Chem. Inf. Model.*, **50**, 662–676.
28. Jänsch,N., Meyners,C., Muth,M., Kopranovic,A., Witt,O., Oehme,I. and Meyer-Almes,F.J. (2019) The enzyme activity of histone deacetylase 8 is modulated by a redox-switch. *Redox. Biol.*, **20**, 60–67.
29. Ewels,P., Magnusson,M., Lundin,S. and Käller,M. (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, **32**, 3047–3048.

30. Patro,R., Duggal,G., Love,M.I., Irizarry,R.A. and Kingsford,C. (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods*, **14**, 417–419.
31. Love,M.I., Soneson,C., Hickey,P.F., Johnson,L.K., Pierce,N.T., Shepherd,L., Morgan,M. and Patro,R. (2020) Tximeta: reference sequence checksums for provenance identification in RNA-seq. *PLoS Comput. Biol.*, **16**, e1007664.
32. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
33. Yu,G., Wang,L.G., Han,Y. and He,Q.Y. (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*, **16**, 284–287.
34. Di Giorgio,E., Paluvai,H., Dalla,E., Ranzino,L., Renzini,A., Moresi,V., Minisini,M., Picco,R. and Brancolini,C. (2021) HDAC4 degradation during senescence unleashes an epigenetic program driven by AP-1/p300 at selected enhancers and super-enhancers. *Genome Biol.*, **22**, 129.
35. Li,Q., Brown,J.B., Huang,H. and Bickel,P.J. (2011) Measuring reproducibility of high-throughput experiments. *Ann. App. Stat.*, **5**, 1752–1779.
36. Di Giorgio,E., Franforte,E., Cefalù,S., Rossi,S., Dei Tos,A.P., Brenca,M., Polano,M., Maestro,R., Paluvai,H., Picco,R. *et al.* (2017) The co-existence of transcriptional activator and transcriptional repressor MEF2 complexes influences tumor aggressiveness. *PLoS Genet.*, **13**, e1006752.
37. Chou,C.J., Herman,D. and Gottesfeld,J.M. (2008) Pimelic diphenylamide 106 is a slow, tight-binding inhibitor of class I histone deacetylases. *J. Biol. Chem.*, **283**, 35402–35409.
38. Lobera,M., Madauss,K.P., Pohlhaus,D.T., Wright,Q.G., Trocha,M., Schmidt,D.R., Baloglu,E., Trump,R.P., Head,M.S., Hofmann,G.A. *et al.* (2013) Selective class IIa histone deacetylase inhibition via a nonchelating zinc-binding group. *Nat. Chem. Biol.*, **9**, 319–325.
39. Di Giorgio,E., Gagliostro,E. and Brancolini,C. (2015) Selective class IIa HDAC inhibitors: myth or reality. *Cell. Mol. Life Sci.*, **72**, 73–86.
40. Thome,M., Schneider,P., Hofmann,K., Fickenscher,H., Meinl,E., Neipel,F., Mattmann,C., Burns,K., Bodmer,J.L., Schröter,M. *et al.* (1997) Viral FLICE-inhibitory proteins (FLIPs) prevent apoptosis induced by death receptors. *Nature*, **386**, 517–521.
41. Korb,O., Stuetzle,T. and Exner,T.E. (2009) Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J. Chem. Inf. Model.*, **49**, 84–96.
42. Bottomley,M.J., Lo Surdo,P., Di Giovine,P., Cirillo,A., Scarpelli,R., Ferrigno,F., Jones,P., Neddermann,P., De Francesco,R., Steinkühler,C. *et al.* (2008) Structural and functional analysis of the human HDAC4 catalytic domain reveals a regulatory structural zinc-binding domain. *J. Biol. Chem.*, **283**, 26694–26704.
43. Witt,A.E., Lee,C.W., Lee,T.I., Azzam,D.J., Wang,B., Caslini,C., Petrocca,F., Grosso,J., Jones,M., Cohick,E.B. *et al.* (2017) Identification of a cancer stem cell-specific function for the histone deacetylases, HDAC1 and HDAC7, in breast and ovarian cancer. *Oncogene*, **36**, 1707–1720.
44. Yu,G. and He,Q.Y. (2016) ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst.*, **12**, 477–479.
45. Warren,C.F.A., Wong-Brown,M.W. and Bowden,N.A. (2019) BCL-2 family isoforms in apoptosis and cancer. *Cell Death. Dis.*, **10**, 177.
46. Lamparter,C.L. and Winn,L.M. (2016) Valproic acid exposure decreases Cbp/p300 protein expression and histone acetyltransferase activity in P19 cells. *Toxicol. Appl. Pharmacol.*, **306**, 69–78.
47. Brown,J.A., Bourke,E., Eriksson,L.A. and Kerin,M.J. (2016) Targeting cancer using KAT inhibitors to mimic lethal knockouts. *Biochem. Soc. Trans.*, **44**, 979–986.
48. McCann,A.J., Lou,J., Moustaqil,M., Graus,M.S., Blum,A., Fontaine,F., Liu,H., Luu,W., Rudolffi-Soto,P., Koopman,P. *et al.* (2021) A dominant-negative SOX18 mutant disrupts multiple regulatory layers essential to transcription factor activity. *Nucleic Acids Res.*, **49**, 10931–10955.
49. Algar,S., Martín-Martínez,M. and González-Muñiz,R. (2021) Evolution in non-peptide α-helix mimetics on the road to effective protein-protein interaction modulators. *Eur. J. Med. Chem.*, **211**, 113015.
50. Fischer,G., Rossmann,M. and Hyvönen,M. (2015) Alternative modulation of protein-protein interactions by small molecules. *Curr. Opin. Biotechnol.*, **35**, 78–85.
51. Feng,L., Pan,M., Sun,J., Lu,H., Shen,Q., Zhang,S., Jiang,T., Liu,L., Jin,W., Chen,Y. *et al.* (2013). Histone deacetylase 3 inhibits expression of PUMA in gastric cancer cells. *J. Mol. Med. (Berl.)*, **91**, 49–58.
52. Bolden,J.E., Shi,W., Jankowski,K., Kan,C.Y., Cluse,L., Martin,B.P., MacKenzie,K.L., Smyth,G.K. and Johnstone,R.W. (2013) HDAC inhibitors induce tumor-cell-selective pro-apoptotic transcriptional responses. *Cell Death. Dis.*, **4**, e519.
53. Laporte,A.N., Poulin,N.M., Barrott,J.J., Wang,X.Q., Lorzadeh,A., Vander Werff,R., Jones,K.B., Underhill,T.M. and Nielsen,T.O. (2017) Death by HDAC inhibition in synovial sarcoma cells. *Mol. Cancer Ther.*, **16**, 2656–2667.
54. Dong,Z., Yang,Y., Liu,S., Lu,J., Huang,B. and Zhang,Y. (2017) HDAC inhibitor PAC-320 induces G2/M cell cycle arrest and apoptosis in human prostate cancer. *Oncotarget*, **9**, 512–523.
55. Laszig,S., Boedicker,C., Weiser,T., Knapp,S. and Fulda,S. (2020) The novel dual BET/HDAC inhibitor TW09 mediates cell death by mitochondrial apoptosis in rhabdomyosarcoma cells. *Cancer Lett.*, **486**, 46–57.
56. Rafehi,H., Balcerczyk,A., Lunke,S., Kaspi,A., Ziemann,M., Kn,H., Okabe,J., Khurana,I., Ooi,J., Khan,A.W. *et al.* (2014) Vascular histone deacetylation by pharmacological HDAC inhibition. *Genome Res.*, **24**, 1271–1284.
57. Halsall,J.A., Turan,N., Wiersma,M. and Turner,B.M. (2015) Cells adapt to the epigenomic disruption caused by histone deacetylase inhibitors through a coordinated, chromatin-mediated transcriptional response. *Epigenetics Chromatin.*, **8**, 29.
58. Kim,Y.J., Greer,C.B., Cecchini,K.R., Harris,L.N., Tuck,D.P. and Kim,T.H. (2013) HDAC inhibitors induce transcriptional repression of high copy number genes in breast cancer through elongation blockade. *Oncogene*, **32**, 2828–2835.
59. Greer,C.B., Tanaka,Y., Kim,Y.J., Xie,P., Zhang,M.Q., Park,I.H. and Kim,T.H. (2015) Histone deacetylases positively regulate transcription through the elongation machinery. *Cell Rep.*, **13**, 1444–1455.
60. Sanchez,G.J., Richmond,P.A., Bunker,E.N., Karman,S.S., Azofeifa,J., Garnett,A.T., Xu,Q., Wheeler,G.E., Toomey,C.M., Zhang,Q. *et al.* (2018) Genome-wide dose-dependent inhibition of histone deacetylases studies reveal their roles in enhancer remodeling and suppression of oncogenic super-enhancers. *Nucleic Acids Res.*, **46**, 1756–1776.
61. Gryder,B.E., Pomella,S., Sayers,C., Wu,X.S., Song,Y., Chiarella,A.M., Bagchi,S., Chou,H.C., Sinniah,R.S., Walton,A. *et al.* (2019) Histone hyperacetylation disrupts core gene regulatory architecture in rhabdomyosarcoma. *Nat. Genet.*, **51**, 1714–1722.
62. Lahm,A., Paolini,C., Pallaoro,M., Nardi,M.C., Jones,P., Neddermann,P., Sambucini,S., Bottomley,M.J., Lo Surdo,P., Carfí,A. *et al.* (2007) Unraveling the hidden catalytic activity of vertebrate class IIa histone deacetylases. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 17335–17340.
63. Muscat,A., Popovski,D., Jayasekara,W.S., Rossello,F.J., Ferguson,M., Marini,K.D., Alamgeer,M., Algar,E.M., Downie,P., Watkins,D.N. *et al.* (2016) Low-Dose histone deacetylase inhibitor treatment leads to tumor growth arrest and multi-lineage differentiation of malignant rhabdoid tumors. *Clin. Cancer Res.*, **22**, 3560–3570.
64. Brancolini,C., Di Giorgio,E., Formisano,L. and Gagliano,T. (2021) Quis custodiet ipsos custodes (Who controls the controllers)? Two decades of studies on HDAC9. *Life (Basel)*, **11**, 90.
65. Di Giorgio,E. and Brancolini,C. (2016) Regulation of class IIa HDAC activities: it is not only matter of subcellular localization. *Epigenomics*, **8**, 251–269.
66. Park,S.Y., Kim,G.S., Hwang,H.J., Nam,T.H., Park,H.S., Song,J., Jang,T.H., Lee,Y.C. and Kim,J.S. (2018) Structural basis of the specific interaction of SMRT corepressor with histone deacetylase 4. *Nucleic Acids Res.*, **46**, 11776–11788.
67. Vidal-Sancho,L., Fernández-García,S., Solés-Tarrés,I., Alberch,J. and Xifró,X. (2020) Decreased myocyte enhancer factor 2 levels in the hippocampus of huntington's disease mice are related to cognitive dysfunction. *Mol. Neurobiol.*, **57**, 4549–4562.
68. Frank,C.L., Manandhar,D., Gordân,R. and Crawford,G.E. (2016) HDAC inhibitors cause site-specific chromatin remodeling at PU.1-bound enhancers in K562 cells. *Epigenetics Chromatin.*, **9**, 15.

69. Qu,K., Zaba,L.C., Satpathy,A.T., Giresi,P.G., Li,R., Jin,Y., Armstrong,R., Jin,C., Schmitt,N., Rahbar,Z. *et al.* (2017) Chromatin accessibility landscape of cutaneous T cell lymphoma and dynamic response to HDAC inhibitors. *Cancer Cell*, **32**, 27–41.
70. Angelelli,C., Magli,A., Ferrari,D., Ganassi,M., Matafora,V., Parise,F., Razzini,G., Bachi,A., Ferrari,S. and Molinari,S. (2008) Differentiation-dependent lysine 4 acetylation enhances MEF2C binding to DNA in skeletal muscle cells. *Nucleic Acids Res.*, **36**, 915–928.
71. Ma,K., Chan,J.K., Zhu,G. and Wu,Z. (2005) Myocyte enhancer factor 2 acetylation by p300 enhances its DNA binding activity, transcriptional activity, and myogenic differentiation. *Mol. Cell. Biol.*, **25**, 3575–3582.
72. Grégoire,S., Xiao,L., Nie,J., Zhang,X., Xu,M., Li,J., Wong,J., Seto,E. and Yang,X.J. (2007) Histone deacetylase 3 interacts with and deacetylates myocyte enhancer factor 2. *Mol. Cell. Biol.*, **27**, 1280–1295.
73. Morgan,H.L. (1965) The generation of a unique machine description for chemical Structures-A technique developed at chemical abstracts service. *J. Chem. Doc.*, **5**, 107–113.
74. Perkel,J.M. (2015) Programming: pick up python. *Nature*, **518**, 125–126.
75. Landrum,G. (2019) RDKit: open-source cheminformatics from machine learning to chemical registration. In: *Abstracts of Papers of the American Chemical Society*. American Chemical Society, Washington, DC, Vol. **258**.