Sede Amministrativa: Università degli Studi di Padova
Dipartimento di *Salute della Donna e del Bambino*
Clinica di Oncoematologia Pediatrica

SCUOLA DI DOTTORATO DI RICERCA IN MEDICINA DELLO SVILUPPO E SCIENZE DELLA PROGRAMMAZIONE SANITARIA
INDIRIZZO: EMATOONCOLOGIA, GENETICA, MALATTIE RARE E MEDICINA PREDITTIVA
CICLO XXVII

DNA METHYLATION ANALYSIS IN RHABDOMYOSARCOMA

Direttore della Scuola: Ch.mo Prof. GIUSEPPE BASSO
Coordinatore d'indirizzo: Ch.mo Prof. GIUSEPPE BASSO
Supervisore: Dott.ssa ANGELICA ZIN

Dottorando: ELENA POLI

INDEX

SUMMARY

Rhabdomyosarcoma (RMS) is a highly aggressive pediatric soft-tissue sarcoma. It is mainly classified into two major subtypes characterized by alveolar (ARMS) and embryonal (ERMS) histologies. ARMS are characterized by a more aggressive behavior with a higher tendency to present metastasis at diagnosis and to relapse after treatment. Approximately 80% of ARMS harbour the reciprocal chromosomal translocation t(2;13)(q35;q14) and, less commonly, the variant translocation t(1;13)(p36;q14), in which *PAX3* and *FOXO1*, or *PAX7* and *FOXO1* genes, respectively, are juxtaposed. Unfortunately, no such specific genetic aberrations are known in ERMS, and myogenic factors as myogenin and MyoD1 are the only diagnostic indicators that can be used. Despite aggressive multimodal therapies, the prognosis of high-risk RMS patients has not been improved, with a 5-year overall survival rate less than 20-30%, which prompts a need for new therapeutic strategies. In the last decade many scientific studies have demonstrated that gene expression signature distinguishes *PAX3-FOXO1* positive RMS from *PAX3-FOXO1* negative, but the reasons of the different expression are still unknown. Aberrant DNA methylation patterning is a hallmark of cancer and could be responsible for the different gene expression of RMS tumor subtypes.

We performed genome-wide methylation profile by microarray experiments followed by *Reduced-Representation Bisulfite Sequencing* (RRBS).

Microarray analysis demonstrated a different methylation profile between *PAX3-FOXO1* positive and negative RMS, besides among metastatic and non-metastatic RMS. We confirmed *HOXC11* as one of the gene differentially methylated between *PAX3-FOXO1* positive and negative RMS cell lines using *in vitro* demethylating agents and bisulfite sequencing. Unfortunately, we did not validate the result in the cohort of RMS biopsies. Moreover, we performed another analysis on microarray data comparing metastatic *vs* non-metastatic RMS. We found an elevated numbers of *differentially methylated regions* (DMRs) and many of them map to promoter regions of genes implicated in tumors development. In particular we found DMRs linked to clustered protocadherins, known as tumor suppressor genes. We confirmed a different expression pattern of *PCDHA4*, as well as a different methylation level of its promotorial region, comparing metastatic and non-metastatic RMS samples. Nevertheless, the methylation status and the expression level of *PCDHA4* did not have significant correlation with clinical features and are not a predictor of poor prognosis in RMS.

Then, we performed an RRBS sequencing, in order to validate data obtained with microarray platforms. We observed a very low concordance between the two approaches, probably caused by a low quality DNA used in microarray experiments. The RRBS sequencing had demonstrated again that *PAX3-FOXO1* positive and negative RMS have a different methylation pattern. Moreover, we demonstrated that *GADD45G* and *NELL1*, already described as tumor suppressors in other cancers and often downregulated by methylation processes, had also an involvement in RMS biology. Our experiments confirmed an epigenetic regulation by DNA methylation for *GADD45G* and *NELL1* and that their expression were correlated to RMS histology, presence of fusion status and IRS group staging. Furthermore, *GADD45G* and *NELL1* expression levels affect the progression free survival of RMS patients suggesting their association with a poor prognosis.

In conclusion, we demonstrated that *GADD45G* and *NELL1* could be novel potential biomarkers in RMS and we evidenced that the DNA methylation pattern in RMS could be interesting for new therapeutic strategies.

We hope that our efforts could contribute to a better molecular classification of RMS tumors and to the identification of new targets for improving standard therapy.

RIASSUNTO

Il rabdomiosarcoma (RMS) è una sarcoma pediatrico dei tessuti molli altamente aggressivo. Viene classificato principalmente in due sottotipi, caratterizzati da istologia alveolare (RMSA) o embrionale (RMSE). Nei RMSA si osserva un comportamento più aggressivo e una maggiore tendenza a presentare metastasi alla diagnosi e alla ricaduta dopo trattamento. Circa l'80% dei RMSA presentano la traslocazione cromosomica reciproca t(2; 13) (q35; q14) e, meno comunemente, la variante t(1; 13) (p36; q14), in cui i geni *PAX3* e *FOXO1*, o *PAX7* e *FOXO1*, rispettivamente, sono giustapposti. Purtroppo, non si conoscono aberrazioni genetiche specifiche nei RMSE e i fattori miogenici, come miogenina e MyoD1, sono gli unici indicatori diagnostici che possono essere utilizzati. Nonostante l'applicazione di terapie aggressive multimodali, la prognosi dei pazienti affetti da RMS, della categoria alto rischio, non è migliorata, con un tasso di sopravvivenza a 5 anni inferiore al 20-30%. Questo dato indica la necessità di sviluppare nuove strategie terapeutiche. Nell'ultimo decennio molti studi scientifici hanno dimostrato che in base al profilo di espressione genica è possibile distinguere RMS *PAX3-FOXO1*-positivi e *PAX3-FOXO1*-negativi, ma le ragioni di questa diversa espressione sono ancora sconosciute. L'anomala metilazione del DNA è un indicatore di neoplasia e potrebbe essere la causa responsabile della diversa espressione genica dei due sottotipi di tumore.

In questo studio, per mezzo di esperimenti di microarray, abbiamo realizzato un'analisi dello stato di metilazione del DNA su tutto il genoma, proseguendo poi con esperimenti di sequenziamento sfruttando la tecnica *Reduced-Representation Bisulfite Sequencing* (RRBS).

L'analisi dei risultati ottenuti con gli esperimenti di microarray ha dimostrato, non solo un profilo di metilazione diverso tra i RMS *PAX3-FOXO1*-positivi e negativi, ma anche tra i RMS metastatici e non metastatici. Abbiamo confermato che il gene *HOXC11* risulta essere differenzialmente metilato tra linee cellulari di RMS *PAX3-FOXO1*-positive e negative, sfruttando trattamenti con agenti demetilanti *in vitro* e sequenziamento del DNA dopo conversione con bisolfito; purtroppo, non abbiamo confermato il risultato nella coorte di biopsie di RMS. Inoltre, abbiamo effettuato un'ulteriore analisi sui dati di microarray confrontando i RMS metastatici con i non metastatici. Abbiamo trovato un elevato numero di regioni differenzialmente metilate (DMR) e molte di queste sono risultate coincidere con le regioni promotoriali di geni implicati nello sviluppo di tumori; in particolare, abbiamo trovato DMR connesse alla famiglia delle *clustered* protocaderine, note come geni soppressori di

tumore. Abbiamo poi confermato un diverso profilo di espressione del gene *PCDHA4*, così come un diverso stato di metilazione a livello della sua regione promotoriale, confrontando campioni di RMS metastatici e non metastatici. Tuttavia, lo stato di metilazione e il livello di espressione di *PCDHA4* non hanno dimostrato una correlazione significativa con le caratteristiche cliniche del RMS. Il gene *PCDHA4* non risulta quindi essere un predittore prognostico nel RMS.

Successivamente, abbiamo effettuato un sequenziamento RRBS, al fine di validare i dati ottenuti con le piattaforme dei microarray. Ne è risultata una bassa concordanza tra i due approcci, probabilmente a causa della bassa qualità del DNA utilizzato negli esperimenti di microarray. Il sequenziamento RRBS ha dimostrato ancora una volta che i RMS *PAX3-FOXO1*-positivi hanno un profilo di metilazione diverso dai RMS *PAX3-FOXO1*-negativi. Inoltre, abbiamo dimostrato che *GADD45G* e *NELL1*, già descritti come soppressori tumorali in altri tipi di tumore e spesso regolati in maniera negativa da processi di metilazione, sono anche coinvolti nella biologia del RMS. Con i nostri esperimenti abbiamo confermato una regolazione epigenetica, mediata dalla metilazione del DNA ,per i geni *GADD45G* e *NELL1*, e come la loro espressione sia correlata alla istologia del RMS, alla presenza dei geni di fusione e alla stadiazione in gruppi IRS. Inoltre, abbiamo dimostrato che i livelli di espressione di *GADD45G* e *NELL1* influenzano la sopravvivenza libera da progressione di malattia nei pazienti affetti da RMS, suggerendo la loro associazione con una prognosi sfavorevole.

In conclusione, il nostro lavoro ha dimostrato che *GADD45G* e *NELL1* potrebbero essere nuovi potenziali biomarcatori nel RMS, evidenziando come il profilo di metilazione del DNA nel RMS potrebbe favorire lo sviluppo di nuove strategie terapeutiche.

Ci auguriamo che i nostri sforzi possano contribuire ad una migliore classificazione molecolare dei tumori nei pazienti affetti da RMS e alla identificazione di nuovi bersagli farmacologici per una terapia più mirata.

1. INTRODUCTION

1.1. Rhabdomyosarcoma

Rhabdomyosarcoma (RMS) is a soft tissue malignant tumor of mesenchymal origin accounting for 4-8% of all pediatric malignancies (Egas-Bejar and Huh, 2014; Yang *et al.*2014). It is the most prevalent soft tissue tumor in children and adolescents, with an incidence of approximately 350 new cases per million children per year in the USA (Parham and Ellison, 2006; Paulino and Okcu, 2008). In contrast, RMS is extremely rare in adults. There is a slight male predominance (male:female ratio of 1.4:1), but there are no significant differences in the incidence rates among races or different ethnic groups (Breneman *et al.*, 2003).

Since RMS is derived from primitive mesenchymal stem cells directed towards myogenesis, it can arise virtually everywhere in the body, including anatomic sites that lack skeletal muscle, such as biliary and genitourinary tract (Dagher and Helman, 1999). Indeed, the most common primary sites are the genitourinary tract (24%), followed by parameningeal region (16%), limbs (>19%), orbit (9%), other head and neck areas (10%) and assorted sites (22%) (Meyer and Spunt, 2004).

RMS can occur either as a primary malignancy or as a component of a heterogeneous malignancy, such as a malignant teratomatous tumor (Glass *et al.*,2014). Additionally, a small percentage of cases are associated with known genetic disorders, such as neurofibromatosis type 1 and the Li-Fraumeni familial cancer syndrome (Ji *et al.*, 2007; Li and Fraumeni, 1969). Most cases of RMS occur sporadically, with no recognized predisposing factors or risk factors. Genetic conditions associated with RMS include also pleuropulmonary blastoma (with *DICER1* mutations),(Dehner *et al.*, 2012; Doros *et al.*, 2012) Costello syndrome (with germline *HRAS* mutations), (Gripp *et al.*, 2006; Kratz *et al.*, 2011) Beckwith-Wiedemann syndrome (with which Wilms tumor and hepatoblastoma are more commonly associated), (Samuel *et al.*, 1999; DeBaun and Tucker, 1998) and Noonan syndrome. (Kratz *et al.*, 2011; Moschovi *et al.*, 2007; Hasle, 2009).

1.1.2. Histo-pathological classification

RMS is a small, round blue cell neoplasm that must be differentiated from neuroblastoma, the Ewing family of tumors and lymphoma (Gallego Melcòn and Sànchez de Toledo Codina, 2007). The key to the diagnosis of RMS is the presence of markers of skeletal differentiation. Immunohistochemical staining with myogenin, MyoD, muscle-specific actin, myoglobin and/or desmin is commonly required for accurate diagnosis.

The *World Health Organization* (WHO) recently revised the classification of RMS subtypes as alveolar rhabdomyosarcoma (ARMS), embryonal rhabdomyosarcoma (ERMS), pleomorphic rhabdomyosarcoma (PRMS), and sclerosing/spindle cell rhabdomyosarcoma (SRMS) (Fletcher *et al.*, 2013).

ARMS is a high-grade malignancy occurring mostly in adolescents and young adults (costant incidence from ages 0 to 19 years). The most common site for ARMS is in the deep tissue of extremities. It was described for the first time in 1956 by Riopelle and Theriault that revealed a key defining pattern composed by alveolar structure formed by fibrous septa (Riopelle and Theriault, 1956). These fibrous septa are composed of collagenous fibrovascular tissue that forms a scaffolding suspending attached rhabdomyoblasts and circumscribing central clusters of tumour cells that appear to float in alveolar spaces (Parham, 2001).

ERMS represents approximately 70% of all childhood RMS, usually afflicting infants or children under 10 years of age. ERMS often affects the head and neck regions, especially the orbit, and genitourinary tract. The tumor cells composing embryonal RMS variably exhibit all cellular phases of myogenesis, from stellate unidifferentiated mesenchymal cells to elongated myoblasts, multinucleated myotubes and fully differentiated myofibers (Parham, 2001).

Pleomorphic subtype, was first described in 1946 as a variant of RMS (Stout, 1946). It usually occurs in adult males in the deep tissue of extremities, but may occur at any site. In adult patients, the pleomorphic variant is associated with the worst prognosis (Sultan *et al.*, 2009). PRMSs form spindle cell lesions with a whorled pattern and containing cells with enlarged, hyperchromatic nuclei (Pharm, 2001).

1.1.3. Molecular classification

Chromosomal analysis have demonstrated two translocations associated with ARMS: t(2;13)(q35;q14) and t(1;13)(p36;q14) (Lizard-Nacol *et al.*, 1987). Initial studies detected these two gene fusions in 80% of ARMS. Translocation t(2;13)(q35;q14), which occurs in approximately 60-70% of ARMS, and translocation t(1;13)(p36;q14), which occurs less frequently in approximately 10% of ARMS, result in the expression of chimeric transcription factors *PAX3-FOXO1* or *PAX7-FOXO1* (Marshall and Grosveld, 2012) (**Figure 1**). These characteristic chromosomal translocations are adjacent to the 5′ DNA-binding domains of *PAX*, a member of the paired box transcription factor family, and the transactivation domain at the 3′ end of *FOXO1*, a member of the forkhead/HNF-3 transcription factor family. PAX-FOXO1 fusion proteins are tumor specific transcription factors and are expressed at higher levels than their corresponding wild type proteins. PAX-FOXO1 fusion proteins are transcription factors more potent (10 to 100 fold) than PAX3 and PAX7 wild type proteins (Bennicelli *et al.*, 1996).

Overexpression of PAX-FOXO1 induces oncogenic transformation of chicken embryo fibroblasts (Scheidler *et al.,* 1996) and murine NIH3T3 cells (Lam *et al.,* 1999) as assessed by growth in soft agar, but PAX-FOXO1 alone is generally not sufficient for full oncogenic transformation leading to *in vivo* formation of RMS, thus indicating the need for additional genetic lesions. In cooperation with other events, such as decreased expression of p16INK4/p14ARF, telomeres stabilization, MYCN amplification, mutated p53 or mutated HRAS, PAX-FOXO1 transforms human myoblasts into cells able to give rise to ARMS (Naini *et al.,* 2008). PAX3-FOXO1 activity is necessary for maintenance of the transformed phenotype. Gene silencing with antisense oligonucleotides or siRNA induced apoptosis (Bernasconi *et al.,* 1996) and repression of the malignant phenotype *in vitro* (Kikuchi *et al.*, 2007). Oncogenic activity of PAX3-FOXO1 seems to be due both to an altered regulation of targets of wild type PAX3 and to the recruitment of new targets (Linardic, 2008; Begum *et al.*, 2005). Target genes of PAX3-FOXO1, such as *MYCN*, *MYOD1*, *FOXF1*, *KCNN3*, *IGFBP3* and *GADD45A* are involved in the development of nervous and muscular system, besides in regulation of transcription (Mercado e*t al.*, 2008; Robson *et al.*, 2006). Moreover, PAX-FOXO1 is involved in cell proliferation, cell survival/anti-apoptosis and in inhibition of terminal myogenic differentiation (Ahn *et al.*, 2013).

Another novel translocation, t(2;2)(q35;p23), was identified in ARMS biopsy samples by gene expression signatures (Wachtel *et al.*, 2004). The chromosomal translocation generates a fusion protein composed of PAX3 and the nuclear receptor coactivator NCOA1, which has similar transactivation properties as PAX3-FOXO1 (Wachtel *et al.*, 2004). These biologic effects contribute to tumorigenesis by modulating myogenic differentiation, altering growth and apoptotic pathways, and stimulating motility and other metastatic pathways (**Figure 1**).

The remaining 20% of ARMS is *PAX* gene fusion-negative and forms a more heterogeneous group, which remains a challenge to detect due to the lack of consistent chromosomal rearrangements. Whatever, *PAX-FOXO1* negative ARMS has a similar clinical course to ERMS, besides similar biology and expression profile (Davicioni *et al.*, 2006; De Pittà *et al.*, 2006).

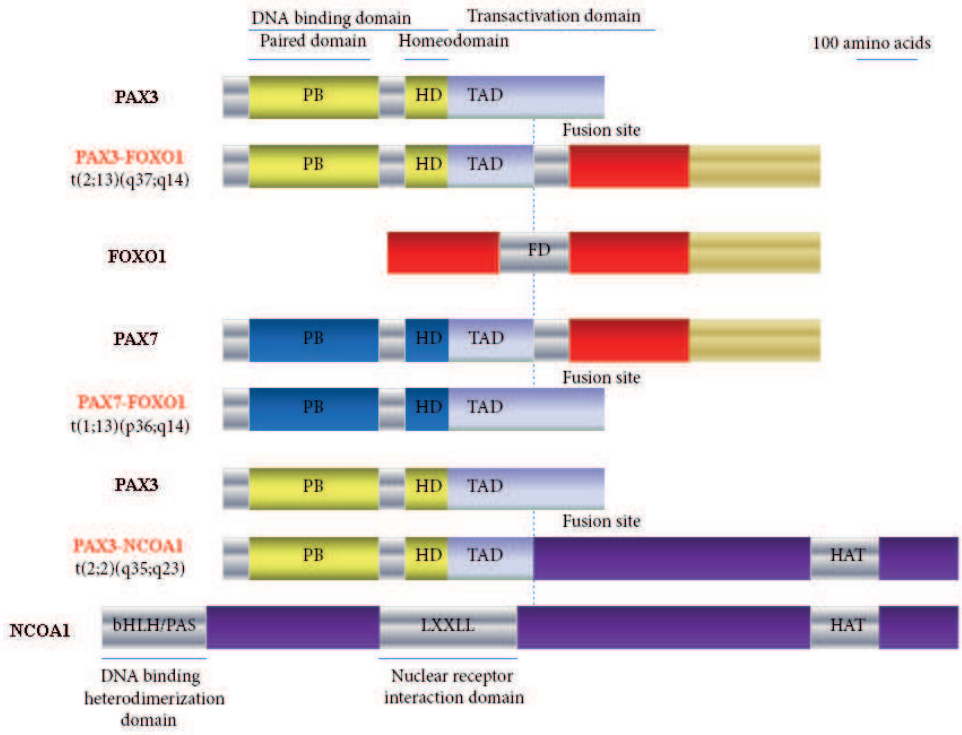


Figure 1. Chromosomal rearrangements in ARMS.

Almost all (94%) RMS show at least one 15-mb region with loss of heterozygosity (LOH), but the overall degree of LOH differed greatly between histological groups (Davicioni *et al.*, 2009), with ERMS showing complex karyotypes, which are characteristics of severe genomic instability (Helman and Meltzer, 2003). Fusion-negative ARMS have allelic imbalance and LOH patterns indistinguishable from conventional ERMS cases (Davicioni *et al.*, 2009). The most frequent LOH regions are found in chromosome 11, both in the long and in the short arm (nearly 70% of all RMS and 80% of ERMS). LOH in 11p15.5 has long been considered a hallmark of ERMS (Visser *et al.*, 1997), and, in a recent study on a wide series of RMS, showed a frequency of 77% in fusion-negative RMS, as compared with only 24% of fusion-positive cases (Davicioni *et al.*, 2009). This LOH region includes the genes *IGF2*, *H19* and *CDKN1C*, all subject to parental imprinting. In normal muscle, imprinting leads to the expression of the paternal allele of IGF2 and the maternal allele of H19 and CDKN1C (p57KIP2), instead LOH in ERMS results in loss of the maternal (silenced) IGF2 allele and duplication of the active paternal allele, resulting in overexpression of IGF2 (Visser *et al.*, 1997; Smith *et al.*, 2007).

Other LOH regions were found in considerable proportions of *PAX-FOXO1* negative RMS (chromosomal region 8q) and *PAX3/7-FOXO1* positive RMS (chromosomal region 4q), increased in RMS with *PAX3-FOXO1* compared with *PAX7-FOXO1* (Davicioni *et al.*, 2009).

Approximately 30% of ERMS show genomic loss at the 9q22 region, which includes the *PTCH1* gene (Bridge et al., 2000).

1.1.4. Risk-group stratification

Three Cooperative Groups have been working in Europe on paediatric soft tissue sarcoma for the last thirty years: the *Malignant Mesenchymal Tumor Committee of the International Society of Pediatric Oncology* (SIOP MMT), the *Cooperative Weichteilsarkomen Studie* (CWS) and the *Associazione Italiana di Ematologia e Oncologia Pediatrica-Soft Tissue Sarcoma Committee* (AIEOP-STSC). Intensification of cooperations had led to the foundation of the *European paediatric Soft tissue Sarcoma Study Group* (E*p*SSG) and the development of a protocol for the treatment of children and young people presenting with non-metastatic RMS: protocol E*p*SSG RMS2005. This protocol is still the one used and contains a randomised trial for "high risk patients" and observational studies for patients categorized in other risk groups. Patients with metastatic RMS are treated according to the "Very High Risk E*p*SSG RMS2005" protocol or the RMS4.99 protocol, for patients with RMS or other soft tissue sarcomas with evidence of metastases at the onset of the disease.

Patients are subdivided in risk-groups taking into account histology (alveolar *vs* non alveolar RMS), post surgical stage (according to IRS grouping, **Table 2**), tumour site and size, node involvement (according to TNM classification, **Table 3**) and patient age. Based on their risk profile four Groups have been identified: Low Risk, Standard Risk, High Risk and Very High Risk (**Table 1**).

This system is a good predictor of patients outcome and allows correlation between intensity of therapy and outcome. For example, overall survival is low and has not improved for patients with high-risk RMS, whereas patients with low-risk RMS have an excellent prognosis. In low-risk patients, investigators are attempting to decrease the intensity of overall therapy by decreasing the duration of therapy and doses of chemotherapeutic agents, without compromising survival, whereas high-risk patients are receiving more intensive therapy with additional drugs at greater doses.

Risk Group	Subgroups	Pathology	Post surgical Stage (IRS Group)	Site	Node Stage	Size & Age
Low Risk	A	Favourable	I	Any	N0	Favourable
Standard Risk	B	Favourable	I	Any	N0	Unfavourable
	C	Favourable	II, III	Favourable	N0	Any
	D	Favourable	II, III	Unfavourable	N0	Favourable
High Risk	E	Favourable	II, III	Unfavourable	N0	Unfavourable
	F	Favourable	II, III	Any	N1	Any
	G	Unfavourable	I, II, III	Any	N0	Any
Very High Risk	H	Unfavourable	I, II, III	Any	N1	Any

Table 1. Risk stratification for E*p*SSG non metastatic RMS study (Protocol E*p*SSG RMS2005).

Pathology:

- Favourable = all embryonal and spindle cells RMS
- Unfavourable = all alveolar RMS (including the solid-alveolar variant)

Post surgical stage:

The postsurgical staging system (IRS grouping) defines the extent of residual disease after resection, that is one of the most important prognostic factors in RMS. Patients are assigned to a clinical group based on the completeness of tumor excision and the evidence of tumor metastasis to the lymph nodes or distant organs after pathologic examination of surgical specimens (**Table 2**).

IRS Group	Definition
I	Tumour macroscopically and microscopically removed
I A	Tumour confined to organ or tissue of origin
I B	Tumour not confined to organ or tissue of origin
II	Macroscopic complete resection but microscopic residuals
II A	Lymph nodes not affected
II B	Lymph nodes affected but removed
III	Macroscopic complete resection but microscopic residuals and lymph nodes affected and not removed
	Macroscopic residuals after resection or biopsy with malignant effusion
IV	Metastasis present or non-regional lymph nodes involved

Table 2. IRS clinical grouping classification (Protocol E*p*SSG RMS2005).

Site:

- Favourable = orbit, genitourinary non bladder prostate (such as paratesticular and vagina/uterus) and non parameningeal head and neck
- Unfavourable = all other sites (parameningeal, extremities, genitourinary bladder-prostate and "other site")

Node stage:

Node stage is defined according to the tumor, nodes, metastases (TNM) classification system (**Table 3**). This staging is based on the preoperative workup of imaging and physical examination and as the name suggests is defined by the degree of tumor invasion, the nodal status and the presence or absence of metastases.

Tumour		Lymph nodes		Metastasis	
T0	No evidence of tumour	**N0**	No evidence of lymph node involvement	**M0**	No evidence of metastases or non-regional lymph nodes
T1	Tumour confined to organ or tissue of origin	**N1**	Evidence of regional lymph node involvement	**M1**	Evidence of distant metastasis or involvement of non-regional lymph nodes
T1a	Tumour = 5 cm in greatest dimension	**NX**	No information on lymph node involvement	**MX**	No information on metastasis
T1b	Tumour > 5 cm in greatest dimension				
T2	Tumour not confined to organ or tissue of origin				
T2a	Tumour = 5 cm in greatest dimension				
T2b	Tumour > 5 cm in greatest dimension				
TX	No information on size and tumour invasiveness				

Table 1. Pretreatment TNM staging classification (Protocol E*p*SSG RMS2005).

Size and Age:

- Favourable = Tumour size (maximum dimension) <5cm and Age <10 years
- Unfavourable = all others (such as Size >5 cm **or** Age =10 years)

1.1.5. Prognosis

The prognosis of patients with RMS depends on many factors. The main prognostic parameters for RMS are histologic subtype, site of onset, tumor size (widest diameter), age of patient, presence of *PAX3/7-FOXO1* fusion status, presence of metastases, number of metastatic sites or tissues involved, presence or absence of regional lymph node involvement (Parham and Ellison, 2006; Meza *et al.*, 2006; Sultan *et al.*, 2009).

Favourable sites are head and neck (non-parameningeal), genitourinary (excluding bladder or prostate) and bile duct regions. All other sites are classified as unfavourable. ARMS histology is an unfavourable prognostic factors. The 5-year overall survival is approximately 73% for ERMS and 48% for ARMS, ranging from 10-30% for metastatic ARMS to over 95% for localized forms of ERMS. ARMS show a more aggressive behaviour, they are often metastatic at the diagnosis, have a poor response to therapy and a worse prognosis than ERMS. Tumor size is an integral prognostic variable for RMS and plays a major role in clinical grouping. Clinical grouping has also been identified as one of the most important predictors of failed treatment and tumor relapse. These factors become important in the designation of treatment groups for risk-based therapy.

1.2. Epigenetics

The character of a cell is defined by its constituent proteins, which are the result of specific patterns of gene expression. Crucial determinants of gene expression patterns are DNA-binding transcription factors that choose genes for transcriptional activation or repression by recognizing the sequence of DNA bases in their promotorial regions. Interaction of these factors with their cognate sequences triggers a chain of events, often involving changes in the structure of chromatin, that leads to the assembly of an active transcription complex (Cosma *et al.*, 1999). Then, the types of transcription factors present in a cell are not alone sufficient to define its spectrum of gene activity, but different mechanisms contribute to modulate and define the trascriptome of the cells. Epigenetics can be described as all heritable alterations in gene expression potential and chromatin structure, that take place during development and cell proliferation, due to chemical modifications that do not involve changes in the primary gene nucleotide sequence (Wu and Morris, 2001). Epigenetic events explain how two identical genotypes can give rise to different phenotypes in response to the same environmental stimulus. There are four types of mechanistic layers in the field of epigenetics: DNA methylation, post-translational modifications of histone proteins and noncoding RNAs. Disruption of any of these distinct and mutually reinforcing epigenetic mechanisms leads to inappropriate gene expression, resulting in cancer development and other epigenetic diseases (Egger *et al.*, 2004).

For many years cancer research has focused on genetic defects, but during the last decade epigenetic deregulation has been increasingly recognized as a hallmark of cancer. The advent of genome-scale analysis techniques, including the next-generation sequencing, has enabled an invaluable advance in the molecular mechanisms underlying tumor initiation, progression, and expansion, besides advances in the field of cancer epigenomics concerning DNA methylation and its role in oncogenesis.

1.2.1 DNA methylation

DNA methylation is one of the most commonly occurring epigenetic events taking place in the mammalian genome and the first epigenetic mark to be associated with cancer, because modification of DNA methylation causes a different gene regulation (Feinberg and Vogelstain, 1983).

DNA methylation is the covalent chemical addition of methyl (CH3) group at the carbon 5 position of the cytosine ring, resulting in the 5-methylcytosine (5mC). The presence of 5mC was first demonstrated in tubercle bacillus DNA in 1925 (Johnson and Coghill, 1925). Although biological functions of this cytosine modification remained uncharacterised for decades, in 1975, two studies demonstrated important roles of 5mC as an epigenetic modification that influences gene expression (Riggs, 1975; Holliday and Pugh, 1975) and highlighted the significance of the "fifth nucleotide" in eukaryotic biology (Doerfler, 2006).

In prokaryotes, methylation at both adenine (A) and cytosine (C) residues contributes to host restriction systems and protects the cell from foreign genetic materials such as bacterial and viral genomes (Bickle and Kruger, 1993). In contrast, DNA methylation in multicellular eukaryotes occurs predominantly, but not exclusively, at cytosine residues within the strict sequence context composed by 5'CG3', also called CpG dinucleotide (Bestor, 2000). In vertebrates DNA methylation is involved in a variety of biological processes, including modulation of gene expression, embryonic development, inactivation of transposable elements, chromatin structure, X-chromosome inactivation, genomic imprinting, chromosome stability and maintenance of epigenetic memory (Bird, 2002).

CpG dinucleotide is under-represented in the mammalian genome; it occurs at only about 1/5 of the roughly 4% frequency that would be expected in human DNA, because approximately 70% of CpG dinucleotides are methylated on the cytosine base, and spontaneous hydrolytic deamination of 5mC residues gives rise to T residues (Lander *et al.,* 2001). In contrast, short stretches of CpG rich regions, known as CpG islands (CGIs), which are found in the promoters of approximately 60% of the coding genes in the mammalian genome, are generally unmethylated in normal cells (Cooper, 1983; Bird *et al.*, 1985; Gardiner-Garden *et al.*, 1987). CpG islands, ranging from 0.5 to 5 kb and occurring on average every 100 kb, have distinctive properties. These regions are GC rich (60% to 70%), have a ratio of CpG to GpC of at least 0.6, and thus do not show any suppression of the frequency of the CpG dinucleotide differently from the rest of the genome (Antequera and Bird, 1993; Cross and Bird, 1995). However, the hypermethylation of these promoter regions is found in virtually every type of human cancer and is associated with the inappropriate transcriptional silencing of genes (Jones and Laird 1999; Jones and Baylin 2002; Herman and Baylin 2003).

1.2.1.1 DNA methyltransferases

The establishment and the maintenance of DNA methylation patterns in mammals is performed by a group of enzymes known as DNA methyltransferase (DNMTs). They catalyze the transfer of the methyl group from S-adenosyl-methionine (SAM), the methyl donor, to the cytosine at CpGs sites (Okano *et al.*, 1998).

In mammals, three catalytically active DNA methyltransferases are present: DNMT1, DNMT3a and DNMT3b. They share some common features, such as a N-terminal regulating part and a C-terminal catalytic one. The N-terminal portion is characterized of more than one domain and it can have variable size. It has regulatory functions, indeed it guides nuclear localization and mediates interaction with DNA, chromatin or other proteins. The C-terminal part has catalytic functions and is particularly conserved among different species (Lauster *et al.*, 1989; Robertson, 2001). It includes ten sequence motifs (I to X) which form the binding site for the cofactor S-Adenosyl-L-Methionine (SAM) and three motifs (IV, VI and VII) which play catalytic functions.

1.2.1.1.1 DNMT3A and DNMT3B

DNMT3A and DNMT3B are *de novo* methyltransferases that target cytosine within CpG dinucleotides. These enzymes act on hemimethylated or unmethylated DNA, which are essential for their roles in *de novo* methylation of the genome during development (Okano *et al.*, 1999).

Following the wave of genome-wide demethylation during the preimplantation embryo, DNMT3A and DNMT3B are highly expressed at implantation and re-establish a bimodal methylation pattern that effects more than 80% of the genome (Okano *et al.*, 1999), whereas most CGIs are protected by unknown mechanisms and therefore remain unmodified. Genetic and functional analyses indicate that DNMT3A and DNMT3B have different functions during development with different phenotypes and lethality stages (Okano *et al.*, 1999). This suggests that each enzyme has regional specificity that reflects their respective N-terminal domains. Accordingly, DNMT3A is necessary for maternal imprinting at differentially methylated regions, and DNMT3B is required for methylation of pericentromeric repeats and CGIs on inactive X-chromosomes (Kim *et al.*, 2009).

1.2.1.1.2 DNMT1

DNMT1 was the first mammalian DNA methyltransferase cloned (Bestor, 1988) and biochemically characterised. It works for the preservation of the methylation patterns during cell divisions. Indeed, it is known as a maintenance enzyme that shows a preference for

hemimethylated DNA over unmethylated DNA (Fatemi *et al.*, 2001; Goyal *et al.*, 2006) and a cellular localisation close to DNA replication foci during the S phase. DNMT1 operates with its co-factor UHRF1 (Np95), constituting an enzymatic platform that provides a maintenance methyltransferase function for CpG methylation (Bostick *et al.*, 2007).

1.2.1.1.3 DNMT2 and DNMT3L

In addition to the canonical DNA methyltransferase there are other minor, such as DNMT2 and DNMT3L. DNMT2 has weak methyltransferase activity *in vitro*, and its depletion has little impact on global CpG methylation levels and no discernible effects on developmental phenotypes (Hermann *et al.*, 2003). DNMT3L (DNMT3-like) does not have methyltransferase activity because it is catalytically inactive. It is highly expressed in germ and ES cells and acts as an obligatory cofactor for *de novo* methyltransferase in ES cells. DNMT3L stimulates the methyltransferase activity of DNMT3a or DNMT3b through physical interaction (Kareta *et al.*, 2006). Crystallographic analyses of DNMT3a and DNMT3L indicate that these interactions may be mediated by a heterotetrameric complex formation, which may prevent DNMT3a oligomerisation and heterochromatic localisation. A recent study showed that DNMT3L is a positive regulator of DNA methylation at gene bodies of housekeeping genes and a negative regulator of DNA methylation at promoters of bivalent genes in mouse ES cells, suggesting a dual role in ES cell differentiation (Neri *et al.*, 2013).

In addition to the DNMTs, the others machinery involved in the process of DNA methylation include demethylases, methylation centers triggering DNA methylation, and methylation protection centers (Costello and Plass, 2001; Szyf, 2003).

1.2.2 DNA methylation and gene regulation

The regulation of eukaryotic gene expression is a complex process. Transcription initiation is a highly controlled and integrated event that involves *cis*-acting and *trans*-acting factors. The *cis*-acting elements are DNA sequences that act as the substrate for the *trans*-acting factors, and the DNA in the vicinity is prepared for transcription. Generally, increased methylation in the promoter region of a gene leads to reduced expression, whereas methylation in the transcribed region has a variable effect on gene expression (Jones, 1999; Singal *et al.*, 2002).

Several mechanisms have been proposed to account for transcriptional repression by DNA methylation. The first mechanism involves direct interference with the binding of specific transcription factors to their recognition sites in their respective promoters (**Figure 2**). Several

transcription factors, such as AP-2, c-Myc/Myn, the cyclic AMP-dependent activator CREB, E2F, and NFkB, recognize sequences that contain CpG residues, and the binding of each has been shown to be inhibited by methylation (Singal and Ginder, 1999; Tate and Bird, 1993).

The second mode of repression involves a direct binding of specific transcriptional repressors to methylated DNA (**Figure 2**). The DNA methylation signals are recognized by a family of proteins know as methyl-CpG–binding proteins (MBPs), that target the 5′ methylated CpG sequence (Prokhortchouk and Hendrich, 2002). MBPs elicit the repressive potential of methylated DNA. MBPs fall into two families: first, the methyl-CpG binding domain (MBD) proteins, MeCP2, MBDl, MBD2, MBD3, and MBD4, which share an approximately 80 amino acid MBD and repress transcription through a transcriptional repression domain (TRD) (Hendrich and Bird, 1998), and second, the Kaiso-like proteins, Kaiso, ZBTB4 (zinc finger and BTB domain containing 4), and ZBTB38, which lack the MBD, but recognize DNA sequences containing methyl-CpG sequences through a zinc-finger domain and repress transcription through a POZ/BTB domain (Prokhortchouk *et al.,* 2001; Filion *et al.,* 2006).

DNA methylation can also affect histone modifications and chromatin structure, which, in turn, can alter gene expression. The underlying patterns of methylated cytosine are important in guiding histone deacetylation to certain residues (Bender, 2004).

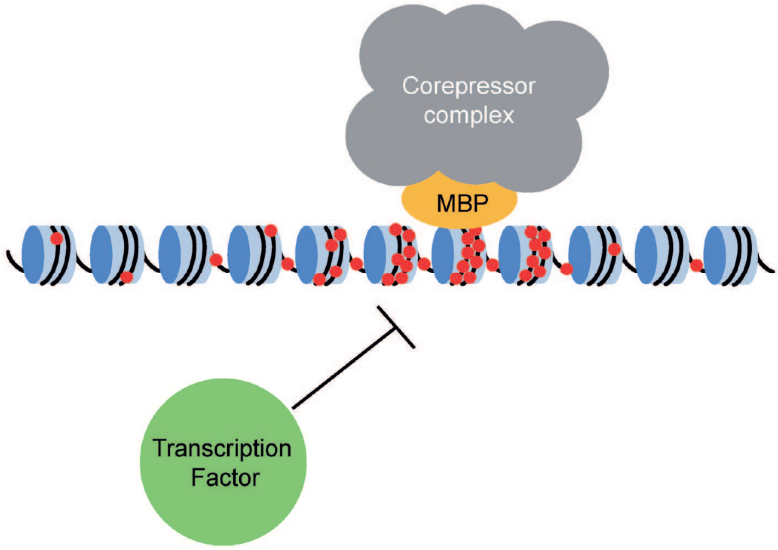


Figure 2. Mechanisms of transcriptional repression by DNA methylation.

1.2.3 DNA methylation and cancer

Earlier it was thought that normal cells become progressively transformed to malignant cells as a consequence of damage to the genome, which could be a gain, loss, or mutation of the genetic

information. These events cause critical loss of gene activity and thereby predispose to cancer. DNA methylation can modify the gene activity without changing the gene sequence and has been proposed as one of the two hits in Knudson's two hits hypothesis for oncogenic transformation (Jones and Laird, 1999).

Aberrant DNA methylation was the first epigenetic mark to be associated with cancer as a consequence of the alteration it causes in normal gene regulation (Feinberg and Vogelstain, 1983). As compared with normal cells, the malignant cells show major disruptions in their DNA methylation patterns (Baylin and Herman, 2000). These alterations are of two types: hypermethylation and hypomethylation. Hypomethylation usually involves repeated DNA sequences, such as long interspersed nuclear elements, whereas hypermethylation involves CpG islands (Ehrlich, 2002).

DNA hypermethylation refers principally to the gain of methylation at specific sites that are unmethylated under normal conditions. This aberrant methylation occurs mainly in promoter CpG islands (CGIs). This phenomenon of aberrant promoter CGI hypermethylation has been associated with the stabilization of transcriptional repression and loss of gene function, and occurs fundamentally in tumor suppressor genes (Bird, 1992; Esteller, 2008) (**Figure 3**). It is believed that epigenetic gene inactivation is at least as common as, if not more frequent than, mutational events in the development of cancer (Jones and Baylin, 2002). There are several protective mechanisms that prevent the hypermethylation of the CpG islands. These include active transcription, active demethylation, replication timing, and local chromatin structure preventing access to the DNA methyltransferase (Clark and Melki, 2002). To date, numerous genes have been found to undergo hypermethylation in cancer. The genes that are susceptible are the genes involved in cell cycle regulation (*$p16^{INK4a}$, $p15^{INK4a}$, Rb, p^{14ARF}*), genes associated with DNA repair (*BRCA1, MGMT*), apoptosis (*DAPK, TMS1*), drug resistance, detoxification, differentiation, angiogenesis, and metastasis (Sugimura and Ushijima, 2000).

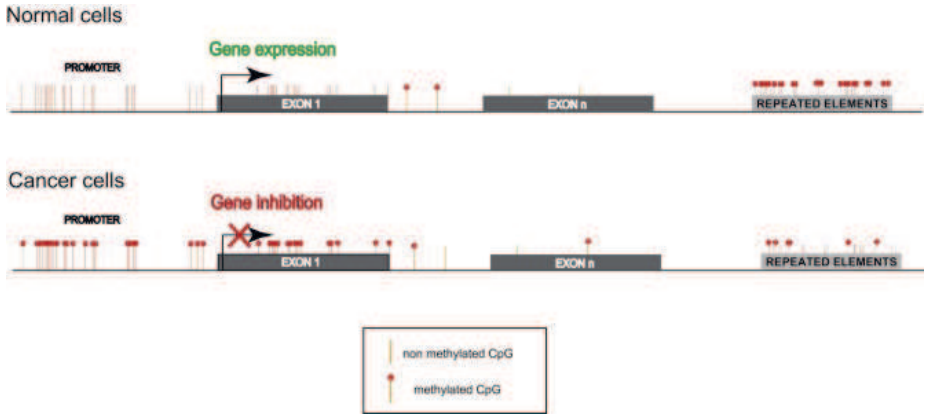


Figure 3: DNA methylation in normal and cancer cells. Stricks represent CpG sites and the red circles represent mehtylated CpGs.

DNA hypermethylated genes frequently targets genes involved in normal development. For example, these genes include cyclin-dependent kinase inhibitor 2A (CDKN2A), which regulates stem cells number and cell cycle functions (Park and Lee, 2003), GATA binding protein 4 and 5 (GATA-4and -5), which are crucial for proper epithelial differentiation (Akiyama *et al.*, 2003), E-cadherin (CDH1),which controls cell-cell adhesion (Hirohashi, 1998), and death-associated protein kinase (DAPK), which functions as anti-apoptotic gene (Michie *et al.*, 2010). These observations support the idea of a stem or precursor cell of origin for cancer and explain the targeting of at least some genes for DNA methylation and gene silencing during tumor initiation and progression. Thus, genes that undergo to abnormal gene silencing are called epigenetic gatekeepers because they may help push the early aberrant clonal expansion of cells, providing a substrate for risk of subsequent genetic and epigenetic alterations that further foster tumor progression (Baylin and Ohm, 2006; Feinberg *et al.*, 2006).

Furthermore, like the majority of gene mutations, the roles of the hundreds of DNA hypermethylated genes other than genes functioning as tumor suppressors may be their aggregation in the same signaling pathway, thus helping derive the cancer phenotype.

DNA hypomethylation is associated mainly with the losses of DNA methylation in genome-wide regions, although it can also occur locally. A number of studies have described DNA hypomethylation in several tumor types, such as colorectal and gastric cancers, melanomas, among others (Kulis and Esteller, 2010). DNA hypomethylation occurs in many gene-poor genomic areas, including repetitive elements, retrotransposons and introns, where it leads to genomic instability (Esteller, 2008). In repeat sequences, this is achieved by a higher rate of chromosomal rearrangements and, in retrotransposons, by a higher probability of translocation to other genomic regions (Eden *et al.*, 2003; Rodriguez *et al.*, 2006). During tumor progression, the degree of hypomethylation of genomic DNA increases as the lesion derives from a benign proliferation of cells to an invasive cancer (Fraga *et al.*, 2004).

1.2.4 DNA methylation as a marker for tumor diagnosis and prognosis

Several knowledge about altered methylation patterns in human cancers has been gained. Tumor-specific methylation changes in different genes have been identified and documented. The potential clinical application of this information is in cancer diagnosis, prognosis and therapeutics.

An early diagnosis is critical for the successful treatment of many types of cancer. The discovery of new molecular markers can improve the classification of tumours and then could help addressing of the standard therapy. The clinical utilization of methylated biomarkers for the early diagnosis of tumors, such as for colorectal cancer and lung carcinoma, is already underway for several years. Hypermethylation of *SHOX2* in bronchial aspirates is a clinically

tumor biomarker used in routine diagnostics for identifying subjects with lung carcinoma, especially when histological and cytological findings after bronchoscopy are ambiguous (Schmidt *et al.*, 2010). *SEPT9* gene methylation has been used as a biomarker for colorectal cancer (CRC) for more than 10 years and has been used clinically for more than 6 years (Song and Li, 2015). It permits the screening of CRC with a blood-based approach, allowing early diagnosis in a noninvasively way.

Because DNA methylation is closely related to the development of cancer, it would be interesting to know whether its presence or absence affects the prognosis as well. Many studies have shown several methylated genes to be closely related to the prognosis. For example, methylation of the promoter region of four genes, p16, H-cadherin, Ras association domain family 1A and APC, in primary tumor of non-small-cell lung cancer (NSCLC) and mediastinal lymph node biopsy samples strongly correlates with early recurrence and short survival (Brock *et al.*, 2008). The current method for risk assessment of recurrence in patients with stage I NSCLC is imprecise. The validation of these findings may allow re-staging of NSCLC at the molecular level and may thus identify high risk patients who require special adjuvant therapies.

Methylation profile may also help in predicting response to a chemotherapeutic agent. Methylation of the promoter region of the DNA repair gene O^6-methylguanine-DNA methyltransferase (*MGMT)* is associated with the sensitivity of gliomas to alkylating agents, determining a better response to the treatment (Esteller *et al.*, 2000; Hegi *et al.*, 2005).

1.2.5 DNA methylation as therapeutic target

Given the critical role of DNMTs, intense interest has focused on developing drugs able to interfere with aberrant DNMT activities, and using them to correct epigenetic defects such as tumor suppression gene (TSG) silencing. DNMTs modulators represent a useful tool in epigenetic therapies. Several epi-drugs, interfering with DNMT activity, are currently in pre-clinical and clinical trials (Foulks *et al.*, 2012). Most of these trials have involved various types of cancer, such as solid and hematological tumors (Chaib *et al.*, 2011; Fandy, 2009; Song *et al.*, 2011). Currently, however, the main challenge in using epigenetic modulators for therapy, especially for interfering with DNMT enzymes is their specificity (Veeck and Esteller, 2010).

The reversibility of epigenetic changes, unlike genetic modifications, make them powerful therapeutic targets. Moreover, since methylation of CpG islands occurs infrequently in normal cells, methylation provides a selective tumor-specific therapeutic target. Pharmacologic inhibition of methylation-mediated suppression could therefore derepress inappropriately silenced genes and restore normal gene function. Several clinical trials are presently underway using azacytidine,

histone deacetylase inhibitors, and phenylbutyrates to reactivate the silenced genes in hematologic (Lubbert, 2000) and solid tumors (Brown and Strathdee, 2002).

The commonly used drugs targeting methylation are azacytidine (5-azacytidine), decitabine (5-aza-2′-deoxycytidine), fazarabine (1-β-D-arabinofurasonyl-5-azacytosine), and dihydro-5-azacytidine (Goffin and Eisenhauer, 2002). These are all derivatives of deoxycytidine with some modification at the fifth position in the pyrimidine ring. Other drugs include zebularine (Cheng *et al.*, 2003) and antisense oligodeoxynucleotides. Histone deacetylase (HDAC) inhibitors are also being tried as potential chemotherapeutic agents (Thiagalingam *et al.*, 2003).

To date, only azacitidine and decitabine are the hypomethylating agents approve by FDA (Food and Drug Administration) and EMA (European Medicines Agency) for the treatment of myelodysplastic syndrome (MDS), acute myeloid leukemia (AML), and chronic myelomonocytic leukemia (CMML). Azacitidine is incorporated into RNA, where it suppresses RNA synthesis and has cytotoxic effects, while 5-aza-2'-deoxycytidine (the deoxy derivative of azacitidine) is incorporated into DNA in place of the natural base cytosine. Because of the substitution of the 5' nitrogen atom in place of the carbon, the DNMTs are trapped on the substituted DNA strand and methylation is inhibited.

Clinical studies clearly demonstrate that such agents, used alone or in combination, have clinical benefit in many patients with hematologic malignancies (Daskalakis *et al.*, 2002; Ruter *et al.*, 2004). While these exciting results strongly suggest that the clinical activity of agents, such as azacitidine and decitabine, is mediated by targeting reversal of gene silencing, formal proof of this causal relationship and identification of the critical gene targets are still lacking. Future research will undoubtedly provide this evidence and suggest additional therapeutic targets regulating methylation and histone acetylation pathways in cancer.

1.3. Methods for DNA methylation analysis

Over the past decade, several techniques have been developed to assess the level of methylation in genomic DNA, isolated from different sources, on a genome-wide or gene-specific basis (Tost and Herman, 2009). Essentially, these can be divided into two categories, typing and profiling technologies, which are both suitable for high-throughput applications. Typing technologies are used when few loci need to be assayed in many samples and exploit bisulfite conversion techniques, restriction enzymes and specific PCR assays. Profiling technologies arise with the advent of high throughput methodologies such as microarray, but particularly next generation sequencing. They allowed mapping DNA methylation on a genome-wide scale, at a high resolution and in a large number of samples (Laird, 2010).

1.3.1. Gene-target techniques

1.3.1.1. Restriction enzyme-based methods

The use of restriction enzymes was the first technology used for the investigation of DNA methylation. Restriction endonucleases are such powerful tools in molecular biology that, their biological role in modification systems in bacteria and archaea, is sometimes overlooked. Each sequence-specific restriction enzyme has an accompanying DNA methyltransferase that protects the endogenous DNA from the restriction defence system by methylating bases in the recognition site. The 5mC-methylation-sensitive restriction enzymes (MSRE) are inhibited by 5mC in the sequence context containing CpG motifs in their recognition sequence and can only cut unmethylated sites. Thus, methylated DNA is protected from cleavage (Hashimoto *et al.*, 2007). As a consequence the patterns of cutting by such enzymes can provide a read-out of DNA methylation.

The most widely used MSRE for DNA methylation studies are HpaII and SmaI, in part because they each have an isoschizomer (MspI for HpaII) or neoschizomer (XmaI for SmaI) that is not inhibited by CpG methylation. Moreover, a handful of restriction enzymes are available to cut methylated but not unmethylated DNA, such as GlaI, McrBC and SgeI (Tarasova *et al.*, 2008). Methylation-sensitive restriction digestion followed by PCR across the restriction site is a very sensitive technique that is still used in some applications today. However, it is extremely prone to false-positive results caused by incomplete digestion for reasons other than DNA methylation. A multiple restriction enzymes combined with quantitative PCR analysis might overcome the false-positive results. An assays covering at least three restriction enzyme sites per target region have been developed and successfully applied using the MSRE technology to elucidate aberrantly methylated genes in syndromal disease (Weinhaeusel *et al.*, 2008) and different cancers, such as breast, lung and colon (Agrawal *et al*., 2007). The results have been successfully validated by bisulfite-based approaches confirming the restriction enzyme-based findings.

1.3.1.2. Bisulfite deamination-based methods

Sodium bisulfite ($NaHSO_3$) treatment of genomic DNA revolutionized DNA methylation studies (Frommer *et al.*, 1992). Indeed, bisulfite-based approaches are the most commonly used for qualitative and quantitative DNA methylation measures enabling the assessment of absolute methylation levels at single-base resolution. Incubation of genomic DNA with sodium bisulfite leads to deamination of unmethylated cytosine to uracil, whereas methylated cytosine are not affected (Frommer *et al.*, 1992). This chemical treatment of DNA effectively turns an epigenetic difference into a genetic difference. Thus by the subsequent PCR of interested region are

generated amplicons containing a C (derived from a 5mC) or a T (derived only from unmethylated cytosine) depending on the original methylation status of the sample.

1.3.1.2.1 Bisulfite sequencing PCR

The gold standard for assessing the DNA methylation status on a single-allele basis of every given CpG, within a region of interest, is bisulfite genomic sequencing (BSP). In this method, the interested region is amplified from bisulfite-modified DNA, using PCR primers not overlapping CpG sites, in order to amplify both methylated and unmethylated alleles (Clark et al., 1994; Frommer et al., 1992). PCR products are ligated to a cloning vector and transfected to competent cells. The plasmidic DNA is isolated and subjected to sequencing to generate detailed methylation patterns at single CpG resolution. This technique is labour-intensive and quite expensive for large sets of samples. These inconveniences were be overcome with the advent of bisulfite sequencing pyrosequencing, whom principles are the same of BSP and permits quantitative analyses of single methylation sites.

1.3.1.2.2 Bisulfite-pyrosequencing

Using bisulfite-specific primers and bisulfite sequencing PCR amplification, quantitative methylation analyses on a single CpG can be determined using BSP, when single colonies of cloned PCR amplicons are sequenced and 5mC per CpG are deduced (Grigg and Clark, 1994). This principle, that relies on bisulfite conversion and PCR amplification (Colella *et al.*, 2003) is also used in the pyrosequencing approach, but omitting the cloning step. To facilitate the conversion of PCR products to single-stranded DNA for later pyrosequencing, the PCR reaction is performed with either one primer biotinylated or using a tailed primer in combination with a biotin-labeled universal primer in the same reaction (this avoids biotin-labeling for each primer for each assay). The sequencing primer is then annealed to single-stranded DNA and the samples are ready for pyrosequencing analysis. Pyrosequencing is a primer extension method for the analysis of short to medium length DNA sequences. Incorporation of a nucleotide into the template strand leads to the release of pyrophosphate, which is quantified with a luciferase reaction. The signal produced is proportional to the amount of pyrophosphate released, therefore the percentage of unconverted C and converted T nucleotides at each CpG site can be detected and quantified. This method has the advantages of introducing an internal control (DNA sequence including a control for unconverted cytosines) and allowing accurate quantification of multiple CpG methylation sites in the same reaction. The only significant drawback is that few (only 25–

30) base pairs can be sequenced in each reaction, limiting the number of CpG sites that can be assessed.

1.3.1.2.3. Methylation-specific PCR

Methylation specific-PCR (MS-PCR or MSP) is the most commonly used technique to analyze DNA methylation levels. It requires two parallel PCR reactions amplifying either methylated or unmethylated alleles from bisulfite-converted DNA using primers annealing to CG versus TG sequence, respectively (Herman *et al.*, 1996). The detection of PCR products is originally performed by gel electrophoresis. MSP, combined with end point PCR, is particularly prone to amplification bias and methylation artifacts. This technique has been replaced by quantitative MS-PCR (qMS-PCR), in which PCR amplification is monitored in real time by the incorporation of fluorescent molecules. This improvement allows precise quantification of the DNA methylation levels of numerous specific regions and avoids the long electrophoresis step. Since it only provides the methylation status of few CpG sites (contained in PCR primers), qMS-PCR requires perfect knowledge of the most discriminative methylated regions present for example in cancer cells, to design powerful primers for diagnosis. Special care has to be taken when designing the primers to obtain equal amplification efficiencies and to avoid a bias towards unmethylated or methylated DNA. As these approaches provide quantitative measurements of DNA methylation, it is necessary to define a cut-off DNA methylation value before declaring that a sample is positive (Herman *et al.*, 1996). Nevertheless, qMS-PCR technique is simple, rapid, inexpensive, highly-sensitive and easily standardized. It is currently one of the most commonly used techniques for cancer diagnosis in clinical use.

1.3.1.2.4. Methylation-sensitive high-resolution melting

Methylation-sensitive high-resolution melting (MS-HRM) is based on the fact that the nucleotide sequence of PCR products of bisulfite-treated DNA will differ depending on the methylation status of the DNA region of interest. Thus, the methylation level is determined by comparing the melting dissociation curves to standard PCR products of the same region containing known methylated CpG sites. Despite its high sensitivity, this method requires the acquisition of specific PCR apparatus and skilled operators.

1.3.2. Global DNA methylation profiling: the methylome

1.3.2.1. Gel-based methylation profiling

Restriction landmark genomic scanning (RLGS), developed in 1991, was the first genome-wide profiling method that allowed some genomic positional information about where changes in methylation have occurred (Hatada *et al.*, 1991). This was achieved using either methylation-sensitive endonucleases, which cut preferentially in certain genomic regions such as CpG islands, or computational prediction of the respective restriction fragments where genome sequences were available (Rouillard *et al.*, 2001). A key feature of RLGS entails the use of two-dimensional electrophoresis, which significantly increases the resolution of potential target sites and enable to interrogate more than 1000 CpG islands in a single experiment. In 2000, RLGS was used to survey aberrant CpG island methylation in 98 primary human tumours (Costello *et al.*, 2000), providing the first genome-wide insight into the global changes of DNA methylation affecting CpG islands in cancer.

1.3.2.2. Array-based methylation profiling

The development of microarrays in the mid-1990s opened the door for methylome profiling, permitting to researchers to assess locus-specific DNA methylation on a genome-wide scale. Three important DNA methylation analysing techniques were adapted for array-based DNA methylation profiling: bisulphite conversion, methylation-sensitive restriction, and immunoprecipiation.

1.3.2.2.1. Bisulfite-based array for DNA methylation profiling

The first array-based DNA methylation profiling was described in 2002 (Gitan *et al.*, 2002) and involved conventional sodium bisulphite treatment of genomic DNA, followed by PCR amplification of regions of interest (about 300–400 bp in size). The bisulfite converted PCR products were hybridized to custom microarrays that contained probes to discriminate converted versus unconverted cytosine at the CpG site of interest, thereby providing a readout of the original methylation state at that CpG site. Although this approach potentially offers single-CpG resolution, the major drawback is that, after bisulphite conversion, most cytosine are converted to thymine, mostly the ones restricted to CpG sites and this results in a reduction of sequence complexity and consequently an increasing difficult to design enough unique probes to extend this approach to a genome-wide scale. Therefore, it remains to be determined whether whole mammalian genomes can be assayed using this technique.

Bisulphite conversion-based DNA methylation profiling on arrays seems not well suited for *de novo* delineation of DNA methylation profiles, but it could be useful for high-throughput validation or follow-up studies of a limited number of CpG sites in hundreds to thousands of samples. Nowadays, the most commonly platforms used with these features is the Infinium HumanMethylation450 BeadChip (Illumina), which is based on whole genome amplification of bisulfite-converted DNA, followed by fragmentation and hybridization to chip. On the chip, there are only two bead types for each CpG site per locus. One of the bead types will correspond to the methylated cytosine locus and the other will correspond to the unmethylated cytosine locus, which has been converted into uracil during bisulfite treatment and later amplified as thymine during whole genome amplification. Hybridization is followed by single-base extension and after staining, the chip is scanned to show the intensities of the unmethylated and methylated bead types.

1.3.2.2.2. Methylation-sensitive enzymes-based array for DNA methylation profiling

Tompa *et al*. in 2002 (Tompa *et al.*, 2002) developed a new method consisted in fragmentation of DNA by the methylsensitive restriction endonuclease (MSRE) followed by size fractionation and hybridization to custom microarrays. Recently, such approaches have been improved with respect to sensitivity and genomic coverage than bisulphite-based methods coupled to single-nucleotide resolution (because in theory it is known where the enzyme has restricted the DNA) (Khulan *et al.*, 2006; Schumacher *et al.*, 2006). The drawback is that only those regions that contain the restriction site of interest can be analysed and hence can never really be used for truly whole genome profiling.

1.3.2.2.3. Enrichment-based array for DNA methylation profiling

Two distinct approaches were developed to enrich the fraction of DNA methylated with respect to all genomic DNA. The first is based on the high affinity of some proteins to methylated DNA, such as the complex of methyl-binding domain protein 2 (MBD2) and methyl-binding domain protein 3L1 (MBD3L1). The methyl-CpG binding domain of human MBD2 or MBD3L1 proteins are coupled to beads and the methylated fragments are captured, obtaining a single enriched population or distinct subpopulations based on the degree of methylation, varying NaCl concentration in the elution buffer. The second method, called methylated DNA immunoprecipitation (MeDIP) (Wilson *et al.*, 2006; Keshet *et al.*, 2006), consists in immunoprecipitate the methylated fraction of the genomic DNA sample with a monoclonal antibody against methylated cytosine.

Afterwards, in both cases the enriched fraction from sheared genomic DNA is amplified and hybridizes against the total fraction of genomic DNA on a microarray. Although the resolution of enrichment-based methods is a few hundred base pairs at best, there is less sequence bias compared with the other two methods, based on bisulphite conversion and methylation sensitive restriction. In this context, it is worth noting that methylation of neighbouring CpG sites has been shown to be frequently correlated up to 1000 bp (Eckhardt *et al.*, 2006) suggesting that, although desirable, it is not always necessary to obtain methylation values at single base pair resolution.

Overall, the application of microarray technology has provided an important platform to profile DNA methylation. However, the use of microarrays as the platform of choice for methylome analysis is decrease with the advancement of sequencing-based approaches, that have developed in recent years.

1.3.2 3. Sequencing-based methylation profiling

Throughout the 1990s, DNA sequencing underwent a major transformation because of numerous genome projects, including the Human Genome Project. This resulted in much improved sequencing technology, aiding the development of sequencing-based approaches for genome-wide methylation profiling. To date high-throughput DNA sequencing combined to sodium bisulfite conversion could be applied for determining DNA methylation states at individual cytosine in all the genome. These methods, generally referred to as *whole-genome bisulfite sequencing* (WGBS) are cost prohibitive for sequencing of large numbers of individual samples. As a result, *Reduced Representation Bisulfite Sequencing* (RRBS) was developed (Meissner *et al.*, 2005; Gu *et al.*, 2010), making possible to investigate large numbers of individuals.

RRBS is a DNA wide methylation analysis technique that specifically enriches genomic regions with a high density of potential methylation sites and enables investigation of DNA methylation at single-nucleotide resolution. RRBS encompasses random DNA fragmentation over 3′-CCGG-5′ sequences by MspI. This step is biased in areas with low CpG content, although it grants single-base resolution. After end repair, A-tailing, and ligation with methylated adapters, the DNA fragments from 50 to 250 bp are purified by gel electrophoresis. These are representative of most promoter length and CpG islands. Next, the modified DNA is amplified and sequenced. This reduced representation implies a lower number of reads necessary to yield an accurate sequencing, involving less cost and time compared with WGBS. Various companies have launched platforms based on RRBS methodology with improved versions (Methyl-MiniSeq/Methyl-MidiSeq) able to detect 3-4 to 8-9 millions of unique CpG sites, capturing more

than 85% of all CpG islands and more than 80% of gene promoters. Several factors must be considered with this approach. First, the choice of a restriction enzyme to fractionate the DNA will bias the portion of the genome that is represented. A second consideration is the process of mapping reads of bisulfite converted DNA to the genome. Compared to other sequencing methods, RRBS provides an efficient way to generate absolute quantification of methylation of more than 1 million CpG sites at single base pair resolution.

2. OBJECTIVES

PAX3-FOXO1 positive and negative RMS show a different and characteristic gene expression signature, but the mechanisms involved in such a differential regulation are poorly understood. We performed genome-wide methylation analysis of *PAX3-FOXO1* positive and *PAX3-FOXO1* negative RMS tumors to determine whether epigenetic phenomena may explain the differences between these RMS subgroups and to investigate how epigenetic mechanisms contribute to their different biological behavior. Furthermore, we explored the different methylation profiling between metastatic and non-metastatic RMS tumors, in order to understand whether epigenetic changes affect the expression of genes associated to the severity of tumor. The results of this study will help to identify genes and signaling pathways relevant for RMS tumors classification, patients' risk stratification and targeted therapy.

3. MATERIALS AND METHODS

3.1. Patients

A total of 73 tumor biopsies from RMS patients enrolled in the pediatric sarcoma protocols RMS96, RMS4-99 and RMS2005 of the Italian Association of Pediatric Haematology and Oncology (AIEOP) were included after obtaining institutional review board approval. Diagnosis was reviewed by the AIEOP central panel of pathologists in all of the cases and confirmed by RT-PCR using primers for *PAX3/7-FOXO1* fusion gene and MyoD1 transcript. Overall, the cohort of patients can be subdivided into 23 *PAX3-FOXO1* positive ARMS, 7 *PAX7-FOXO1* positive ARMS, 9 *PAX3/7-FOXO1* negative ARMS and 37 ERMS. The clinicopathological characteristics of RMS cases are reported in (**Table 4**).

Sample Name	Gene fusion status	Hystology	Gender	Size	IRS groups	DNA methylation array	RRBS seq	qRT-PCR
ERMS 1	PFN	Embryonal	M	<=5 cm	I			**X**
ARMS 1	PFN	Alveolar	F	>5 cm	III a	**X**		**X**
ARMS 2	P3F	Alveolar	F	<=5 cm	IV		**X**	**X**
ERMS 2	PFN	Embryonal	F	<=5 cm	III a			
ERMS 3	PFN	Embryonal	F	>5 cm	III a		**X**	**X**
ARMS 3	P3F	Alveolar	F	>5 cm	IV	**X**		**X**
ARMS 4	P3F	Alveolar	M	>5 cm	IV	**X**		**X**
ARMS 5	P3F	Alveolar	F	n.e.	IV			**X**
ARMS 6	P3F	Alveolar	F	>5 cm	III a		**X**	**X**
ERMS 4	PFN	Embryonal	M	>5 cm	II a			**X**
ERMS 5	PFN	Embryonal	F	>5 cm	III a			**X**
ARMS 7	P7F	Alveolar	F	>5 cm	IV			**X**
ERMS 6	PFN	Embryonal	M	>5 cm	III a			**X**
ERMS 7	PFN	Embryonal	F	<=5 cm	III a		**X**	**X**
ERMS 8	PFN	Embryonal	M	>5 cm	III b	**X**	**X**	**X**
ARMS 8	P3F	Alveolar	F	n.e.	III a			**X**
ARMS 9	P3F	Alveolar	M	>5 cm	III a			**X**
ARMS 10	PFN	Alveolar	M	>5 cm	IV	**X**		**X**
ERMS 9	PFN	Embryonal	F	>5 cm	III a	**X**		**X**
ARMS 11	P3F	Alveolar	F	n.e.	IV			**X**
ERMS 10	PFN	Embryonal	M	>5 cm	IV		**X**	**X**
ERMS 11	PFN	Embryonal	M	<=5 cm	III a			**X**
ERMS 12	PFN	Spindle cells/Leiomiomatous	F	>5 cm	III a			**X**
ARMS 12	PFN	Alveolar	M	<=5 cm	III b			**X**
ARMS 13	P7F	Alveolar	M	<=5 cm	III a			**X**
ARMS 14	P7F	Alveolar	F	>5 cm	IV			**X**
ARMS 15	P7F	Alveolar	M	>5 cm	III a			**X**
ERMS 13	PFN	Embryonal	M	>5 cm	III a			**X**
ARMS 16	P3F	Alveolar	F	<=5 cm	IV	**X**		**X**
ARMS 17	P3F	Alveolar	M	>5 cm	IV	**X**		**X**
ARMS 18	PFN	Alveolar	M	>5 cm	III		**X**	**X**
ERMS 14	PFN	Botryoid	M	<=5 cm	II a			**X**
ARMS 19	PFN	Solid Alveolar	M	>5 cm	III a		**X**	**X**

Sample Name	Gene fusion status	Hystology	Gender	Size	IRS groups	DNA methylation array	RRBS seq	qRT-PCR
ARMS 20	PFN	Alveolar	M	>5 cm	III b			**X**
ERMS 15	PFN	Botryoid	F	>5 cm	III a			**X**
ERMS 16	PFN	Embryonal	F	>5 cm	III a			**X**
ARMS 21	P3F	Alveolar	F	<=5 cm	IV	**X**		**X**
ERMS 17	PFN	Embryonal	F	>5 cm	III a			**X**
ERMS 18	PFN	Embryonal	M	<=5 cm	III a			**X**
ERMS 19	PFN	Embryonal	M	>5 cm	III a			**X**
ARMS 22	P7F	Alveolar	F	>5 cm	III a			**X**
ARMS 23	P3F	Alveolar	F	n.e.	IV			**X**
ARMS 24	PFN	Alveolar	F	>5 cm	III a			**X**
ERMS 20	PFN	Embryonal	M	<=5 cm	III b	**X**		**X**
ARMS 25	P7F	Alveolar	F	>5 cm	III a			**X**
ARMS 26	P3F	Alveolar	M	>5 cm	III a	**X**	**X**	**X**
ERMS 21	PFN	Embryonal	F	<=5 cm	III		**X**	**X**
ARMS 27	PFN	Alveolar	F	<=5 cm	II a	**X**		**X**
ARMS 28	P3F	Alveolar	M	>5 cm	III		**X**	**X**
ERMS 22	PFN	Embryonal	M	>5 cm	III a		**X**	**X**
ARMS 29	P3F	Alveolar	M	>5 cm	IV			**X**
ARMS 30	P3F	Alveolar	F	<=5 cm	IV			**X**
ERMS 23	PFN	Embryonal	M	<=5 cm	III			**X**
ERMS 24	PFN	Embryonal	M	>5 cm	IV	**X**		**X**
ERMS 25	PFN	Embryonal	F	<=5 cm	III a		**X**	**X**
ERMS 26	PFN	Embryonal	M	<=5 cm	III b			**X**
ERMS 27	PFN	Embryonal	M	>5 cm	III a			**X**
ERMS 28	PFN	Embryonal	M	>5 cm	IV			**X**
ARMS 31	P3F	Alveolar	F	<=5 cm	III a		**X**	**X**
ERMS 29	PFN	Embryonal	M	>5 cm	III a			**X**
ARMS 32	P3F	Alveolar	M	>5 cm	III a			**X**
ARMS 33	P3F	Alveolar	M	>5 cm	IV			**X**
ARMS 34	P3F	Alveolar	F	>5 cm	IV			**X**
ARMS 35	PFN	Alveolar	F	>5 cm	III a			**X**
ARMS 36	P3F	Alveolar	F	>5 cm	IV	**X**		**X**
ERMS 30	PFN	Embryonal	M	>5 cm	IV	**X**		**X**
ERMS 31	PFN	Embryonal	F	<=5 cm	IV			**X**
ARMS 37	P3F	Alveolar	F	unknow	IV			**X**
ARMS 38	P3F	Alveolar	M	<=5 cm	IV			**X**
ARMS 39	P7F	Alveolar	M	>5 cm	III a			**X**
ERMS 32	PFN	Embryonal	M	>5 cm	III a			**X**
ERMS 33	PFN	Embryonal	M	<=5 cm	III a			**X**
ARMS 40	PFN	Alveolar	F	<=5 cm	III a	**X**		**X**

Table 4. Main clinical characteristics of RMS patients. Abbreviations: P3F: *PAX3-FOXO1* positive RMS; P7F: *PAX7-FOXO1* positive RMS; PFN: *PAX3/7-FOXO1* negative RMS; M: male; F: female; n.e.: not evaluable.

3.2. Cell cultures

The human RMS cell lines, RH30 and RD, were purchased from ATCC (Manassas, VA), whereas RH4 and RH28 cells were a gift of Dr P.J. Houghton (St Jude Children's Hospital, Memphis, TN). SMS-CTR, RH36 and CCA were obtained from Dr M. Tsokos (NCI, Bethesda, MD). HEK-293T cells were a generous gift of Dr. S. Indraccolo (Istituto

Oncologico Veneto, IOV, Padova, Italy). RMS and HEK-293T cells were all grown in Dulbecco's Modified Eagle's Medium supplemented with 10% heat-inactivated fetal calf serum (FCS) (Gibco, Life Technologies Co., Carlsbad, CA, USA), 2 mmol/l glutamine, 100U/ml penicillin and 100 µg/ml streptomycin (SIGMA-Aldrich Co., St. Louis, MO, USA), under standard tissue-culture conditions. The features of cell lines are summarize in **Table 5**.

Cell lines	Karyotype/Gene fusion status	Hystology	Origins
RH4	t(2;13)(p25;q14); TP53 mutation	Alveolar	Lung metastasis 7-year-old female
RH30	t(2;13)(p25;q14); TP53 mutation; amplification of 12q13-15; region including CDK4	Alveolar	Bone marrow metastasis 16-year-old male
RH28	t(2;13)(p25;q14); near tetraploid	Alveolar	Axillary metastasis 17-year-old male
RH36	Unknow	Embryonal	Paratesticular relapse 15-year-old male
RD	51-hyperdiploid; MYC amplification; Q61H mutation of NRAS; TP53 mutation	Embryonal	Pelvic mass 7-years-old female
CCA	Multiple chromosomal rearrangements; Q61L mutation of KRAS	Embryonal	Vescical mass 8-year-old male
SMS-CTR	Hypertriploid	Embryonal	Pelvic mass 1-years-old male

Table 5. Human RMS cell lines.

3.3. DNA and RNA extraction

High-molecular-weight DNA and total RNA were extracted from tumor biopsies or cell lines using QIAamp DNA Mini Kit (Qiagen Co., Hilden Germany) and TRIzol reagent (Invitrogen, Life Technologies Co., Carlsbad, CA, USA), respectively, according to the manufacturer's protocols.

3.4. Genome-wide DNA methylation profile

Four µg of genomic DNA were fragmented by sonication, purified using Mini-Elute coloumns (Qiagen Co., Hilden Germany) and the amount of double-strand DNA (dsDNA) was measured using Qubit instrument (Invitrogen, Life Technologies Co., Carlsbad, CA, USA). The success of fragmentation was evaluated using the Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). The MethylMiner Methylated DNA enrichment kit (Invitrogen, Life Technologies Co., Carlsbad, CA, USA) was used to enrich the fraction of methylated dsDNA, starting from 2 µg of fragmented whole genomic DNA. Ten ng of methylated dsDNA for each sample was amplified using Whole Genome Amplification (WGA, Sigma-Aldrich Co., St. Louis, MO, USA). Genomic DNA for each

samples was used as control. DNA methylation profiling was carried out in RMS tumour samples using the *Human DNA Methylation Microarray* (Agilent Technologies, Santa Clara, CA, USA), consisting of about 244,000 (60-mer) probes designed to interrogate about 27,000 known CpG islands. The control genomic DNA and methylated dsDNA were labeled with Cy3 and Cy5 dye using Agilent Genomic DNA labeling kit PLUS (Agilent Technologies, Santa Clara, CA, USA) and competitively hybridized to Human DNA Methylation microarrays platforms (GEO ID: GPL10878). The hybridization was carried out at 67°C for 40 hours in a hybridization oven rotator (Agilent Technologies, Santa Clara, CA, USA). The arrays were washed with Agilent ChiP-on-chip wash buffers as suggest by the supplier. Slides were scanned on an Agilent microarray scanner (model G2565CA), and Agilent Feature Extraction software version 10.7.3.1 was used for image analysis. Raw data are available on the GEO website using accession number GSE67201.

3.5. Statistical analysis of DNA methylation data

Intra-array normalization of methylation levels was performed with linear and loess normalization. While inter-array normalization was performed with quantile normalization (Bolstad *et al.*, 2003; Risso *et al.*, 2009) in order to correct possible experimental distortions. The normalization function was applied to the methylation data of all the experiments. With regard to methylation expression data, Feature Extraction Software (Agilent Technologies, Santa Clara, CA, USA) provided spot quality measures in order to evaluate the quality and the reliability of the hybridization data. In particular, flag "glsFound" and “rlsFound”(set to 1 if the spot had an intensity value that was significantly different from the local background or to 0 in any other cases) was used to filter out unreliable probes: flag equal to 0 was to be noted as "not available (NA)." In order to make more robust and unbiased statistical analyses, probes with a high proportion of NA values were removed from the dataset. Twenty-five percent of NA was used as the threshold in the filtering process, obtaining a total of 90’591 available probes. To identify the differentially methylated regions (DMRs) we used iChip R Bioconductor Package (http://www.bioconductor.org/packages/release/bioc/html/iChip.html). Specifically we directly compared two group of samples (class 1 *versus* class 2). The iChipZ function was run with β=l, following the iChip reference manual instruction for low-resolution arrays. Enriched regions were called with a FDR ≤ 0.2.

3.6. Reduced-Representation Bisulfite Sequencing

Reduced-Representation Bisulfite Sequencing (RRBS) library generation and sequencing were performed at the BGI technology centre (BGI HongKong, Shenzhen, China; www.bgi-international.com). The reduced representation library is enriched for CpG islands and is predicted to include 84% of the CpG islands in the human genome and about 3.4 million of unique CpG sites (Meissner *et al.*, 2008; Smith *et al.*, 2009). Briefly, genomic DNA from tumour biopsies, which passed the quality tests, was digested by methylation–insensitive restriction enzyme, MspI (New England Biolabs, Ipswich, MA). End repair was necessary to fill the 3' terminal of the ends of the strands. The next step was adding of an extra adenosine to both strands (A-tailing step). In the subsequent step methylated sequence adapters, with 3′-T overhang, were ligated to the DNA fragments. The methylated adapter oligonucleotides had all cytosine replaced with 5mC, in order to prevent the deamination of these cytosine in the bisulfite conversion reaction. For reduce representation, DNA fragments of 40-220 bp were then selected to be purified. The different sizes of the fragments were separated using gel electrophoresis and purified using gel excising. The DNA from the excised gel pieces was recovered with the QIAGEN Gel Extraction Purification Kit (Qiagen Co., Hilden Germany), followed by bisulfite treatment using ZYMO EZ DNA Methylation-Gold kit (Zymo Research, Orange, CA, USA). The bisulfite-converted library was then amplified using PCR with primers that were complementary to the sequence adapters. Before sequencing, the PCR product must be free of unused reaction reagents such as unincorporated dNTPs or salts. Thus, a step for PCR purification was required. The fragments were then sequenced on Illumina sequencer, 50 base single-end reads are most commonly performed. FASTQ sequence files were obtained containing sequenced reads for each sample. Finally, sequencing data were mapped to reference genome and only uniquely mapped reads were used for bioinformatics analysis.

3.7. Trichostatin A and 5-aza-2'-deoxycytidine treatments

Cell lines were seeded in Petri dishes at a concentration of $0.25*10^6$ cells/mL and the following day treated with different concentrations (100 nM ÷ 2 μM) of 5-aza-2'-deoxycytidine (Selleck Chemicals, Houston; TX, USA). The same treatment was repeated after 48 hours from the seeding, while the collection of the cells was performed at a distance of 72 hours. Treatment with 200 ng/ml of trichostatin A (Selleck Chemicals, Houston; TX,

USA) was performed after 48 hours from the seeding and cells were collected the following day after 16 hours. Cells were differently processed for RNA or DNA extraction.

3.8. Reverse transcription and qRT-PCR

One µg of total RNA was reverse transcribed using SuperScript II reverse transcriptase (Invitrogen, Life Technologies Co., Carlsbad, CA, USA) and random hexamers. Briefly, samples were denatured at 75°C for 3 min and cooled at 4°C for 5 min. Subsequently, 10 µl of reaction mixture were added, containing 1X of buffer, 2 mM of each dNTP (GE Healthcare Life Sciences, Uppsala, Sweden), 250 nM random hexamers and 40 U of RNAse inhibitor (Roche Applied Science, Penzberg, Germany). After 10 min at 20°C, 10 mM DDT and 200 U of SuperScript II reverse transcriptase were added. Retrotranscription was performed as follows: 60 min at 37°C, 5 min at 99°C, 5 min at 4°C. Quantitative reverse polymerase chain reaction (qRT-PCR) was performed in the VIIa 7 thermal cycler (Applied Biosystems, Life Technologies, Foster City, CA) using standard amplification conditions. Amplification reaction was set up in triplicate in a final volume of 10 µl containing 8 ng of cDNA, 1X of SYBR Green PCR mastermix (Applied Biosystems, Life Technologies, Foster City, CA) and variables concentrations (100 nM ÷ 500 nM) of forward and reverse primers (Invitrogen, Life Technologies Co., Carlsbad, CA, USA). Gene-specific primers were designed using Primer Express Software v3.0 (Applied Biosystems, Life Technologies, Foster City, CA). To evaluate differences in gene expression, we chose a relative quantification method in which the expression of target gene is standardized by the housekeeping glyceraldehyde 3-phosphate dehydrogenase (*GADPH*), as reference gene. The mathematical method presented by Pfaffl (Pfaffl, 2001), which calculates the efficiency of each PCR using a standard curve, was applied. To calculate the relative expression ratio we used REST software tool (Pfaffl *et al.*, 2002), which permits the comparison of more than one target genes with a reference gene in two experimental groups, exploiting pair wise fixed reallocation randomization test. For *GADD45G* and *NELL1* relative gene expression we used the $\Delta\Delta C_T$ method implemented in the software of VIIa 7 thermal cycler (Applied Biosystems, Life Technologies, Foster City, CA), using fetal human skeletal muscle (Stratagene, La Jolla, CA, USA) as calibrator.

3.9. Sodium bisulfite treatment of DNA

One µg of genomic DNA, isolated from RMS biopsies or cell lines, was subjected to conversion with sodium bisulfite using EZ DNA Methylation-Gold ™ kit (Zymo Research,

Orange, CA, USA), following the manufacturer's instructions. This treatment allows the conversion of unmethylated cytosine residues to uracil, whereas methylated cytosine remain unchanged.

3.10. Bisulfite sequencing PCR

One hundred ng of bisulfite-coverted DNA of RMS tumors or cell lines was used as template for the amplification of promotorial regions of *HOXC11*, *PCDHA4*, *NELL1* and *GADD45G* genes. First, polymerase chain reaction (PCR) was performed in a total volume of 50 µl containing with 1.25 U of AmpliTaq Gold DNA polymerase (Applied Biosystems, Life Technologies Co., Carlsbad, CA, USA), 0.8 mM of deoxynucleoside triphosphate (GE Healthcare Life Sciences, Uppsala, Sweden), 1X of GeneAmp PCR Buffer II (Applied Biosystems, Life Technologies Co., Carlsbad, CA, USA), 3mM MgCl2 (Applied Biosystems, Life Technologies Co., Carlsbad, CA, USA) and 200 nM of forward and reverse primer (Invitrogen, Life Technologies Co., Carlsbad, CA, USA). Methylation-independent primers were designed with the free online tool MethPrimer (http: /itasa.ucsf..edu/~urolab/methprimer; Li and Dahiya, 2002). Amplification consisted in an initial denaturation of 7 min at 95°C, followed by 45 cycles of 30 sec at 94°C, 30 sec at 60, 1 min at 72°C and 10 sec at 72°C. The PCR products were purified using QIAquick PCR purification kit (Qiagen Co., Hilden Germany) and subcloned into pSC-A-amp/kan vector, supplied with the StrataClone PCR Cloning Kit (Agilent Technologies, Santa Clara, CA, USA). Competent cells were transformed with ligation reaction product and grown in LB (Luria-Bertani) agar plates supplemented with 40 µg/ml of X-Gal (Promega Co., Madison, WI, USA) and 50 µg/ml of ampicillin for 16 hours at 37 °C. Blue-white screening permitted identification of recombinant bacteria. Selected clones were evaluated by colony PCR performed with 0.5 U of TaqDNA Polymerase (Roche, Basel, Switzerland), 1X PCR Reaction Buffer 10X (Roche, Basel, Switzerland), 40 nM of deoxynucleoside triphosphate (GE Healthcare Life Sciences, Uppsala, Sweden) and 100 nM of the universal primers M13R and T7 (Invitrogen, Life Technologies Co., Carlsbad, CA, USA). Then, by agarose electrophoresis PCR products were checked for insert cloning presence. PCR products corresponding to positive clones were purified by QIAquick PCR purification kit (Qiagen Co., Hilden Germany) and then sequenced by 3500 Dx Genetic Analyzer sequencer (Applied Biosystems, Life Technologies Co., Carlsbad, CA, USA) using BigDye® Terminator v3.1 Cycle

Sequencing Kit (Applied Biosystems, Life Technologies Co., Carlsbad, CA, USA) according to manufacturer's protocols.

3.11.Quantitative methylation-specific PCR

Quantitative methylation-specific polymerase chain reaction (qMS-PCR) was performed to assess the methylation status of *PCDHA4* promotorial region. This method allows the analysis of the methylation status of target genes by real-time PCR making use of primer pairs able to discriminate between methylated and unmethylated sequences. Primers were designed using the free software MethPrimer (http: /itasa.ucsf..edu/~urolab/methprimer; Li and Dahiya, 2002). The amplification reaction was set up in triplicate in a final volume of 10 µl containing 8 ng of bisulfite-converted DNA, 1X SYBR green PCR mastermix (Applied Biosystems, Foster City, CA) and variable concentrations of forward and reverse primer (Invitrogen, Life Technologies Co., Carlsbad, CA, USA). Amplifications were carried out in VIIa7 thermal cycler (Applied Biosystems, Life Technologies, Foster City, CA), using standard amplification conditions. Primers were designed also to amplified the internal reference gene glyceraldehyde 3-phosphate dehydrogenase (*GADPH*). These were located in areas without CpG dinucleotides, thus amplifying the modified *GAPDH* gene independently of the methylation status. Four µg of genomic DNA of 293T cell line were treated with 20 U of SssI methyltransferase (New England Biolabs Inc., Beverly, Mass, USA) to generate completely methylated DNA, whereas 10 ng were subjected to the whole-genome amplification using REPLI-g Mini Kit (QIAGEN, Hilden, D) to obtain totally demethylated DNA. Serial dilutions (8 ng to 0.125 ng) of bisulfite-converted SssI-treated DNA were used to construct a calibration curve for each plate. To determine the relative levels of methylated promoter DNA in each sample, the values of the gene of interest were compared with the values of the internal reference gene to obtain the ratio GENE/GAPDH. The amount of methylated DNA was then calculated by dividing the GENE/GAPDH ratio of each sample by the GENE/GAPDH ratio of SssI-treated 293T DNA and multiplying by 100 (Kloten *et al.*, 2013).

3.12. Statistical analysis

Correlation analysis of *PCDHA4* expression and methylation levels with histology, *PAX3/FOXO1* fusion gene status, sex, size of tumors and IRS group, were performed by Welch Two t-test and Wilcoxon rank sum test, using R statistical software. The mRNA expression levels of *GADD45G* and *NELL1* were correlated with histology, *PAX3/7-FOXO1*

fusion gene status and IRS group, either by Student's t-test or Wilcox-Mann Whitney test. Receiver-operator characteristic curves (ROCs) were also calculated to understand if *GADD45G* and *NELL1* expression levels were good predictors of *PAX3/7-FOXO1* fusion gene status. Survival analysis was performed according to Kaplan–Meier method and differences were calculated by applying log-rank test. Overall survival (OS) was calculated from the date of diagnosis to the date of death for any cause or the last follow-up, whereas progression-free survival (PFS) was calculated from the date of diagnosis to the date of the first event (tumor progression or relapse) or the last follow-up. All these analysis were performed using PRISM 6 software (GraphPad Software, La Jolla, CA, USA)

4. RESULTS

4.1. DNA methylation profiling of RMS tumors biopsies by microarray analysis

To study the global methylation status of RMS DNA samples we used Agilent Human DNA Methylation platform that is a collection of 244 k probes design to interrogate about 27,000 known CpG islands. Methylome analysis were carried out on the genomic DNA of RMS biopsies chosen to represent the major subtypes of RMS. We analyzed 16 RMS samples including 7 *PAX3-FOXO1* positive ARMS, 4 *PAX3-FOXO1* negative ARMS and 5 ERMS (**Table 4**).

4.1.1. Discovery of novel methylated target genes: *PAX3-FOXO1(+)* RMS *vs PAX3-FOXO1(-)* RMS

4.1.1.1. Microarray data analysis

Initially, we compared the methylation profile obtained by microarray experiments among *PAX3-FOXO1* positive and negative RMS biopsies using the iChip R Bioconductor Package. The analysis (*false discovery rate* (FDR) <0.2) revealed 216 *differentially methylated regions* (DMRs) able to discriminate between the two groups of samples. Therefore, we mapped DMRs to genome using UCSC Genome browser. We observed that only a small number of DMRs aligned with gene promoters while the others were localized on CpG islands distal to known genes or inside the genes body. While the relationship between DNA methylation occurring in the promoter region and gene expression is largely established, the link between methylation in the gene body and the expression of the gene is still a controversial issue. Then, we initially focused our attention on DMRs associated with the putative promoters. We established as promoter a region in proximity of the 5' of the gene that range from -2000 bp to +1000 bp, where the position +1 corresponds to the *transcriptional start site* (TSS). In this way we found that the DMRs associated with regulatory regions matched to 40 genes (**Table 6**).

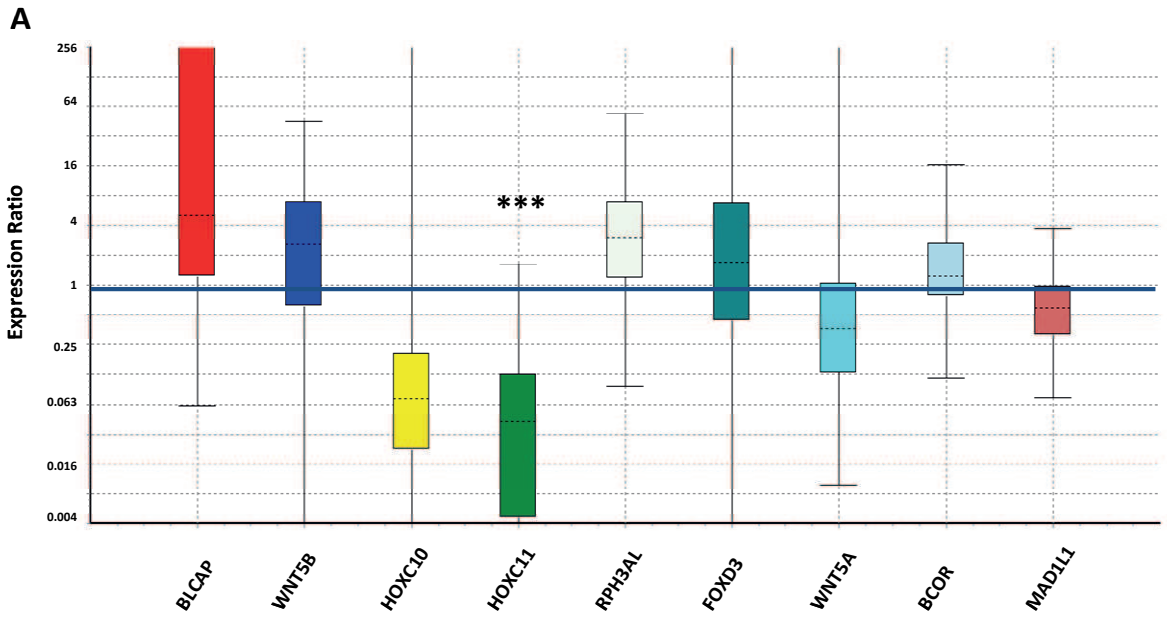
Gene Symbol	Official full name	qRT-PCR
RNF223	ring finger protein 223	
FOXD3	forkhead box D3	**X**
LHX4	LIM homeobox 4	
MYT1L	myelin transcription factor 1 like	
TSSC1	tumor suppressing subtransferable candidate 1	
PDCD1	programmed cell death 1	
ANKRD18DP	ankyrin repeat domain 18D, pseudogene	
FAM157A	family with sequence similarity 157 member A	
MAD1L1	MAD1 mitotic arrest deficient-like 1	**X**
PTPRN2	protein tyrosine phosphatase, receptor type N2	
DIP2C	disco interacting protein 2 homolog C	
ADARB2	adenosine deaminase, RNA-specific, B2	
HMX2	H6 family homeobox 2	
HMX3	H6 homeobox 3	
GPR123	G protein-coupled receptor 123	
DCDC1	doublecortin domain containing 1	
DNAJC24	DnaJ heat shock protein family (Hsp40) member C24	
IMMP1L	inner mitochondrial membrane peptidase subunit 1	
PAX6	paired box 6	
SHANK2	SH3 and multiple ankyrin repeat domains 2	
DHCR7	7-dehydrocholesterol reductase	
FBXL14	F-box and leucine-rich repeat protein 14	
WNT5B	wingless-type MMTV integration site family, member 5B	**X**
HOXC11	homeobox C11	**X**
HOXC10	homeobox C10	**X**
MCF2L	MCF.2 cell line derived transforming sequence like	
LMF1	lipase maturation factor 1	
CPNE7	copine VII	
RPH3AL	rabphilin 3A-like (without C2 domains)	**X**
FAM101B	family with sequence similarity 101, member B	
NFATC1	nuclear factor of activated T cells, cytoplasmic, calcineurin dependent 1	
AP2A1	adaptor related protein complex 2 alpha 1 subunit	
FUZ	fuzzy planar cell polarity protein	
BLCAP	bladder cancer associated protein	**X**
NNAT	neuronatin	
CDH4	cadherin 4	
WNT5A	wingless-type MMTV integration site family member 5A	**X**
PARVB	parvin beta	

Gene Symbol	Official full name	qRT-PCR
BCOR	BCL6 corepressor	**X**

Table 6. Genes associated to DMRs obtained by microarray data comparing *PAX3-FOXO1* positive and *PAX3-FOXO1* negative RMS samples. Last column reports genes used for subsequent validation analysis by qRT-PCR.

4.1.1.2. Evaluation of expression levels of candidate genes

Among the 40 genes associated to DMRs identified comparing *PAX3-FOXO1* positive and negative RMS we focused on genes involved in oncogenetic processes or that are established be tumor suppressor. We selected 9 candidate genes hypermethylated in RMS *PAX3-FOXO1* positive *vs PAX3/FOXO1* negative (**Table 6**). To determine whether the hypermethylation of the selected genes correlates with reduction of expression, we evaluated the transcription level of these genes by qRT-PCR in a total of 25 biopsies (**Table 4**; **Figure 4 A**), of which 16 out of these were used for microarray hybridization experiments We assessed the expression levels of the 9 genes also in 7 RMS cells lines (**Table 5**; **Figure 4 B**). Unfortunately, the methylation status of 8 of the selected genes did not match with a decrease in the expression level, but it showed an extremely varied trend. We observed only for *HOXC11* a statistically significant anti-correlation between methylation status and gene expression level in biopsy specimens (P<0,001) and a moderate inverse correlation (P=0.489) in cell lines (**Figure 4**).



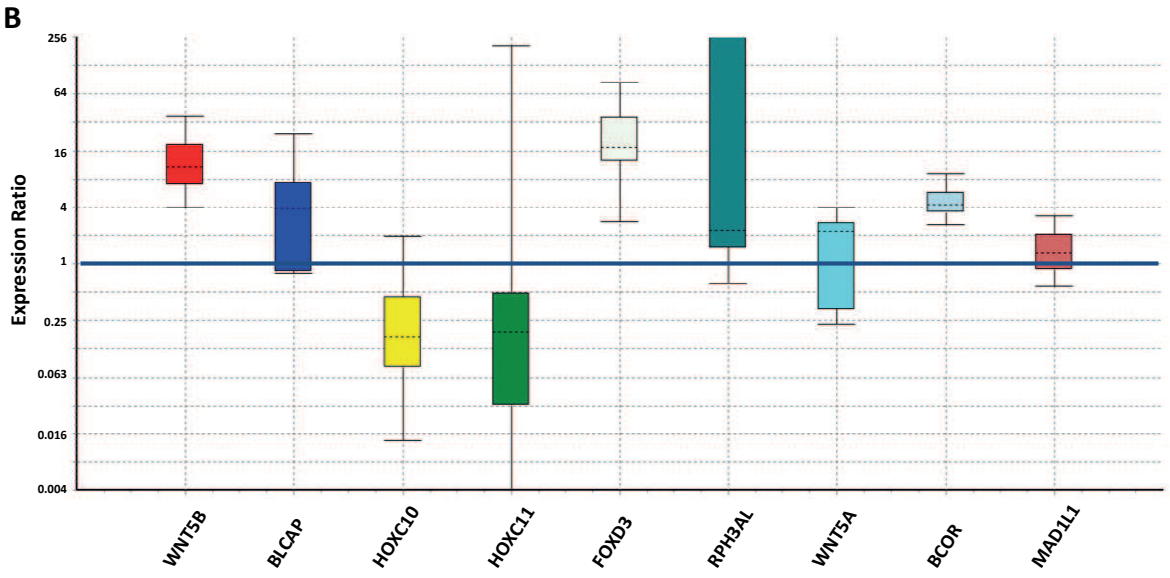


Figure 4. Microarray data validation by qRT-PCR analysis. (**A**) Relative expression levels of 9 selected genes in 10 *PAX3-FOXO1* positive tumor samples compared to 15 *PAX3-FOXO1* negative ones. (**B**) Relative expression levels of the 9 selected genes in 3 *PAX3-FOXO1* positive cell lines compared to 4 *PAX3-FOXO1* negative ones. Distribution of data is represented by box plot analysis. Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was used as housekeeping gene for normalization. Expression ratio >1: high expression in *PAX3-FOXO1* positive RMS samples than *PAX3-FOXO1* negative ones; expression ratio <1: low expression in *PAX3-FOXO1* positive RMS samples than *PAX3-FOXO1* negative ones; expression ratio =1: equal expression in *PAX3-FOXO1* positive RMS samples and *PAX3-FOXO1* negative ones. **P*<0.05; ***P*<0.01; ****P*<0.001.The analysis was performed with REST software.

4.1.1.3. *HOXC11*: pharmacological treatment and bisulfite-sequencing

To understand if the gene expression level of *HOXC11* can be modulated by methylation in the promoter region we performed a set of experiments with demethylating agents. We used 5-aza-2'-deoxycytidine that is a well known pyrimidine analogue that inhibits DNA methyltransferase, impairing DNA methylation. We treated RH4 and RD cell lines, that represented *PAX3-FOXO1* positive and *PAX3-FOXO1* negative cell lines respectively, with increasing doses of 5-aza-2'-deoxycytidine and assessed the expression level of *HOXC11* after 72 hours from the treatment. We observed a strong restoration of *HOXC11* expression in RH4 cells, greater than 100 folds, already with 100 nM of 5-aza-2'-deoxycytidine, than untreated control, while we did not detect any changes in RD cells (**Figure 5**). This result suggested a regulative role of methylation in *HOXC11* expression. To confirm and validate these data we performed a bisulfite Sanger sequencing of the promotorial region of *HOXC11,* ranging from position -161 to +711 and overlapping a CpG island identified with microarray experiments. We bisulfite-converted genomic DNA of RH4 and RD genomic cells and performed a PCR to

amplified the candidate region. Then, we subcloned the PCR product in pSC-A-amp/kan vector and plated on LB agar. Clones were picked out and then sequenced. The results showed that in RH4 cells the 82.87% of the cytosine inside a CpG context were methylated, while were totally unmethylated in RD cells (**Figure 6**). Then, we performed the same experiments of bisulfite sequencing in three RMS patients (ARMS 3, ARMS 26 and ERMS 30; Table 4). We did not observe any difference in the average percentage of methylated CpG dinucleotides between *PAX3-FOXO1* positive and *PAX3-FOXO1* negative tumor specimens (**Figure 6**). Putting together these results and the very low number of DMRs found by the comparison of microarray data between the *PAX3-FOXO1* positive and negative RMS group, we decided to analyzed the microarray data comparing other subgroups of patients.

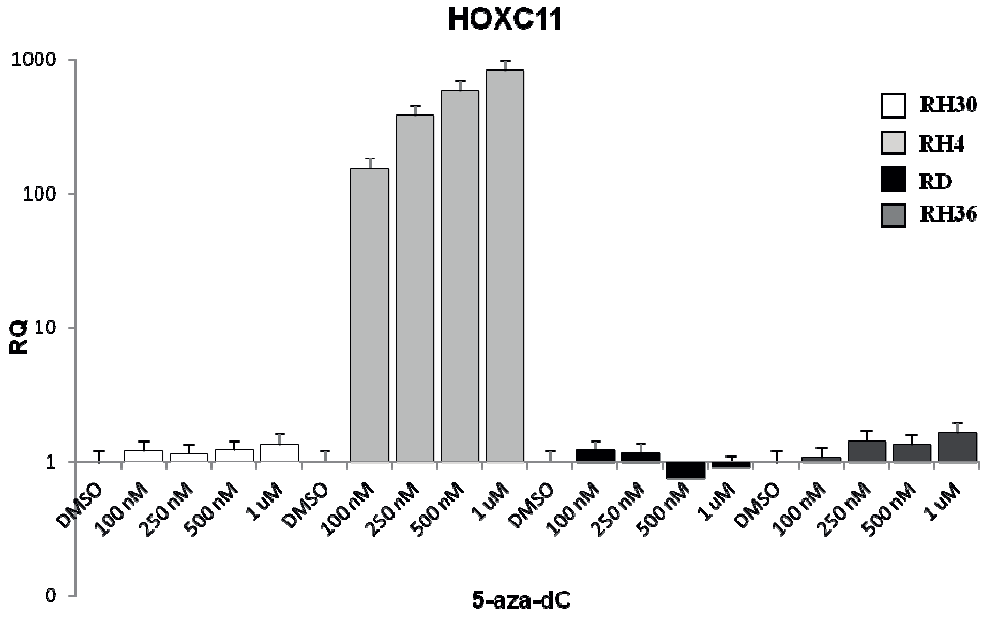


Figure 5. Relative expression of *HOXC11* at 72 hours from treatment with increasing doses of 5-aza-2'-deoxycytidine (100 nM, 250 nM, 500 nM and 1uM) in 4 RMS cell lines. Expression levels were assessed by qRT-PCR. Housekeeping *GAPDH* gene was used as internal control for normalization and DMSO-treated cells as calibrator. RH30, RH4: *PAX3-FOXO1* positive RMS cell lines; RD, RH36: *PAX3-FOXO1* negative RMS cell lines; 5-aza-dC: 5-aza-2'-deoxycytidine, RQ=relative expression ratio.

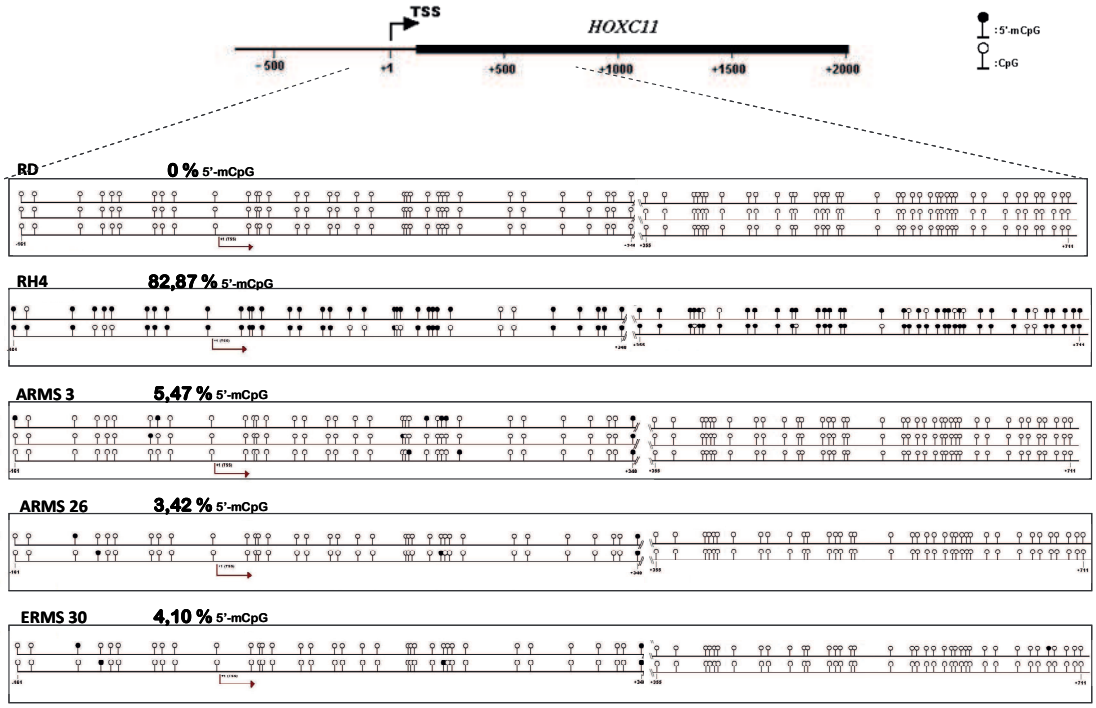


Figure 6. Sanger bisulfite sequencing of *HOXC11* promotorial region in 2 RMS cell lines and 3 biopsies. Sequencing was performed for at least 2 clones obtained by subcloning bisulfite-converted promotorial region. Sequenced region spanning from position -161 to +711, where position +1 corresponds to the *trascriptional start site* (TSS). The % mean value of methylation levels of the sequenced clones is reported next to the sample name. Circles: cytosine within CpG dinucleotides; black circles: methylated cytosine; white circles: unmethylated cytosine; RD: *PAX3-FOXO1* negative RMS cell line; RH4: *PAX3-FOXO1* positive RMS cell line.

4.1.2. Discovery of novel methylated target genes: IRS IV RMS group *vs* IRS I, II and II RMS groups

4.1.2.1. Microarray data analysis

The IRS group system is highly predictive of outcome of RMS (Crist *et al.*, 1995), in particular patients belong to group IRS IV, characterized by metastatic disease, have long term failure-free survival FFS rates of <30% (Breneman *et al.*, 2003; Oberlin *et al.*, 2008). It is not known if the methylation status is different in these subgroups of patients. Then, we used iChip R Bioconductor Package to comparing metastatic (IRS IV) RMS patients *versus* non-metastatic (IRS I, II and III) RMS patients. From this analysis we identified 1394 DMRs (FDR<0.2). Therefore, we mapped DMRs to genome using UCSC Genome browser and we found that only 357 DMRs localize in promoter regions. All the others DMRs mapped in CpG regions within the gene body or distal to known coding sequences (intergenic regions). Interestingly, these results demonstrated that RMS with advanced stage tumors (metastatic)

show an epigenetic pattern peculiar and different with respect to patients with low grade of severity (non–metastatic).

To discovery genes directly or indirectly modulated by DNA methylation, we analyzed the genes associated to DMRs with the functional annotation web tool DAVID that perform a GO-term analysis and identify which functional categories are over-represented. Terms analyzed had an adjusted $P<0.05$. We found a consistent number of genes involved in cell adhesion (48 genes, $P=3.1*10^{-27}$) and in regulation of transcription (65 genes, $P=4.8*10^{-4}$) that are the major significantly enriched functional categories (**Figure 7**). Interestingly, in cell adhesion categories we found many members of protocadherin (*PCDHs*) cluster belonging to the major groups α, β and γ of clustered protocadherins (**Table 7**). Moreover, we observed that all DMRs linked to protocadherins mapped in promoter regions (**Figure 8**). This raised us great interesting since several members of this family are known to be tumor suppressor genes in some cancers and are frequently downregulated by methylation of DNA promoter.

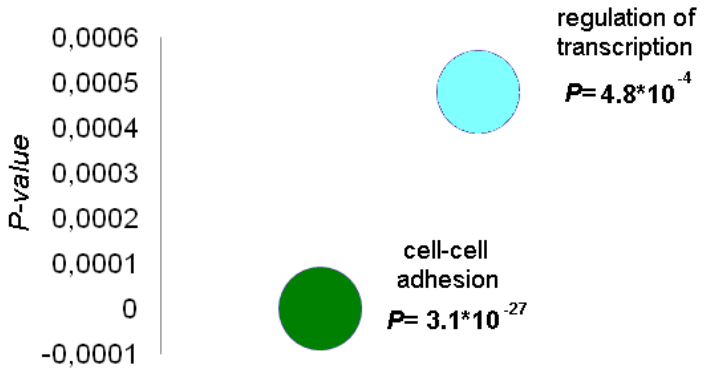


Figure 7. Enriched GO classes. GO-term analysis was performed by DAVID on genes associate to DMRs obtained by microarray data comparison between metastatic and non-metastatic RMS biopsies.

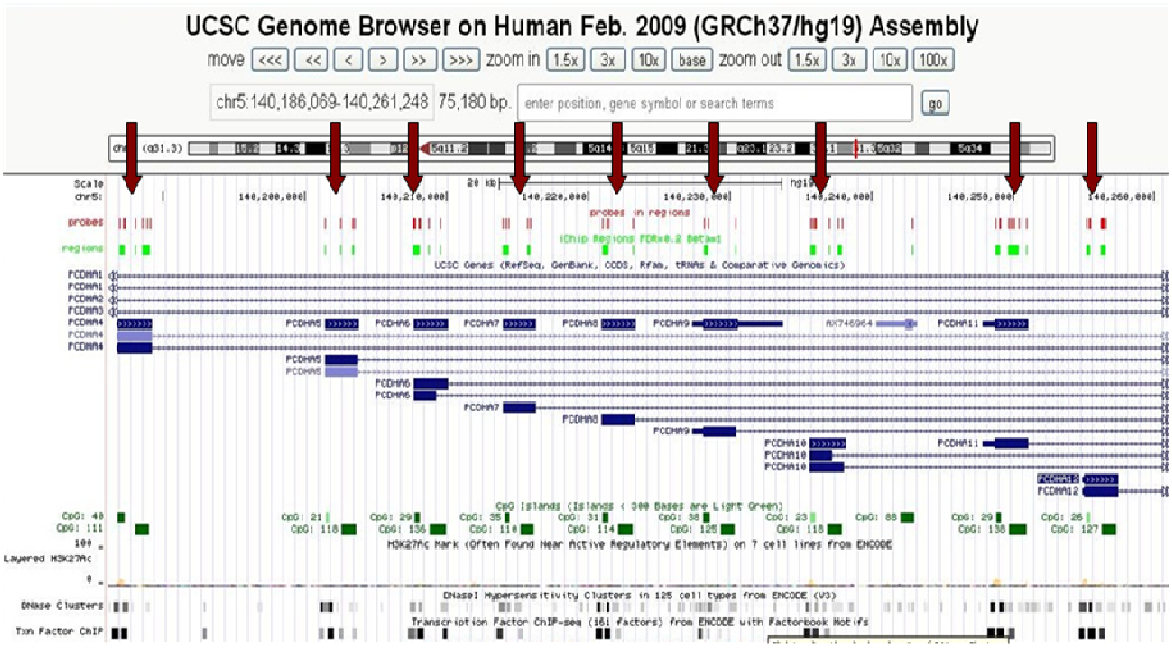


Figure 8. Representation of protocadherins cluster by UCSC genome browser. Microarray probes are represented by red bars. Green bars rapresent the DMRs. Red arrows indicate DMRs, obtained by comparing microarray data of metastatic *vs* non metastatic RMS samples, mapped precisely against regulatory regions of protocadherin genes.

Gene Symbol	Official full name	qRT-PCR
CD9	CD9 molecule	
CNTNAP3B	contactin associated protein-like 3B	
CNTNAP3	contactin associated protein-like 3	
GP1BB	glycoprotein Ib (platelet), beta polypeptide	
GP5	glycoprotein V (platelet)	
PCDHA11	protocadherin alpha 11	**X**
PCDHA12	protocadherin alpha 12	**X**
PCDHA13	protocadherin alpha 13	
PCDHA10	protocadherin alpha 10;	
PCDHAC1	protocadherin alpha subfamily C, 1	
PCDHAC2	protocadherin alpha subfamily C, 2	
PCDHA4	protocadherin alpha 4	**X**
PCDHA1	protocadherin alpha 1	**X**
PCDHA2	protocadherin alpha 2	
PCDHA3	protocadherin alpha 3	
PCDHA5	protocadherin alpha 5	

Gene Symbol	Official full name	qRT-PCR
PCDHA7	protocadherin alpha 7	
PCDHA8	protocadherin alpha 8	**X**
PCDHA6	protocadherin alpha 6	
PCDHA9	protocadherin alpha 9	
PCDHB9	protocadherin beta 9	
PCDHB10	protocadherin beta 10	
PCDHB11	protocadherin beta 11	
PCDHB12	protocadherin beta 12	**X**
PCDHB13	protocadherin beta 13	
PCDHB14	protocadherin beta 14	
PCDHB15	protocadherin beta 15	
PCDHB16	protocadherin beta 16	
PCDHB17	protocadherin beta 17 pseudogene	
PCDHB18	protocadherin beta 18 pseudogene	
PCDHB3	protocadherin beta 3	
PCDHB4	protocadherin beta 4	
PCDHB5	protocadherin beta 5	**X**
PCDHB6	protocadherin beta 6	**X**
PCDHB7	protocadherin beta 7	**X**
PCDHB8	protocadherin beta 8	
PCDHGA1	protocadherin gamma subfamily A, 1	**X**
PCDHGA10	protocadherin gamma subfamily A, 10	
PCDHGA11	protocadherin gamma subfamily A, 11	
PCDHGA2	protocadherin gamma subfamily A, 2	**X**
PCDHGA3	protocadherin gamma subfamily A, 3	**X**
PCDHGA4	protocadherin gamma subfamily A, 4	
PCDHGA5	protocadherin gamma subfamily A, 5	
PCDHGA6	protocadherin gamma subfamily A, 6	
PCDHGA7	protocadherin gamma subfamily A, 7	
PCDHGA8	protocadherin gamma subfamily A, 8	
PCDHGB1	protocadherin gamma subfamily B, 1	
PCDHGB3	protocadherin gamma subfamily B, 3	
PCDHGB4	protocadherin gamma subfamily B, 4	
PCDHGB5	protocadherin gamma subfamily B, 5	X

Gene Symbol	Official full name	qRT-PCR
PCDHGB6	protocadherin gamma subfamily B, 6	
PCDHGB7	protocadherin gamma subfamily B, 7	
PCDHGC5	protocadherin gamma subfamily C, 5	
PCDHGA12	protocadherin gamma subfamily A, 12	
PCDHGC3	protocadherin gamma subfamily C, 3	
PCDHGC4	protocadherin gamma subfamily C, 4	
SCARF2	scavenger receptor class F, member 2	
VWF	von Willebrand factor	

Table 7. Lists of genes associated with the GO-term of cell-cell adhesion. Last column reports genes used for subsequent validation analysis by qRT-PCR.

4.1.2.2. Evaluation of expression levels of candidate genes

To evaluate the correlation among methylation status and expression levels of the protocadherins we selected 15 of them, belonging to the three different subgroups α, β and γ (**Table 7**). We analysed their expression levels by qRT-PCR. We performed this assessment in 7 RMS cell lines and (**Table 5**) in 22 biopsies (**Table 4**). We demonstrated an inverse correlation between methylation and expression for *PCDHA12* (P=0.021) and *PCDHA4* (P<0.001) in the cell lines (**Figure 9 A**) and *PCDHA4* (P=0.030) and *PCDHB7* (P=0.004) in the tumor samples (**Figure 9 B**). Despite, chosen protocadherins shown an hypermethylation in metastatic RMS, a reduced gene expression was not always observed. Indeed, only *PCDHA4* was found to be downregulated at mRNA level, making it a valuable candidate for subsequent *in vitro* studies.

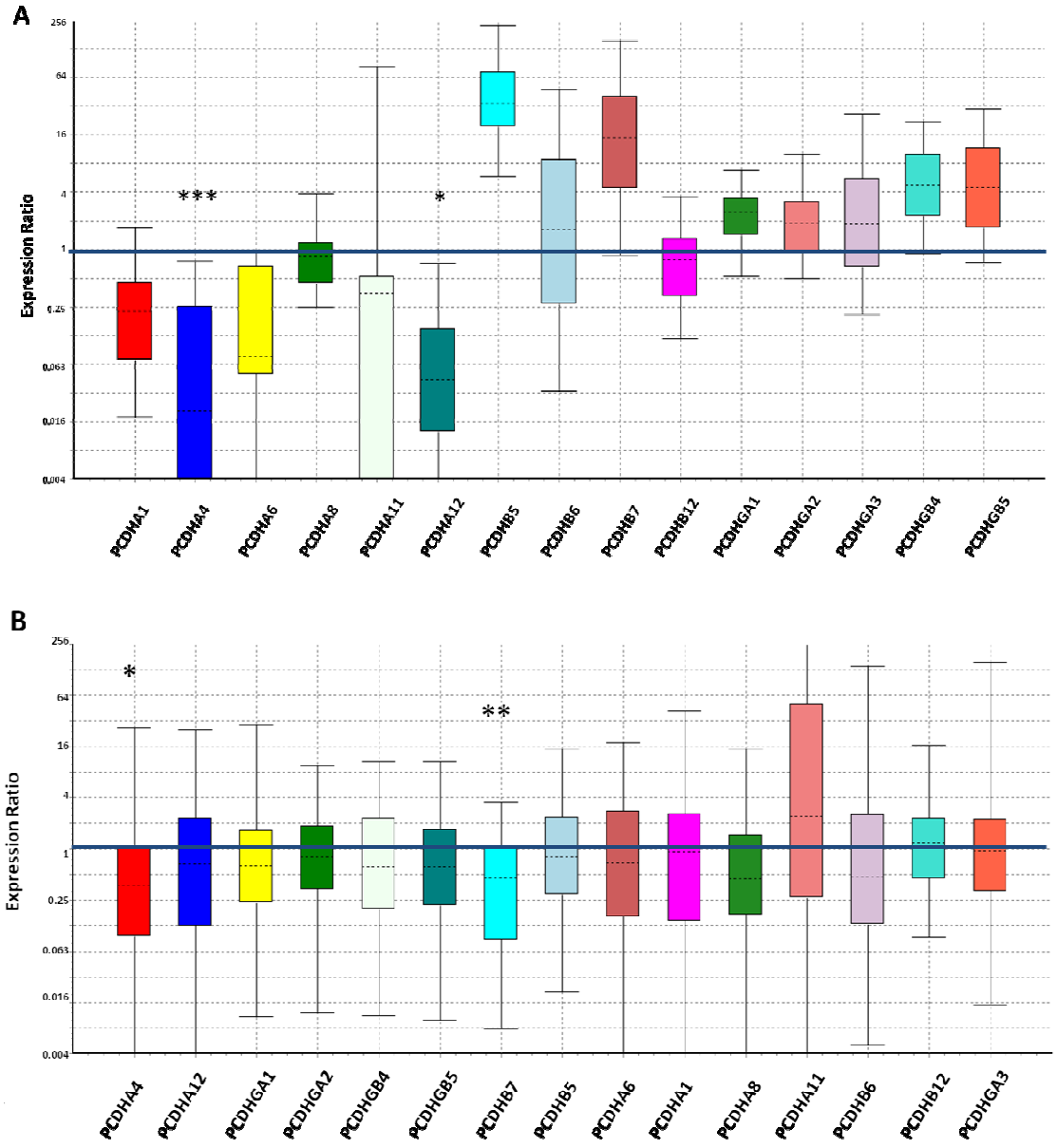


Figure 9. PCDHs genes expression level by qRT-PCR analysis. (**A**) Relative expression levels of 15 PCDHs genes in 3 metastatic cell lines compared to 4 non-metastatic ones. (**B**) Relative expression levels of 15 PCDHs genes in 12 metastatic RMS tumor samples compared to 10 non-metastatic ones. Distribution of data is represented by box plot analysis. Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was used as housekeeping gene for normalization. Expression ratio >1: high expression in metastatic RMS samples than non-metastatic ones; expression ratio <1: low expression in metastatic RMS samples than non-metastatic ones; expression ratio =1: equal expression in metastatic RMS samples and non-metastatic ones. * $P<0.05$; ** $P<0.01$; *** $P<0.001$. The analysis was performed with REST software.

4.1.2.3. PCDHA4 expression is regulate by methylation

To confirm the methylation status of the *PCDHA4* promoter in metastatic RMS samples we performed *in vitro* pharmacological studies with 5-aza-2'-deoxycytidine, trichostatin A (an inhibitor of histone deacetylases) and with the combination of both. We used two *PAX3-*

FOXO1 positive cell lines (RH4 and RH30) and two *PAX3-FOXO1* negative ones (RD and RH36). We assessed *PCDHA4* expression level by qRT-PCR at 72 hours from the treatment with increasing doses of 5-aza-2'-deoxycytidine and as expected we did not observe any change in not metastatic cell lines, whereas we observed a dose dependent restoration of *PCDHA4* expression in RH30 (**Figure 10**). Since there are many literature data that suggest an epigenetic cross-talk between DNA methylation and histone acetylation in the process of gene transcription and aberrant gene silencing in tumours (Vaissierre *et al.*, 2008), we tried to explore if even in our case acetylation mechanisms were involved in *PCDHA4* transcription levels regulation. We treated RMS cell lines for 72 hours with 1uM of 5-aza-2'-deoxycytidine (DAC), for 16 hours with 200 ng/ml of trichostatin A (TSA) and with the combination of both. When we treated with only TSA we could not observe any change whereas when we combined it with DAC we observed a double effect than just treatment with DAC (**Figure 11**). These results confirmed an involvement not just of methylation in *PCDHA4* regulation, but also of histone acetylation as well.

To get further proofs of *PCDHA4* promoter methylation we performed a bisulfite Sanger sequencing in four RMS cell lines. As above we used two *PAX3-FOXO1* positive cell lines (RH4 and RH30) and two *PAX3-FOXO1* negative ones (RD and RH36). We extracted genomic DNA and performed a bisulfite conversion. For primers design, we considered the region overlapping the CpG island proximal to the *PCDHA4*'s TSS, ranging from the position + 94 and +828 and containing the DMR identified comparing metastatic RMS and non-metastatic RMS. We amplified the putative promoter region from bisulfite converted DNA and then subcloned the PCR products in pSC-A-amp/kan vector and plated on LB agar. We performed a Sanger sequencing of at least eight clones obtained by subcloning the amplified *PCDHA4* putative region. The results showed that methylation level was higher in the metastatic RMS cell lines (72.94% of 5'm-CpG in RH4 and 91.42% of 5'm-CpG in RH30; **Figure 12 A, B**) than in non-metastatic ones (41.75% of 5'm-CpG in RH36 and 44.53% of 5'm-CpG in RD; **Figure 12 C, D**). Thus, we confirmed the different methylation levels between metastatic and non-metastatic RMS cell lines.

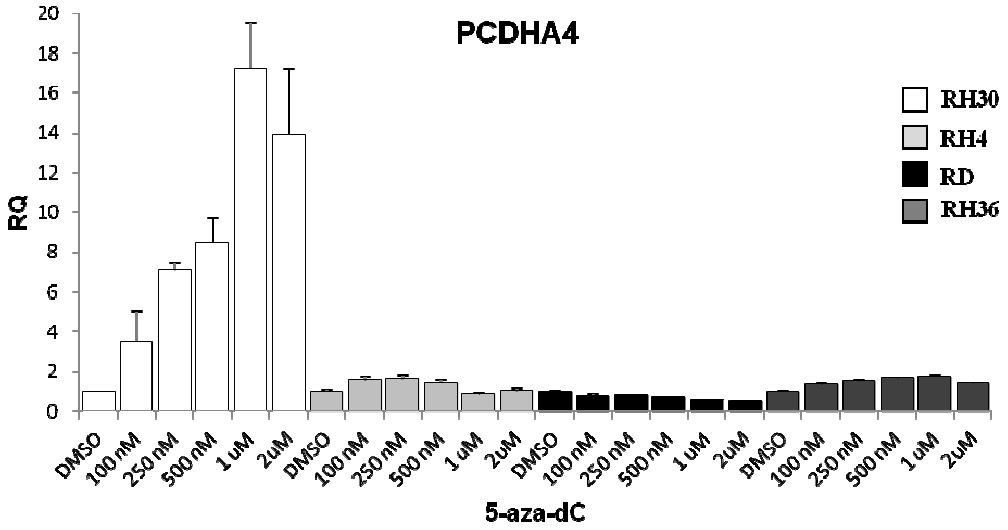


Figure 10. Relative expression of *PCDHA4* at 72 hours from treatment with increasing doses of 5-aza-2'-deoxycytidine (100 nM, 250 nM, 500 nM, 1 uM and 2 uM) in 4 RMS cell lines. Expression levels were assessed by qRT-PCR. Housekeeping *GAPDH* gene was used as internal control for normalization and DMSO-treated cells as calibrator. RH30, RH4: metastatic RMS cell lines; RD, RH36: non-metastatic RMS cell lines; 5-aza-dC: 5-aza-2'-deoxycytidine.RQ=relative expression ratio.

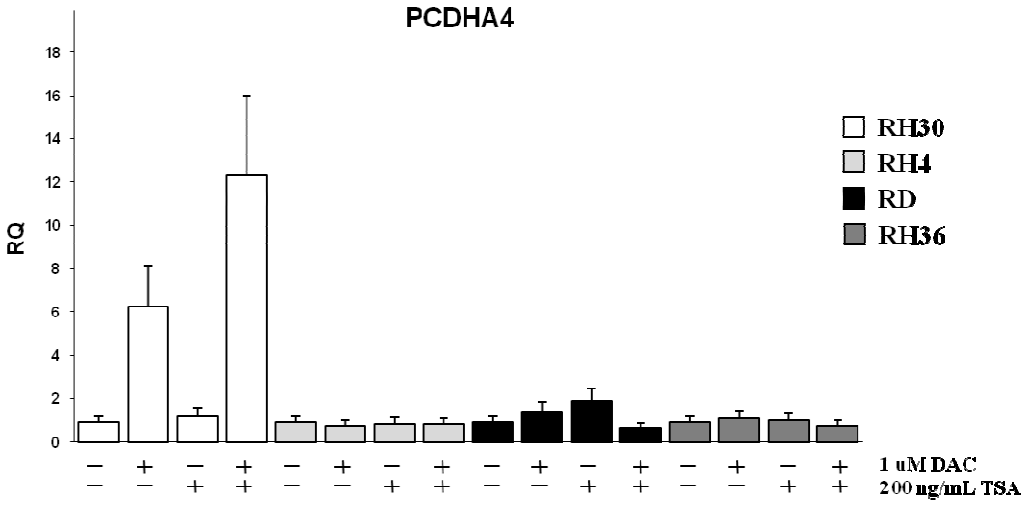


Figure 11. Relative expression of *PCDHA4* after treatment with 1uM of 5-aza-2'-deoxycytidine (72 h), 200ng/mL of trichostatin A (16 h) or the combination of both. Expression levels were assessed by qRT-PCR. Housekeeping *GAPDH* gene was used as internal control for normalization and DMSO-treated cells as calibrator . RH30, RH4: metastatic RMS cell lines; RD, RH36: non-metastatic RMS cell lines; DAC: decitabine or 5-aza-2'-deoxycytidine; TSA: trichostatin A. RQ=relative expression ratio.

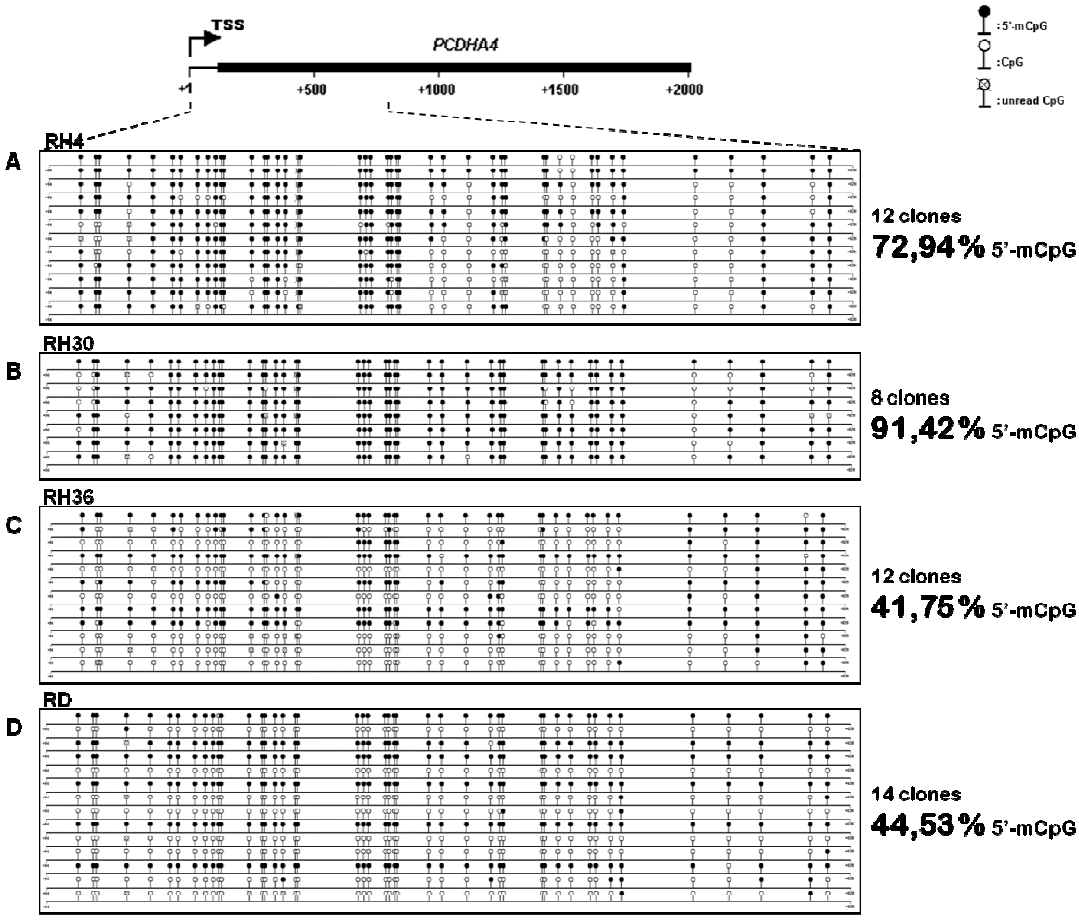


Figure 12. Sanger bisulfite sequencing of *PCDHA4* promotorial region in 4 RMS cell lines. Sequencing was performed for at least 8 clones obtained by subcloning bisulfite-converted promotorial region. Sequenced region spanning from position +94 to +828, where position +1 corresponds to the *trascriptional start site* (TSS). (**A, B**) Sequencing results of metastatic RMS cell lines (RH4 and RH30). (**C, D**) Sequencing results of non-metastatic RMS cell lines (RH36 and RD). The % mean value of methylation levels of the sequenced clones is shown on the right. Circles: cytosine within CpG dinucleotides; black circles: methylated cytosine; white circles: unmethylated cytosine.

4.1.2.4. Assessment of *PCDHA4* expression and methylation status in a larger cohort of patients

To determine the biological impact of *PCDHA4* in RMS tumors we correlated its methylation status and expression levels with clinicophatological variables and we addressed the likely prognostic value in RMS outcome. First of all we expanded the cohort of analyzed patients to 29 RMS biopsies (**Table 4**) and we proceeded with *PCDHA4* expression quantification by qRT-PCR. *PCDHA4* expression analysis confirmed the statistically significance difference between metastatic and non-metastatic RMS samples (P=0.034) (**Figure 13**). The bisulfite Sanger sequencing of promotorial regions is a useful approach to

identify the methylation profile of a region, but is not a method suitable to scale up to many samples. Therefore, we decided to used a quantitative methylation–specific PCR (qMS-PCR) technique that allow to discriminate methylated from unmethylated sequences using specific pairs of primers. The control assay was performed using *in vitro* fully methylated DNA as template. We used the region that we sequenced in cell lines after bisulfite conversion to construct two qMS-PCR assays (**Figure 14**). One assay, amplified the region spanning from position +465 to +574 (site 1) and was selected because already used by Wang *et al.* (Wang *et al.*, 2014) for *PCDHA4* methylation level assessment in cervical cancer. We adapted their assay to a quantitative technique. Then we designed a second assay in a region closest to the TSS (from position +274 to +393, site 2), since our CpG Sanger sequencing data revealed that also this region is able to discriminate between metastatic and non-metastatic samples. The results obtained are highly variable and we did not observe any difference of methylation status between metastatic and non-metastatic samples. Therefore, we performed a Pearson' correlation analysis between *PCDHA4* expression and methylation values of site 1 and site 2 and we detected a statistically moderate significance for the second site (r=-0.32; P= 0.08) (**Figure 15 A**), that mapped in the closest region to TSS. Conversely, none correlation was observed for the first site analyzed (r=0.01138; P=0.9524) (**Figure 15 B**). For these reasons we decided to continue our analysis using only the methylation data obtained for the second site. Therefore, we performed parametric (Welch Two t-test) and non parametric (Wilcoxon rank sum test) tests to correlate the methylation and expression levels with clinical prognostic features of RMS: histology, presence of t(2;13) (q35;q14) translocation, sex, size of tumor and *IRS* group. In these analysis we found that *PCDHA4* expression levels did not show any correlation with clinical variables (**Figure 16 A-E**). Conversely, we observed that *PCDHA4* methylation status correlates with the tumor size (<5cm or >5cm, Wilcoxon rank sum test: P= 0.059) (**Figure 17 A-E**).

To investigate a prognostic value of *PCDHA4* expression and promoter methylation level, we performed survival curve analysis. We selected two groups of patients based on median value of *PCDHA4* expression and methylation levels (high levels and low levels) and built Kaplan-Meier curves for overall survival (OS) and progression free survival (PFS). For both, gene expression and methylation variables, the survival curves were not statistically significant (**Figure 18 A, B**).

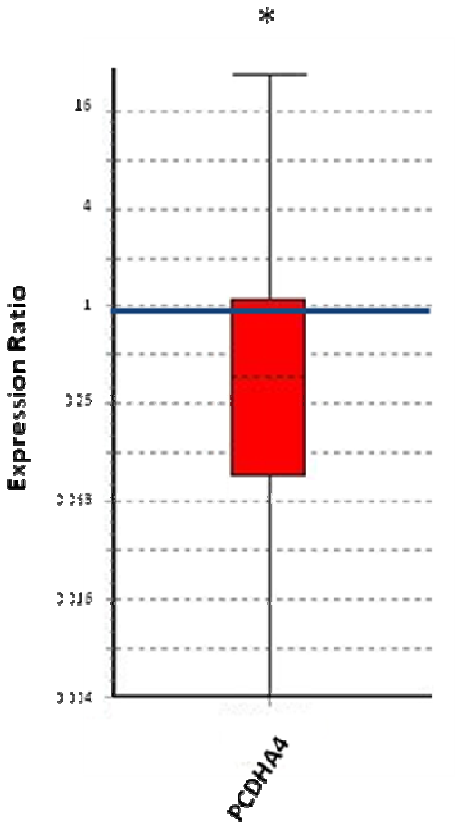


Figure 13. Relative expression analysis of *PCDH4* assessed in 12 metastatic RMS tumor samples compared to17 non-metastatic ones. Distribution of data is represented by box plot analysis. *GAPDH* was used as housekeeping gene for normalization. Expression ratio >1: high *PCDHA4* expression in metastatic RMS samples than non-metastatic ones; expression ratio <1: low *PCDHA4* expression in metastatic RMS samples than non-metastatic ones; expression ratio =1: equal expression of *PCDHA4* in metastatic RMS samples and non-metastatic ones. * $P<0.05$. The analysis was performed with REST software.

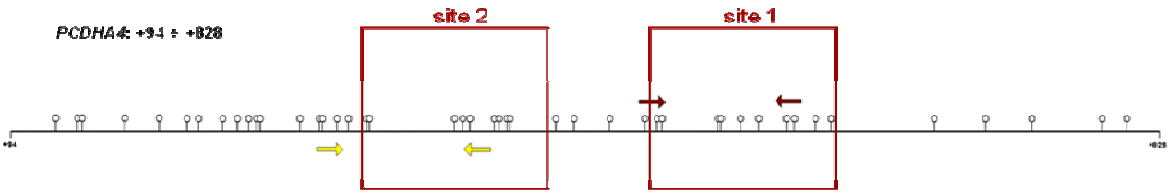


Figure 14. Schematic representation of the sequenced *PCDHA4* promotorial region spanning from position +94 to +828, where position +1 corresponds to the *trascriptional start site* (TSS). Red squares represent the site where we designed qMS-PCR assays. Site1: from +456 to +574; site2: from +274 to + 393; Circles: cytosine within CpG dinucleotides.

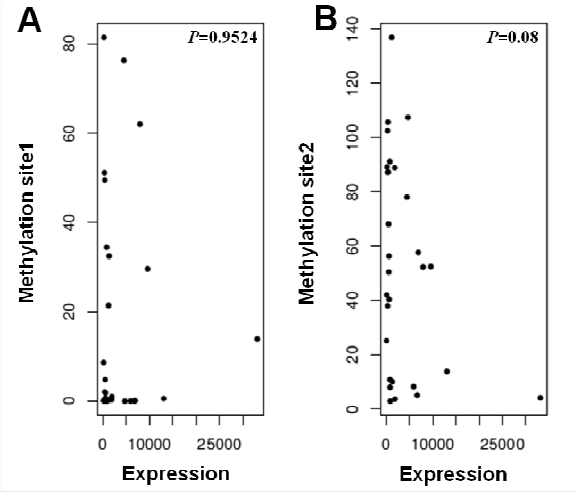


Figure 15. Pearson's correlation between *PCDHA4* expression level and methylation values of site 1 and site 2.

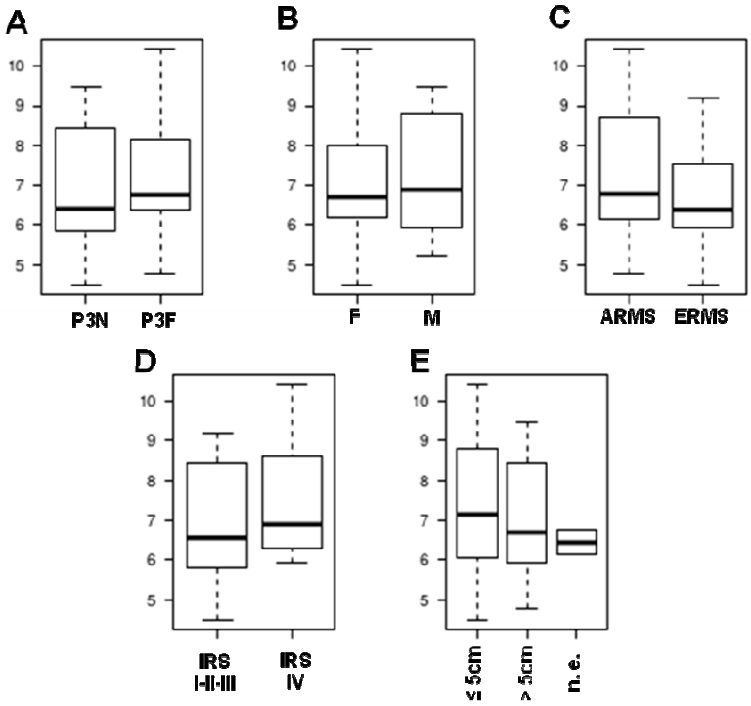


Figure 16. Correlation between *PCDHA4* expression levels and clinicopathological parameters in RMS tumours. Distribution of data is represented by box plot analysis comparing RMS samples based on (**A**) gene fusion status, (**B**) sex, (**C**) histology, (**D**) IRS group and (**E**) tumour size. *PCDHA4* expression levels were assessed by qRT-PCR in 29 RMS biopsies and normalized to *GAPDH*. *P*-values shown were calculated using Wilcoxon rank sum test. P3N: *PAX3-FOXO1* negative RMS biopsies, P3F: *PAX3-FOXO1* positive RMS biopsies, F: female, M: male, n.e.: not evaluable. * $P<0.05$.

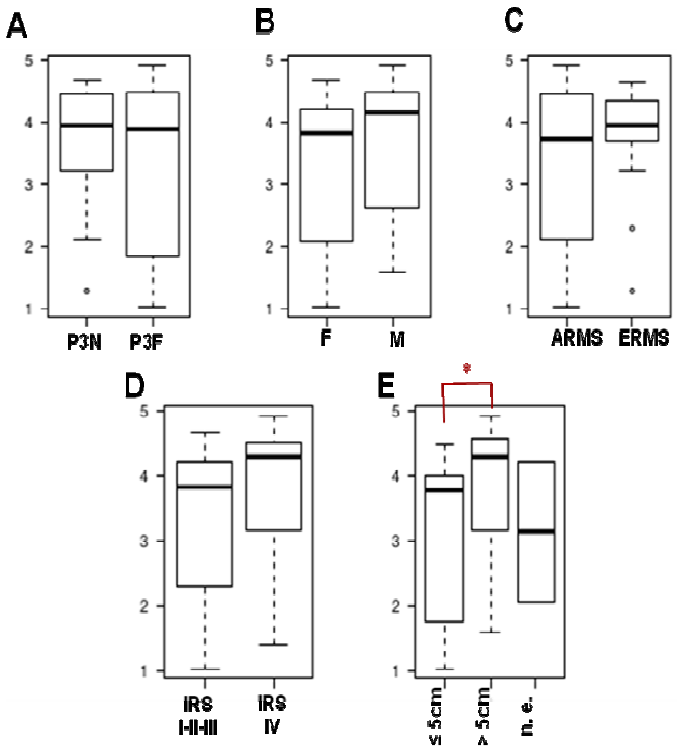


Figure 17. Correlation between PCDHA4 methylation levels of site 2 and clinicopathological parameters in RMS tumours. Distribution of data is represented by box plot analysis comparing RMS samples based on (**A**) gene fusion status, (**B**) sex, (**C**) histology, (**D**) IRS group and (**E**) tumour size. *PCDHA4* methylation levels were assessed by qMS-PCR in 29 RMS biopsies and normalized to *GAPDH*. As calibrator we used an *in vitro* fully methylated DNA. *P*-values shown were calculated using Wilcoxon rank sum test. P3N: *PAX3-FOXO1* negative RMS biopsies, P3F: *PAX3-FOXO1* positive RMS biopsies, F: female, M: male, n.e.: not evaluable. * $P<0.05$

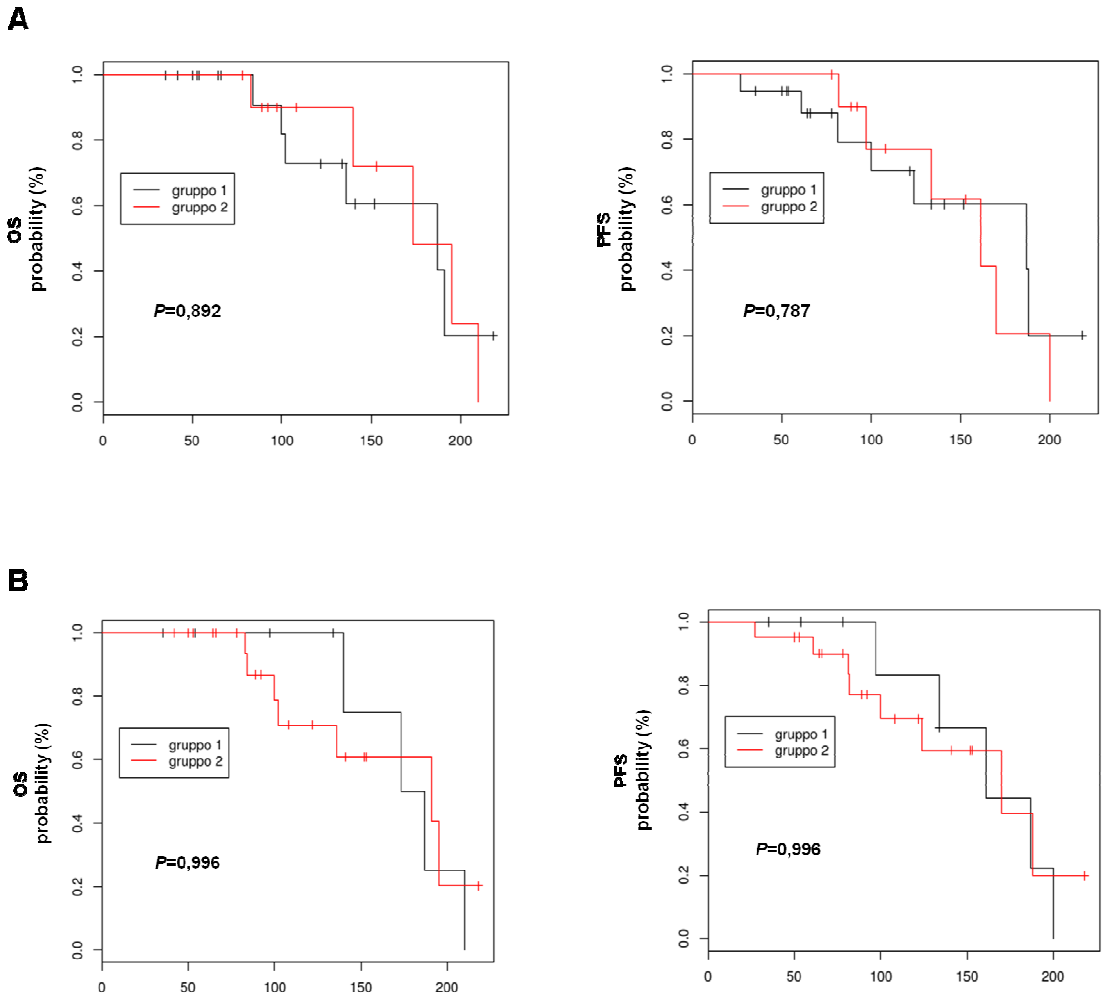


Figure 18. Kaplan-Meier curves considering (**A**) *PCDHA4* expression and (**B**) *PCDHA4* methylation level. The two groups indicate high or low levels of *PCDHA4* expression or methylation. Group 1: low *PCDHA4* expression or methylation levels (≤ median values); group 2: high *PCDHA4* expression or methylation levels (>median values). PFS=progression-free survival; OS= overall survival.

4.2. Evaluation of DNA methylation in RMS tumors biopsies using *Reduced-Representation Bisulfite Sequencing*

4.2.1. *Reduced-Representation Bisulfite Sequencing* analysis

We performed an Illumina sequencing to detect DNA methylation pattern of 15 RMS tumor biopsies including 7 ERMS and 8 *PAX3-FOXO1* positive ARMS (**Table 4**). The sequencing was performed using *Reduced-representation bisulfite sequencing* (RRBS) method at the BGI technology centre (BGI HongKong, Shenzhen, China; www.bgi-international.com).

RRBS is a bisulfite-based method that enriches CG-rich parts of the genome, thereby reducing the amount of sequencing required while capturing the majority of promoters and other relevant genomic regions. The approach provides single-nucleotide resolution and quantitative DNA methylation measurements as well.

The quality of DNA samples was evaluated using Qubit fluorimeter to assess the concentration of nucleic acids and gel agarose electrophoresis was used to test the integrity of DNA. All the 15 DNA analyzed passed the quality control (**Figure 19**).

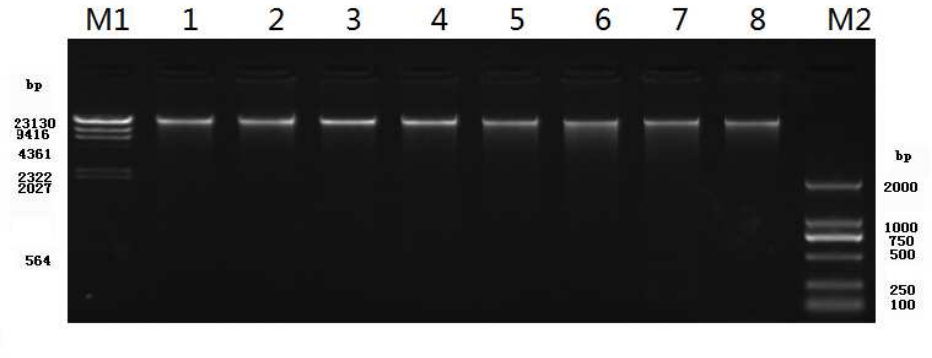


Figure 19. Agarose gel electrophoresis of some genomic DNA biopsies to determine samples integrity. Presence of a single band of high molecular weight confirmed the integrity of the analyzed samples. M1, M2: molecular weight marker; Lane 1: ARMS 2 genomic DNA; Lane 2: ARMS 6 genomic DNA; Lane 3: ERMS 8 genomic DNA; Lane 4: ERMS 7 genomic DNA; Lane 5: ARMS 19 genomic DNA; Lane 6: ARMS 26 genomic DNA; Lane 7: ARMS 28 genomic DNA; Lane 8: ERMS 22 genomic DNA.

Standard bioinformatic analysis was performed at BGI Biotechnology centre and are summarized below. The raw reads were filtered, including removing adaptor sequencing, contamination and low-quality reads (**Table 8**). Then, the sequenced reads were aligned to the reference by SOAP aligner (**Table 9**). Every hit with a single placement with minimum numbers of mismatches and a clear strand assignment was defined as unambiguous alignment

(uniquely mapped reads). Uniquely mapped reads which have restriction enzyme cutting site were used for the following analysis.

Sample Name	Insert size (bp)	Read lenght (bp)	Clean reads	Clean base (Gb)
ARMS 2	40-220	49	65'306'124	3.20
ARMS 3	40-220	49	65'306'124	3.20
ARMS 6	40-220	49	65'306'124	3.20
ERMS 7	40-220	49	65'306'124	3.20
ERMS 8	40-220	49	65'306'124	3.20
ERMS 10	40-220	49	68'496'770	3.36
ARMS 18	40-220	49	65'306'124	3.20
ARMS 19	40-220	49	64'050'928	3.14
ARMS 26	40-220	49	65'306'124	3.20
ERMS 21	40-220	49	65'306'124	3.20
ARMS 28	40-220	49	65'306'124	3.20
ERMS 22	40-220	49	65'306'124	3.20
ERMS 25	40-220	49	65'306'124	3.20
ARMS 31	40-220	49	65'306'124	3.20

Table 8. Data production by RRBS experiments. Gb: giga base pairs; bp: base pairs.

Sample Name	Clean reads (Mb)	Mapped reads (Mb)	Map rate	Uniquely mapped reads (Mb)	Uniquely mapped rate (%)	Enzyme cutting reads (Mb)	Enzyme cutting rate (%)	Bisulfite conversion rate (%)
ARMS 2	65.31	60.34	92.40	45.56	69.77	44.09	96.77	99.78
ARMS 3	65.31	58.92	90.23	44.55	68.22	43.13	96.80	99.77
ARMS 6	65.31	59.40	90.96	44.03	67.42	43.20	98.12	99.82
ERMS 7	65.31	59.35	90.87	45.38	69.48	43 .81	96.53	99.79
ERMS 8	65.31	59.97	91.83	45.29	69.35	44.15	97.48	99.80
ERMS 10	68.50	60.65	88.54	45.57	66.53	44.20	97.00	99.77
ARMS 18	65.31	59.97	91.83	45.82	70.16	44.89	97.98	99.78
ARMS 19	64.05	59.79	93.34	45.73	71.40	44.84	98.05	99.41
ARMS 26	65.31	57.27	87.69	41.98	64.29	41.10	97.90	99.33
ERMS 21	65.31	59.98	91.84	46.25	70.82	45.29	97.92	99.49
ARMS 28	65.31	51.64	79.07	37.98	58.16	36.20	95.32	99.18
ERMS 22	65.31	59.72	91.45	45.53	69.71	44.42	97.57	99.45

Sample Name	Clean reads (Mb)	Mapped reads (Mb)	Map rate	Uniquely mapped reads (Mb)	Uniquely mapped rate (%)	Enzyme cutting reads (Mb)	Enzyme cutting rate (%)	Bisulfite conversion rate (%)
ERMS 25	65.31	59.20	90.65	44.78	68.56	43.75	97.71	99.76
ARMS 31	65.31	58.61	89.75	43.25	66.22	42.19	97.55	99.80

Table 9. Results of reads alignment. Enzyme rate is the percentage of unique reads which have enzyme cutting site. Mb: mega base pairs.

Sequencing coverage describes the average number of reads that align to, or "cover," known reference bases. The next-generation sequencing coverage level often determines whether variant discovery can be made with a certain degree of confidence at particular base positions. Sequencing coverage requirements vary by application. However, at higher levels of coverage, each base was covered by a greater number of aligned sequence reads, so base calls can be made with a higher degree of confidence. The RRBS approach can covered most amount of promoters, although not all of genome's. The average coverage rate of CpG islands (CGIs) and promoters was calculated as the % ratio between practical value (number of CGIs or promoter regions covered by sequencing reads) and theoretical value (number of theoretical CGIs or promoter regions in enzyme cutting fragments than whole genome). Considering all the RMS samples the average coverage rate for promoter regions and CGIs were respectively 89.49 % (+/-1.43) and 93.06 % (+/-1.5) (**Table 10**).

Sample Name	Promoter	CGI
Genome	31’420	27’718
Enzyme fragment region	25’969	24’653
ARMS 2	23’335	23’000
ARMS 3	23’474	23’114
ARMS 6	23’238	22’926
ERMS 7	23’340	22’994
ERMS 8	23’441	23’184
ERMS 10	23’460	23’177
ARMS 18	23’366	23’064
ARMS 19	23’338	23’037
ARMS 26	22’798	22’592
ERMS 21	23’526	23’202
ARMS 28	22’041	21’698

Sample Name	Promoter	CGI
ERMS 22	23’435	23’131
ERMS 25	23’246	22’990
ARMS 31	23’235	22’979

Table 10. Number of covered promoter regions and CGIs.

The coverage rate of CpG dinucleotides at promoter regions and CGIs in each different depth was shown in **Figure 20**. Theoretical coverage rate was the percentage of CpG dinucleotides which located in enzyme cutting fragments.

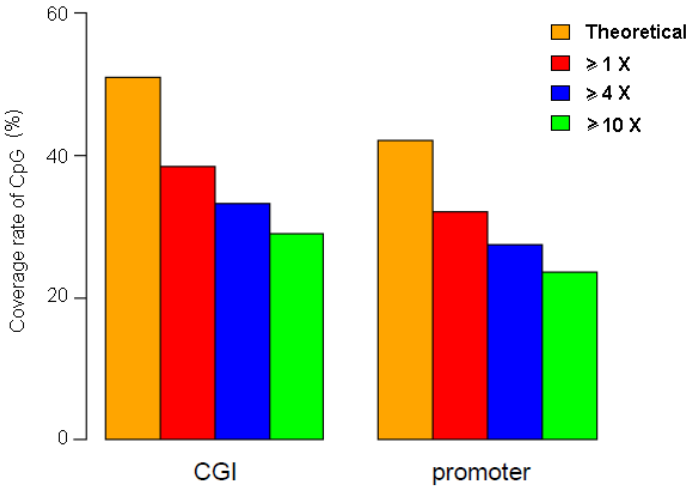


Figure 20. Theoretical CpG coverage and practical coverage in different depth at CGIs and promoter regions. Bar graph compares the different depth with the theoretical value, each region has four bars with different colors representing different depth (1X, 4X and 10X), and the vertical axis represents the coverage rate.

Moreover, it was calculated the average methylation level of cytosine in CpG dinucleotides for each sample dividing the number of the reads covering each 5mC by the total reads covering that cytosine. The average methylation level in CGIs and promoter context ranged from 40 to 60%, showing the same distribution in all analyzed samples (**Table 11**).

Sample Name	Promoter (%)	CGI (%)
ARMS 2	51.94	49.22
ARMS 3	52.16	51.44

Sample Name	Promoter (%)	CGI (%)
ARMS 6	48.48	50.62
ERMS 7	57.82	58.33
ERMS 8	48.40	51.65
ERMS 10	51.57	50.10
ARMS 18	48.57	47.68
ARMS 19	58.70	51.82
ARMS 26	42.76	42.37
ERMS 21	52.55	47.98
ARMS 28	46.16	42.20
ERMS 22	50.85	50.42
ERMS 25	55.46	56.05
ARMS 31	48.10	47.39

Table 11. Average methylation level of cytosine in CpG dinucleotides at CGIs and promoter regions.

Finally, were identified the *differently methylated regions* (DMRs) comparing ERMS and *PAX3-FOXO1* positive ARMS samples using Bsmooth, an opensource software (Kasper *et al.*, 2012). Each DMR was associated to a *P*-value: at a smaller *P*-value matched a greater difference in methylation value. Therefore, a methylation value >0 means a hypermethylation in ERMS than *PAX3-FOXO1* positive ARMS, on the contrary if methylation value is <0 a hypomethylation in ERMS was observed with respect to *PAX3-FOXO1* positive ARMS.

4.2.2 Functional annotation clustering of differentially methylated regions

To analyze the DMRs identified by Illumina sequencing we mapped all obtained DMRs against the genome. First of all, we identified the DMRs associated to the putative promoters (promoter-DMRs), that we defined as regions ranging from position -2000 to +1000, where +1 corresponding to *transcription start site* (TSS). Then, we identified the DMRs that overlapped the gene bodies (genebody-DMRs) and finally, the DMRs "distal" to genes (+/- 50 Kb to known genes) (distal-DMRs). The results are summarized in **Table 12**.

	Number of DMRs
promoter-DMRs	2733
genebody-DMRs	1959
distal-DMRs	1307
not-annotated-DMRs	818
Total	**6817**

Table 12. Distribution of DMRs found by RRBS analysis on the genome. DMRs associated to promoter, gene body and distal regions. It is also reported the number of DMRs associated with not annotated sequences.

We focused our attention first on promoter-DMRs. Then, a filter for methylation value (+/- 0.4) was applied obtaining a list of 1252 promoter-DMRs associated to transcripts. Therefore, we performed a functional annotation clustering using DAVID web tool (https://david.ncifcrf.gov) distinguishing a set of enriched classes and enriched pathways, with *P*-value<0.05. This analysis allowed us to discovery genes directly or indirectly modulated by DNA methylation. We found a consistent number of genes involved in transcription, insulin receptor signaling, chromatin modification, DNA repair, cell migration and actin cytoskeleton organization (**Table 13**). Interestingly, we found also an enrichment of pathways that are linked to altered processes in cancer (**Table 14**).

Annotation Cluster *	Enrichement score	*P*-value	Number of genes
Transcription	8,52	1,20E-006	339
Insulin receptor signaling pathway	3,18	2,30E-005	16
Chromatin modification	3,02	3,60E-003	52
DNA repair	2,4	1,20E-002	51
Positive regulation of cell migration	1,67	8,00E-003	21

Table 13. List of enriched classes found by functional annotation clustering of promoter-DMRs. *Biological process.

ID KEGG pathway	KEGG pathways	Number of genes	*P*-Value
hsa05200	Pathways in cancer	71	1.94E-6
hsa04722	Neurotrophin signaling pathway	35	4.25E-6
hsa04010	MAPK signaling pathway	56	7.44E-5
hsa04916	Melanogenesis	27	1.26E-4
hsa05215	Prostate cancer	25	1.46E-4
hsa04512	ECM-receptor interaction	23	4.31E-4
hsa04810	Regulation of actin cytoskeleton	45	4.57E-4
hsa04310	Wnt signaling pathway	34	7.13E-4
hsa04910	Insulin signaling pathway	31	9.24E-4
hsa04510	Focal adhesion	41	0.001
hsa04914	Progesterone-mediated oocyte maturation	22	0.001
hsa05210	Colorectal cancer	21	0.002
hsa04540	Gap junction	21	0.005
hsa05217	Basal cell carcinoma	15	0.006
hsa05222	Small cell lung cancer	20	0.006
hsa04340	Hedgehog signaling pathway	15	0.007
hsa04912	GnRH signaling pathway	22	0.008
hsa04270	Vascular smooth muscle contraction	24	0.009
hsa04110	Cell cycle	26	0.010
hsa05211	Renal cell carcinoma	17	0.010
hsa04114	Oocyte meiosis	23	0.015
hsa04720	Long-term potentiation	16	0.017
hsa04662	B cell receptor signaling pathway	17	0.020
hsa05214	Glioma	15	0.020
hsa05213	Endometrial cancer	13	0.023
hsa05218	Melanoma	16	0.026
hsa05223	Non-small cell lung cancer	13	0.030
hsa00534	Heparan sulfate biosynthesis	8	0.035
hsa04012	ErbB signaling pathway	18	0.037

Table 14. List of enriched KEGG pathways found by functional annotation clustering of promoter-DMRs.

4.2.3. Correlation of gene expression and promoter methylation status in RMS

The gene ontology analysis allowed us to give functional information of genes and pathway associated with DMRs identified comparing ERMS and *PAX3-FOXO1* positive ARMS. To verify if methylation status of the promoters reflected a different expression of downstream genes we performed bioinformatics analysis using gene expression datasets obtained by microarray experiments on RMS tumor biopsies or RMS cell lines available in literature. We analyzed gene expression profiling of 4 different studies (Davicioni *et al.*,2006, Whachtel *et al.*, 2004, De Pittà *et al.*, 2006, Rapa *et al.*, 2012). We focused our attention in genes differentially expressed comparing *PAX3-FOXO1* positive ARMS and ERMS samples that had matched with our methylation dataset. We found a set of 111 genes for which were available both gene expression and methylation data. We observed an anti-correlation between gene expression and promoter methylation status in only 50-60% of analyzed genes. We used the functional annotation web tool DAVID to performed a GO-term analysis and to identify which functional categories are over-represented. The terms analyzed had an adjusted *P*-value <0.05 and are shown in **Table 15**.

GO-term	Enrichement score	*P*-value
Angiogenesis	2.17	3.0*10-4
Cell proliferation	2.14	3.2*10-2
MAPKKcascade	2.12	2.9*10-4
Cell adhesion	1.60	5,.4*10-2
WNT signalling	1.36	8.4*10-3
Pathway in cancer	1.36	5.6*10-2

Table 15. List of enriched GO-term found by functional annotation clustering of genes which had an inverse correlation between gene expression and methylation status.

4.2.4. Validation of candidate genes by qRT-PCR in RMS cell lines and tumor biopsies

Based on anti-correlation gene expression/methylation and on literature data we selected some target genes for validation studies (**Table 16**).

Gene Symbol	Official full name
SNAI1	snail family zinc finger 1
ADAM9	ADAM metallopeptidase domain 9
ADAM22	ADAM metallopeptidase domain 22
CDC14B	cell division cycle 14B
NELL1	neural EGFL like 1
HDAC11	histone deacetylase 11
GAB2	growth factor receptor bound protein 2-associated protein 2
FAM3C	family with sequence similarity 3 member C
IRS2	insulin receptor substrate 2
TNFAIP3	tumor necrosis factor, alpha-induced protein 3
CYR61	cysteine rich angiogenic inducer 61
NRP2	neuropilin 2
GADD45G	growth arrest and DNA-damage-inducible 45 gamma
CSNK1E	casein kinase 1, epsilon
PIAS1	protein inhibitor of activated STAT 1
CCND1	cyclin D1
CFL1	cofilin 1
POU4F1	POU class 4 homeobox 1

Table 16. Genes selected for validation studies based on assessment of inverse correlation between expression and methylation, besides literature data.

We assessed the expression levels of the 18 genes in 7 RMS cell lines by qRT-PCR assay (**Table 5, Figure 21 A**). We found only for 9 genes an inverse correlation between expression and methylation comparing the two subgroups *PAX3-FOXO1* positive *vs* PAX3-FOXO1 negative. The expression levels is significant for *SNAI1* ($P<0.001$), *CDC14B* ($P=0.034$), *NELL1* ($P=0.023$), *HDAC11* ($P<0.001$), *TNFAIP3* ($P=0.003$), *CYR61* ($P=0.021$), *NRP2* ($P=0.003$), *GADD45G* ($P<0.001$) and *POU4F1* ($P<0.001$) (**Figure 21 A**).

Subsequently, we analyzed in 25 RMS tumor biopsies (**Table 4**) the genes that shown a good anti-correlation in RMS cells. We found that *SNAI1* ($P<0.001$), *CDC14B* ($P=0.006$), *NELL1* ($P=0.015$), *GADD45G* ($P=0.017$), *TNFAIP3* ($P=0.017$) and *POU4F1* ($P=0.002$) were statistically different in *PAX3-FOXO1* positive RMS samples than in PAX3-FOXO1 negative ones (**Figure 21 B**).

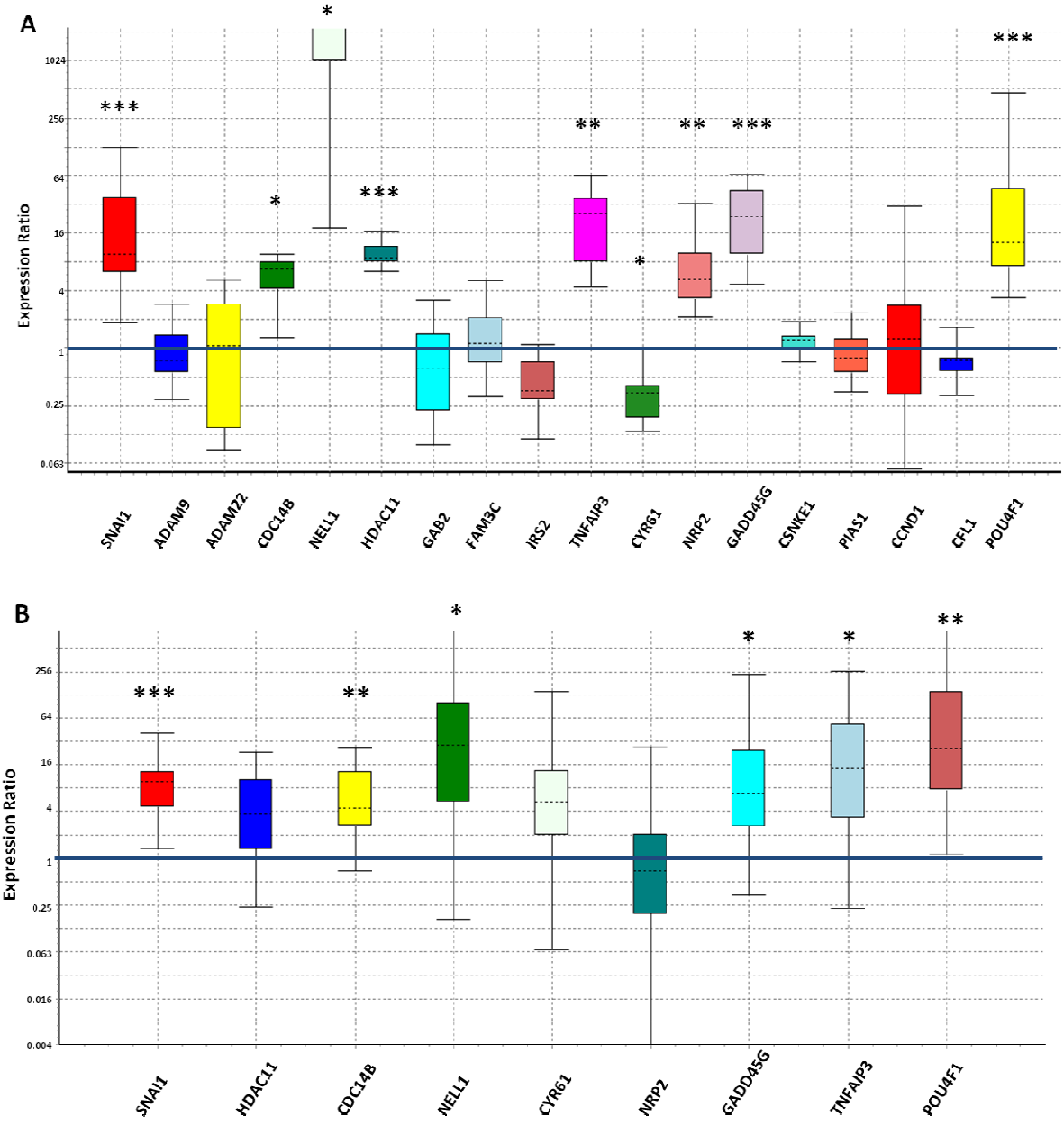
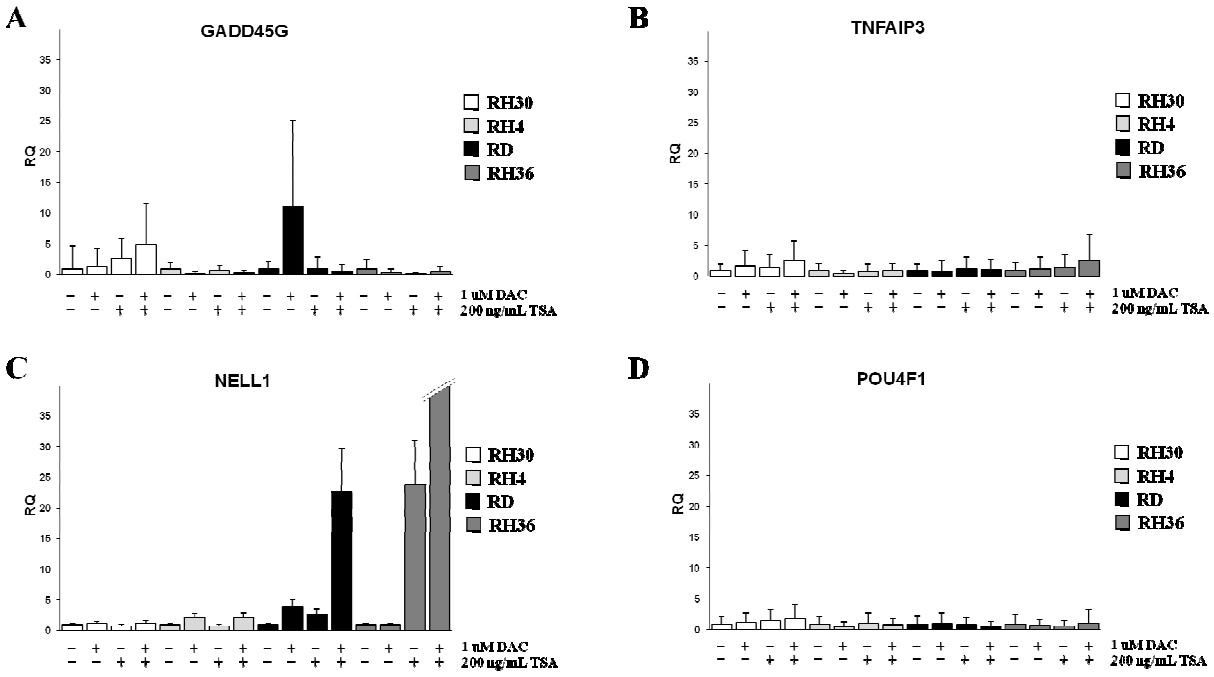


Figure 21. RRBS data validation by qRT-PCR analysis. (**A**) Relative expression levels of 18 selected genes in 3 *PAX3-FOXO1* positive RMS cell lines compared to 4 *PAX3-FOXO1* negative ones. (**B**) Relative expression levels of the 9 selected genes in 10 *PAX3-FOXO1* positive RMS tumor samples compared to 15 *PAX3-FOXO1* negative ones. Distribution of data is represented by box plot analysis. Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was used as housekeeping gene for normalization. Expression ratio >1: high expression in *PAX3-FOXO1* positive RMS samples than *PAX3-FOXO1* negative ones; expression ratio <1: low expression in *PAX3-FOXO1* positive RMS samples than *PAX3-FOXO1* negative ones; expression ratio =1: equal expression in *PAX3-FOXO1* positive RMS samples and *PAX3-FOXO1* negative ones. * $P<0.05$; ** $P<0.01$; *** $P<0.001$.

4.2.5. Restoration of expression after DAC and TSA treatment in RMS cell lines.

To study the possible role of promoter methylation in the regulation of gene expression we performed *in vitro* treatment with 5-aza-2'-deoxycytidine, trichostatin A and a combination of both drugs. For these experiments, we used RH30 and RH4 as *PAX3-FOXO1* positive RMS cell lines and RH36 and RD as *PAX3/7-FOXO1* negative RMS cell lines. We assessed the expression levels by qRT-PCR after 72 hours of treatment with DAC and after 16 hours of treatment with TSA. We observed a moderate restoration of *GADD45G* expression in RD cells after DAC treatment, while no change in RH36 cell lines either with DAC treatment neither with its combination with TSA (**Figure 22 A**). Conversely, the restoration of *NELL1* expression was evident after treatment with DAC+TSA in RD cell lines ,and with only DAC and DAC+TSA (more than 100 folds) in RH36 cells (**Figure 22 C**). *TNFAIP3*, *POU4F1*, *SNAI1* and *CDC14B* did not show significant changes in their expression levels by treatment with DAC, TSA or DAC+TSA (**Figure 22 B, D, E, F**). These data confirmed the involvement of DNA methylation processes in *NELL1* gene expression regulation besides the synergistic action of histone acetylation. Even for *GADD45G* we verified a regulation under methylation control, although more moderate than that observed for *NELL1*.



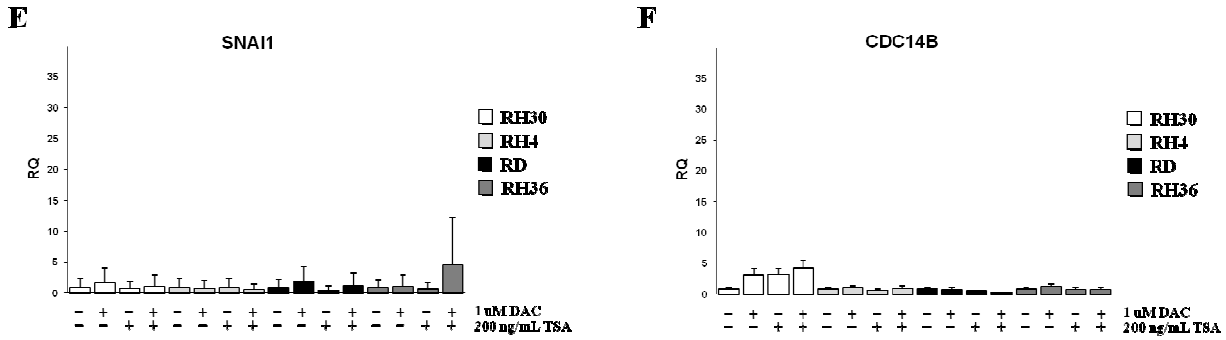


Figure 22. Relative expression of (**A**) *GADD45G*, (**B**) *TNFAIP3*, (**C**) *NELL1*, (**D**) *POU4F1*, (**E**) *SNAI1* and (**F**) *CDC14B,* assessed after treatment with 1uM of 5-aza-2'-deoxycytidine (72 h), 200 ng/mL of trichostatin A (16 h) or the combination of both. Expression levels were assessed by qRT-PCR. Housekeeping *GAPDH* gene was used as internal control for normalization and DMSO-treated cells as calibrator. RH30, RH4: *PAX3-FOXO1* positive RMS cell lines; RD, RH36: *PAX3-FOXO1* negative RMS cell lines; DAC: 5-aza-2'-deoxycytidine; TSA: trichostatin A. RQ= relative expression ratio.

4.2.6. Bisulfite sequencing of *NELL1* and *GADD45G* regulative regions in RMS cell lines reveals a different methylation pattern

Based on *in vitro* treatments we decided to focus our attention on *GADD45G* and *NELL1* genes. To validate previous results we performed a bisulfite Sanger sequencing in *PAX3-FOXO1* positive cell lines (RH30 and RH4) and *PAX3-FOXO1* negative ones (RH36 and RD). To design primers for sequencing we used methPrimer, a free online tool (Li and Dahiya, 2002). It is a program for designing PCR primers for methylation detection at CpG level. Primers are picked around predicted CpG islands or around regions specified by users on the input sequences that is subject to an *in-silico* bisulfite conversion assuming all cytosine in CpG context were methylated. In the output table of methPrimer we selected primers that amplify a region downstream the TSS for NELL1, ranging from position +330 to position +579 and two regions for *GADD45G* promoter, one spanning from -1175 to -915 (region 1) than TSS and one from -796 to -504 (region 2). Therefore, we used those primers to amplified the candidate regions by PCR using genomic DNA previously converted using bisulfite treatment. Then we subcloned PCR products in pSC-A-amp/kan vector and sequenced at least eight clones for cell lines. We found an hypermethylation of *NELL1* in *PAX3-FOXO1* negative cell lines (63.88% of 5'm-CpG in RH36 and 53.46% of 5'm-CpG in RD) than *PAX3-FOXO1* positive cell lines (0% of 5'm-CpG in RH4 and 22.91% of 5'm-CpG in RH30) (**Figure 23**). This result confirmed that methylation of the promoter region of *NELL1* could modulate the gene expression. Indeed, we found that RMS *PAX-FOXO1* positive cells that

showed low level of *NELL1* genes had a hypermethylation of the promoter regions. Also for *GADD45G* we got the same results even if the difference between the two analyzed groups was much less marked than in *NELL1*. Indeed, while *PAX3-FOXO1* positive cell lines were almost totally unmethylated (region 1: 0% of 5'm-CpG in RH4 and RH30; region 2: 0.96% of 5'm-CpG in RH4 and of 1.92% 5'm-CpG in RH30), *PAX3-FOXO1* negative cell lines were slightly more methylated (region 1: 0.73% of 5'm-CpG in RD and 6.61% of 5'm-CpG RH36; region 2: 1.92% of 5'm-CpG in RD and 24.51% of 5'm-CpG in RH36) (**Figure 24**).

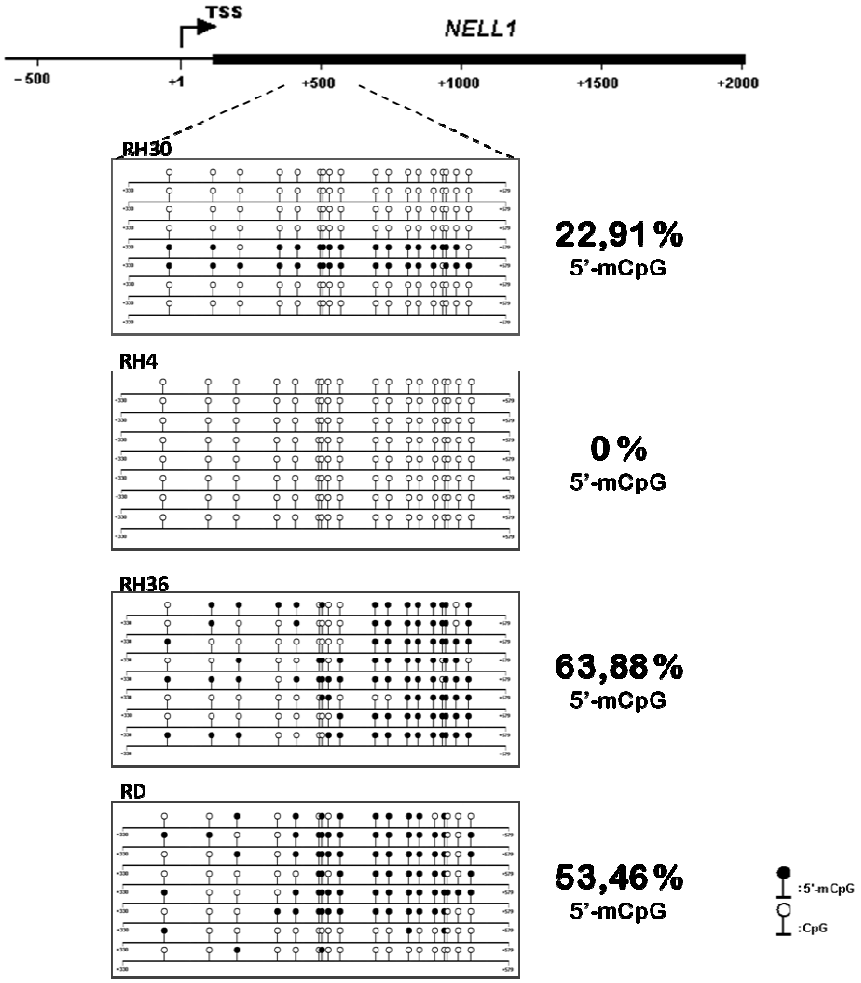


Figure 23. Sanger bisulfite sequencing of *NELL1* promotorial region in 4 RMS cell lines. Sequencing was performed for at least 8 clones obtained by subcloning bisulfite-converted promotorial region. Sequenced region spanning from position +330 to +579, where position +1 corresponds to the *trascriptional start site* (TSS). The % mean value of methylation levels of the sequenced clones is shown on the right. Circles: cytosine within CpG dinucleotides; black circles: methylated cytosine; white circles: unmethylated cytosine; RH4, RH30: *PAX3-FOXO1* positive RMS cell lines. RH36, RD: *PAX3-FOXO1* negative RMS cell lines.

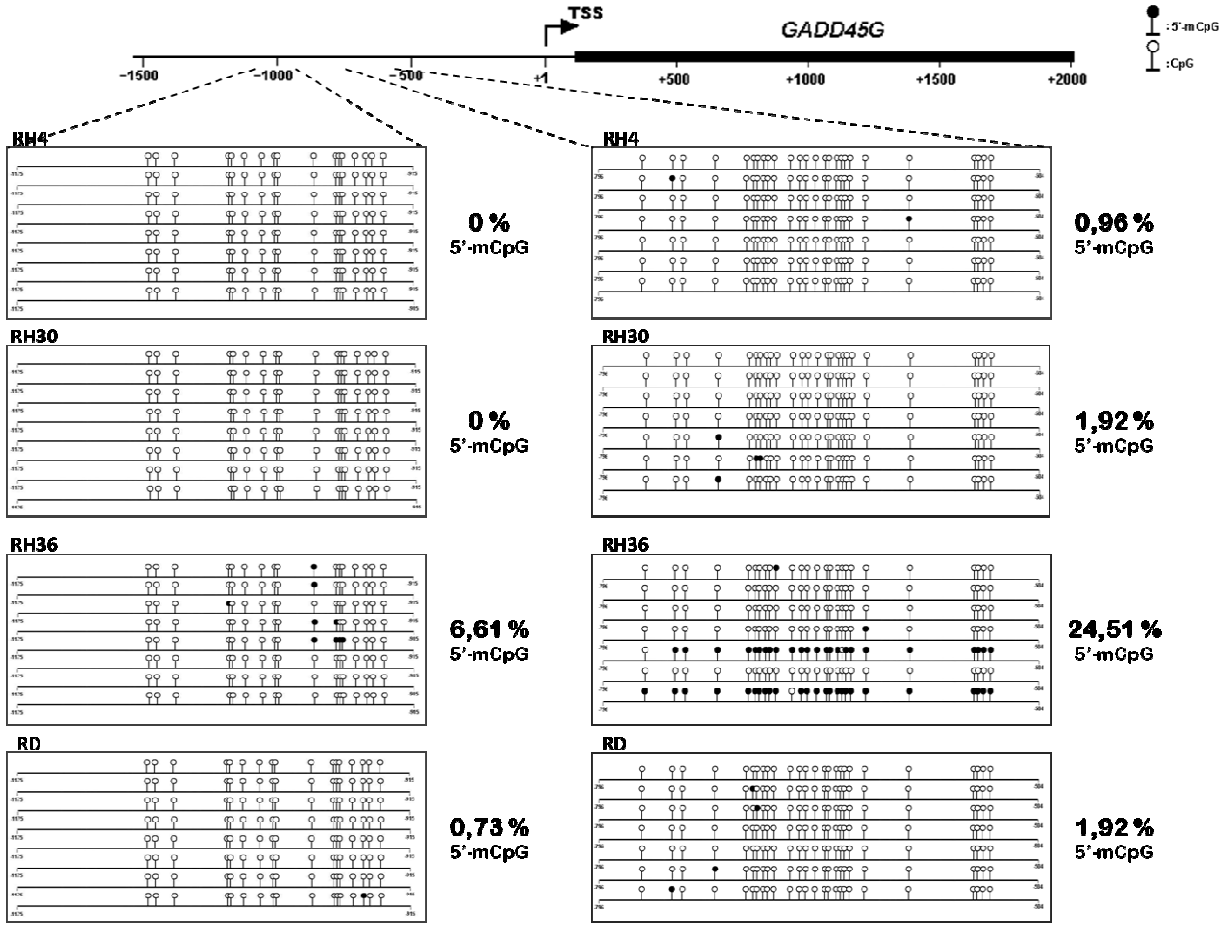


Figure 24. Sanger bisulfite sequencing of *GADD45G* promotorial region in 4 RMS cell lines. Sequencing was performed for at least 8 clones obtained by subcloning two bisulfite-converted regions. Sequenced region 1 (left panels) spanning from position -1175 to -915, while sequenced region 2 ranging from position -796 to -504 (right panels). Position +1 corresponds to the *trascriptional start site* (TSS). The *%* mean value of methylation levels of the sequenced clones is shown to the right of the panels. Circles: cytosine within CpG dinucleotides; black circles: methylated cytosine; white circles: unmethylated cytosine; RH4, RH30: *PAX3-FOXO1* positive RMS cell lines. RH36, RD: *PAX3-FOXO1* negative RMS cell lines.

4.2.7. NELL1 and GADD45G expression in RMS tumor biopsies: statistical analysis and prognostic impact

To determine the biological impact of *NELL1* and *GADD45G* in RMS tumors we evaluated their expression levels in a large cohort of RMS samples by qRT-PCR, using fetal skeletal muscle as calibrator. Moreover, we correlated the expression levels of *NELL1* and *GADD45G* with clinicophatological variables and we addressed the likely prognostic value in RMS outcome.

Survival and correlation analysis was conducted on 73 RMS biopsies divided into 23 *PAX3-FOXO1* positive ARMS, 7 *PAX7-FOXO1* positive ARMS, 9 *PAX3/7-FOXO1* negative

ARMS, 37 ERMS (**Table 4**). We divided the samples considering the most prognostic clinical variables such as the presence or absence of translocations t(2;13)(q35;q14) or t(1;13)(p36;q14), histology (alveolar or embryonal) and the IRS group (metastatic or non-metastatic) and we performed Student's t-tests and Wilcoxon-Mann-Whitney tests. First, we confirmed that in the expanding cohort of RMS patients *NELL1* and *GADD45G* expression levels were statistically significant different comparing RMS samples with or without translocations, independently from histology (*NELL1* Student's t-test and Wilcoxon-Mann-Whitney test: $P<0.0001$; *GADD45G* Student's t-test and Wilcoxon-Mann-Whitney test: $P<0.0001$) (**Figure 25 A, B**). Moreover, we found that the expression level of both *NELL1* and *GADD45G* were statistically significant different also between alveolar and embryonal RMS (*NELL1*: student's t-test: $P= 0.0004$, Wilcoxon-Mann-Whitney test: $P<0.0001$; *GADD45G*: student's t-test: $P=0.0001$, Wilcoxon-Mann-Whitney test: $P<0.0001$) (**Figure 26 A, B**). We also performed a statistical analysis dividing the samples based on IRS group (IRS IV *vs* IRS I-II-III). The presence of distal metastasis at the diagnosis of diseases (patients included in IRS group IV) is known as an important indicator of poor prognosis. The analysis showed statistically significative difference between IRS IV and IRS I-II-III samples for *NELL1* ($P=0.0182$) and even more pronounced for *GADD45G* ($P=0.0006$) (**Figure 27 A, B**). To establish if *NELL1* and *GADD45G* gene expression levels can contribute to discriminate *PAX3/7-FOXO1* positive RMS from *PAX3/7-FOXO1* negative one, we constructed *receiver operating curves* (ROC). The area under the curve (AUC) for *NELL1* was 0.9302 ($P<0.0001$) and for *GADD45G* was 0.9073 ($P<0.0001$) (**Figure 28 A, B**). We concluded that gene expression for both genes selected could be a good predictor to distinguish *PAX3/7-FOXO1* positive RMS and *PAX3/7-FOXO1* negative one.

Finally, we performed survival analysis dividing patients into groups on basis of the median value of *NELL1* expression levels for *NELL1* analysis (high *NELL1* > 12.33; low *NELL1* ≤ 12.33) and the median value of *GADD45G* expression levels for *GADD45G* ones (high *GAGG45G* > 2.41; low *GAGG45G* ≤ 2.41). Kaplan-Meier analysis revealed that high levels of *NELL1* and *GADD45G* gene expression had a significantly poorer prognosis than low levels of *NELL1* and *GADD45G* transcripts, when we consider progression free survival (PFS) (*NELL1*: $P=0.0245$; *GADD45G*: $P=0.0247$) (**Figure 29 A, B**). While, we did not find any statistical difference when we consider the overall survival (OS) for both *NELL1* and *GADD45G* genes (*NELL1*: $P=0.0801$; *GADD45G*: $P=0.3514$) (**Figure 30 A, B**).

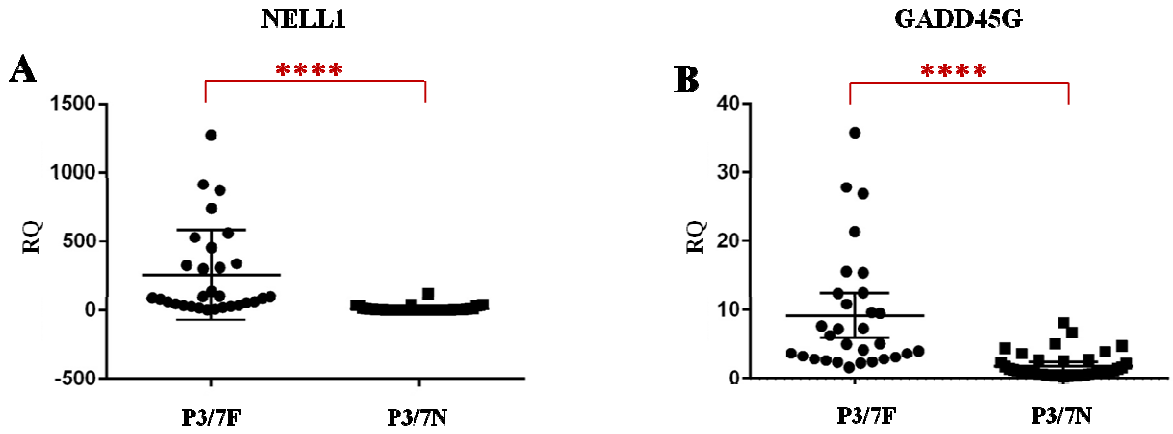


Figure 25. Correlation between (**A**) *NELL1* and (**B**) *GADD45G* expression levels with *PAX3/7-FOXO1* gene fusion status. *NELL1* and *GADD45G* expression levels were assessed by qRT-PCR in 73 RMS biopsies. Housekeeping *GAPDH* gene was used as internal control for normalization and fetal skeletal muscle as calibrator. *P*-values shown were calculated using Wilcoxon-Mann-Whitney test. P3/7N: *PAX3/7-FOXO1* negative RMS biopsies, P3/7F: *PAX3/7-FOXO1* positive RMS biopsies. * $P<0.05$; ** $P<0.01$; *** $P<0.001$;**** $P<0.0001$.RQ= relative expression ratio.

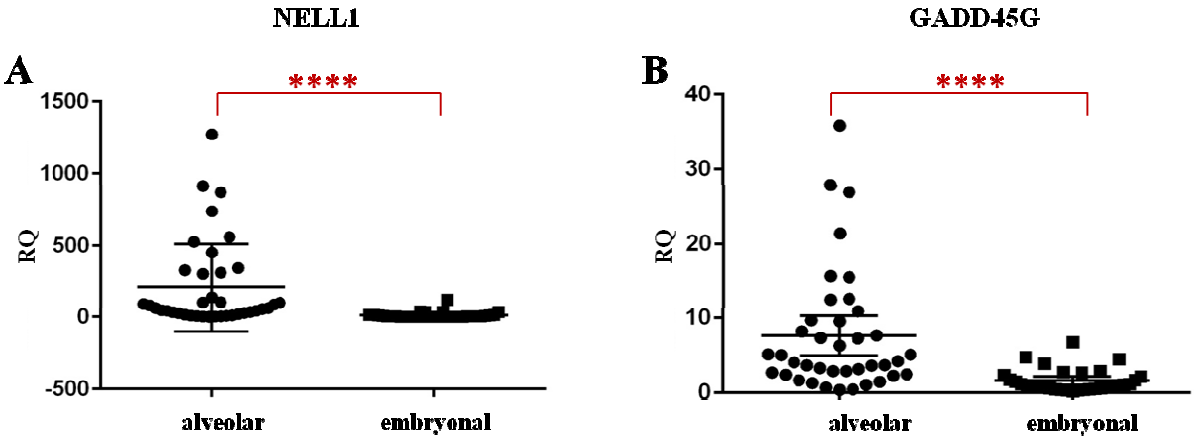


Figure 26. Correlation between (**A**) *NELL1* and (**B**) *GADD45G* expression levels with alveolar or embryonal histology. *NELL1* and *GADD45G* expression levels were assessed by qRT-PCR in 73 RMS biopsies. Housekeeping *GAPDH* gene was used as internal control for normalization and fetal skeletal muscle as calibrator. *P*-values shown were calculated using Wilcoxon-Mann-Whitney test. * $P<0.05$; ** $P<0.01$; *** $P<0.001$;**** $P<0.0001$. RQ= relative expression ratio.

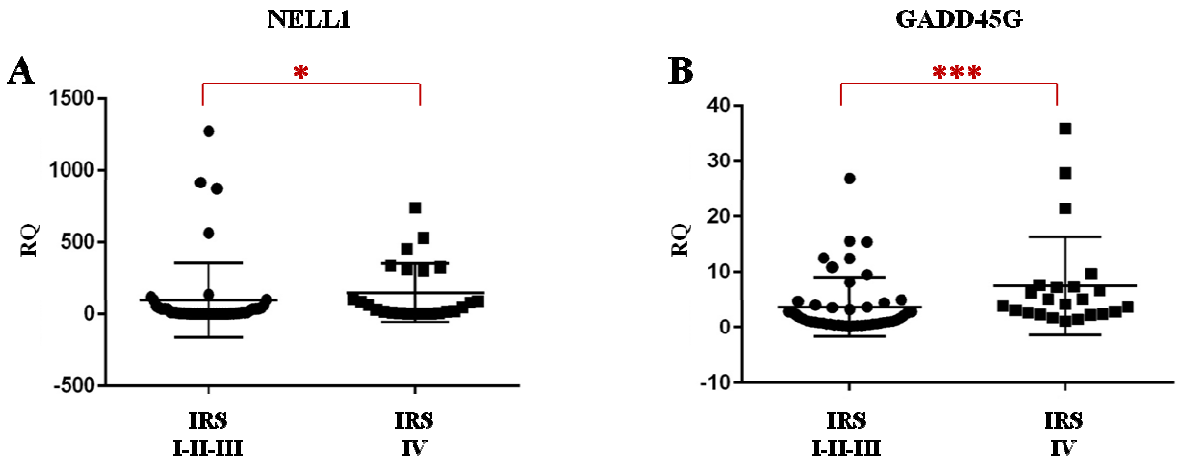


Figure 27. Correlation between (**A**) *NELL1* and (**B**) *GADD45G* expression levels with IRS group (IRS IV *vs* IRS I-II-III). *NELL1* and *GADD45G* expression levels were assessed by qRT-PCR in 73 RMS biopsies. Housekeeping *GAPDH* gene was used as internal control for normalization and fetal skeletal muscle as calibrator. *P*-values shown were calculated using Wilcoxon-Mann-Whitney test. IRS I-II-III: non-metastatic RMS tumors; IRS-IV: metastatic RMS tumors.* $P<0.05$; ** $P<0.01$; *** $P<0.001$. RQ= relative expression ratio.

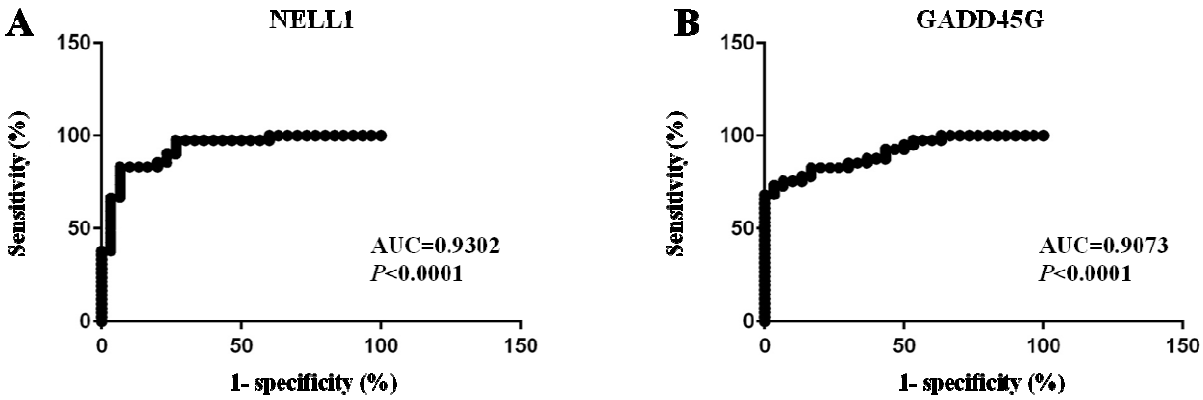


Figure 28. Receiver operating curve (ROC) analysis. ROCs show the sensitivity and specificity of (**A**) *NELL1* and (**B**) *GADD45G* as a parameter to classify RMS patients (n=73) on the basis of *PAX3/7-FOXO1* fusion gene status. *P*-value and the area under the curve (AUC) are reported in the panels.

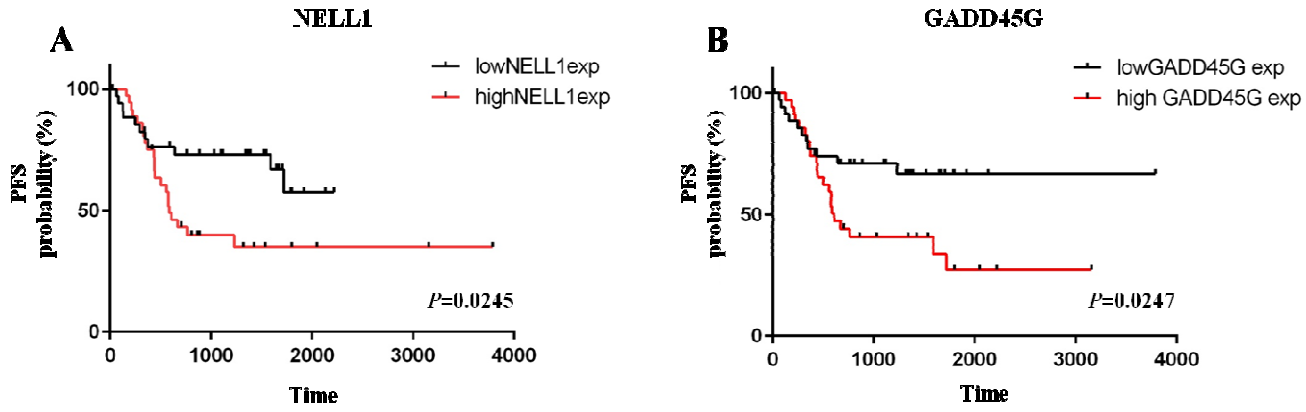


Figure 29. Kaplan-Meier curves for (**A**) *NELL1* and (**B**) *GADD45G* expression level. We performed progression-free survival (PFS) curves dividing samples based on high or low levels of *NELL1* and *GADD45G* expression. Group 1: low *NELL1* and *GADD45G* expression levels (≤ median values); group 2: high *NELL1* and *GADD45G* expression levels (> median values).

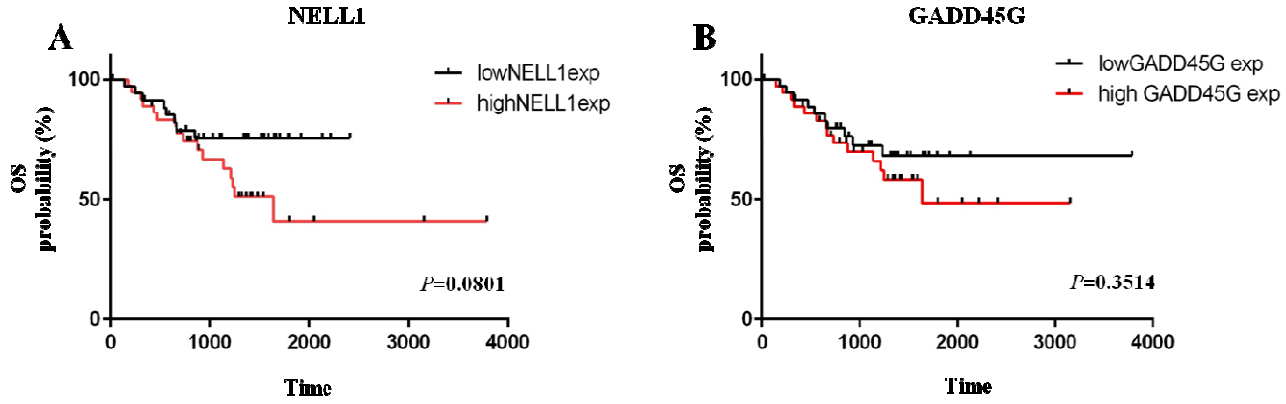


Figure 30. Kaplan-Meier curves for(**A**) *NELL1* and (**B**) *GADD45G* expression level. We performed overall survival (OS) curves dividing the samples based on high or low levels of *NELL1* and *GADD45G* expression. Group 1: low *NELL1* and *GADD45G* expression levels (≤ median values); group 2: high *NELL1* and *GADD45G* expression levels (> median values).

5. DISCUSSION

Several studies have investigated the different behavior and molecular features of *PAX3-FOXO1* positive and negative RMS. In the last decade many scientific studies confirmed that the gene expression profiles distinguish *PAX3-FOXO1* positive RMS from *PAX3-FOXO1* negative RMS (Davicioni *et al.*, 2006; De Pittà *et al.*, 2006; Missiaglia *et al.*, 2012; Hingorani *et al.*, 2015). Moreover, miRNA expression profiles in human RMS cell lines discriminate fusion-positive RMS from fusion-negative (Tombolan *et al.*, 2015). MicroRNAs are a class of small non-coding RNAs that negatively regulate protein-coding gene expression post-transcriptionally by targeting mRNAs. Like microRNAs, other epigenetic modifications such as DNA methylation, can modulate the expression of genes and represents an interesting field to explore in RMS. Definition of genomic methylation patterns will offer the possibility to delineate how these epigenetic changes may be specific to different subgroups of tumors and furthermore, could be a useful approach to development new therapeutic strategies.

We performed a genome-wide DNA methylation study on *PAX3-FOXO1* positive and negative RMS tumors, using microarray platform, to determine the epigenetic pattern of RMS subgroups and to investigate how it contribute to their different biological behaviour. From this comparison we found a small set of *differentially expressed regions* (DMRs) that are associated to promoter region of genes or linked to intergenic regions. The DNA methylation field is continuously evolving, but nowadays the inverse correlation between promoter methylation status and expression level is widely accepted and well documented (Watt and Molloy, 1988; Bird and Wolffe, 1999).

We selected a handful of target genes on which focus the validation studies. These genes are tumour suppressor and transcription factors that showed aberrantly methylation in several cancers or genes implicated in development, differentiation and apoptosis processes. Unfortunately, we observed a poor anti-correlation between methylation status and gene expression level in selected genes in both RMS cell lines and RMS tumours. Only *HOXC11* was found statistically hypermethylated and down-regulated in *PAX3-FOXO1* positive samples than in *PAX3-FOXO1* negative ones. *HOXC11* belongs to the family of homeotic genes, characterized by the presence of highly conserved homeobox domain. It is a transcription factor, whose altered expression has been demonstrated to have a role in tumor progression of cutaneous melanoma (deBlacam *et al.*, 2011) and, more recently, in cellular proliferation in renal cell carcinoma (Liu *et al.*, 2015). To demonstrate the involvement of

methylation in *HOXC11* gene expression regulation we performed *in vitro* pharmacological treatments and sequencing of its promoter after bisulfite treatment. We demonstrated that the treatment with 5-aza-2'-deoxycytidine restored *HOXC11* expression in *PAX3-FOXO1* positive cell lines of 100 folds compared with untreated cells. Moreover, the sequencing of *HOXC11* promoter region, after bisulfite conversion, confirmed that all cytosine in this region were methylated, contrary to what happened in *PAX3-FOXO1* negative cell lines in which all cytosine were unmethylated. Unfortunately, we did not observe the same results in biopsies samples analyzed. Despite cell lines are an excellent model for *in vitro* studies about tumor biology and the response to pharmacological treatment, they have some limitations such as culture and growing conditions leading to the onset of genomic alteration and cellular dynamic differences from the real tumor micro-environment.

Therefore, we explored the different methylation profiling between metastatic and non-metastatic RMS tumor to understand whether the presence of distal metastasis at the onset of the disease could be associated to epigenetic alterations. The treatment of patients with RMS is multidisciplinary and is based on risk factors like histology, tumor site and size, nodal involvement and distal metastasis. In particular, the presence of metastasis is the most powerful independent prognostic factor, as the 5 year survival drops from approximately 70% in localized tumors to less than 30% in metastatic RMS. Children with localized disease that present a metastatic relapse after treatment have an even worse prognosis, with a probability to be long term survivors lower than 5%. Then, define the molecular characteristics of patients with advance stage of tumor respect to that have a localized tumor could contribute to a better molecular classification of RMS and to the identification of new targets for therapy. The IRS group system is highly predictive of outcome of RMS (Crist *et al.*, 1995), in particular patients belong to group IRS IV, characterized by metastatic disease, have long term failure-free survival (FFS) rates of <30% (Breneman *et al.*, 2003; Oberlin *et al.*, 2008).

We performed a new analysis on DNA methylation data dividing the samples based on IRS group: IRS IV group is defined as metastatic group, while IRS I, II and III samples represent the non-metastatic group. The comparison between metastatic and non-metastatic RMS has revealed a high number of DMRs. The gene ontology (GO) analysis of genes linked to DMRs showed that cell adhesion class is enriched in members of the protocadherin family. Protocadherins (PCDHs) are a group of transmembrane proteins and constitute the largest subfamily of the cadherin cell-adhesion molecules. In mammals, two types of PCDH classes have been defined based on their genomic structure: the non-clustered PCDHs, which are scattered throughout the genome (Kim *et al.*, 2011) and the clustered PCDHs, which are

organized in three closely linked gene clusters designated as α, β and γ (Wu and Maniatis, 1999). These gene clusters are generally organized into a variable region containing a tandem array of variable exons and a constant region containing one set of constant exons (Takeshi, 2008). Microarray analysis revealed an enrichment of the clustered prothocadherins and a homogeneous representation of the three major classes. Protocadherins are predominantly expressed in the nervous system of developing vertebrates. It is demonstrated that clustered protocadherins regulate neuronal survival, as well as dendrite self-avoidance. Moreover, there are many studies which show an involvement of protocadherin in tumor processes. It is demonstrated that protocadherins behave as tumor suppressor genes and their involvement in cancer is due to aberrant DNA methylation that determine an altered pattern of expression. *PCDH10* promoter methylation is associated with poor prognosis in non-small-cell lung cancer (Harada *et al.*, 2015), in gastric cancer (Deng *et al.*, 2014) and in prostate cancer (Wang *et al.*, 2014), while *PCDHA8* methylation may be associated with tumor progression and poor prognosis in non-muscle invasive bladder cancer (Lin *et al.*, 2014; Lin *et al.*, 2013). *PCDH10* is also required for the proliferation and tumorigenicity of glioblastoma cells (Echizen *et al.*, 2014), while silencing of *PCDH17* protein expression through hypermethylation of the promoter leads to loss of its tumor-suppressive activity, which may be an event in the carcinogenesis of a subgroup of esophageal squamous cell carcinoma (Haruki *et al.*, 2010).

Our results, obtained with the microarray experiments, suggest that the aberrant methylation of protocadherins could be important in the biology of RMS and in particular in the metastatization processes. Unfortunately, when we assessed expression levels of some members of protocadherins cluster, in biopsies and in cells lines, we demonstrated an inverse correlation with methylation level only for one of them: *PCDHA4*. We demonstrated the role of DNA promoter methylation status in *PCDHA4* expression with 5-aza-2'-deoxycytine treatment and bisulfite sequencing. Indeed, when we treated metastatic cell lines with 5-aza-2'-deoxycytine, we restored the expression of *PCDHA4*. The bisulfite sequencing of its promotorial region confirmed once again the different methylation status of the *PCDHA4* promoter in metastatic RMS samples and non-metastatic ones. Coupling to 5-aza-2'-deoxycytine treatment the trichostatin A one, we demonstrated also a synergic involvement of histones acetylation mechanisms in *PCDHA4* regulation. Although, we demonstrated a different expression pattern of *PCDHA4* between metastatic and non-metastatic RMS samples, we did not find an involvement in the RMS outcome. Indeed, we did not observe any correlation between expression and methylation level of *PCDHA4* (assessed in a wider

cohort of patients) and clinical prognostic features of RMS, such as histology, presence of t(2;13)(q35;q14) translocation, sex and IRS group. Furthermore, the survival curve analysis did not reveal any correlation between *PCDHA4* expression or methylation levels and outcome of patients.

During the first part of the project, which involved the use of microarray-based technology for the study of DNA methylation, we encountered several experimental problems. The method to enrichment of methylated-DNA was not efficient, probably due to the low quality of started DNA samples. Indeed, we had at our disposal only TRIzol preparations of the selected biopsies for the analysis and we observed that DNA extracted with this method is highly degradated and showed a high concentration of salts which interfered with followed analysis. Moreover, we observed an high signal of background noise after hybridization on microarray platform. Furthermore, signals resulted from each array were variable and it was difficult performing inter-array normalization.

The second part of the project is consisted of the validation of data obtained with microarray experiments with a different technology. We decided to use the *Reduced-Representation Bisulfite Sequencing* (RRBS) approach carried out with an Illumina sequencer. We observed a low correlation and concordance between bisulfite sequencing and hybridization on DNA microarray platform. It is known that methylated DNA enrichments methods, used in our microarray experiments, provided the best coverage of the whole genome and gene body regions, while RRBS was superior for recognize CpGs in CpG islands and promoters. Moreover, the differences of our results could be caused by experimental difficulties occurred in microarray experiments as previously described. Of note, next-generation sequencing (NGS), as RRBS approach, offers remedies to such microarray problems: (i) knowledge of genome annotation is helpful, but not required, while microarray design requires a priori knowledge of the genome or genomic features. This directly affects array effectiveness in cases of incomplete, incorrect, or outdated genome annotations; (ii) material is directly sequenced and not interrogated by hybridization to user defined sequences. This remove experimental bias and cross-hybridization issues from the analysis, that represent the major obstacle in microarray analysis. Since NGS offers single-nucleotide resolution one can monitor differences in as little as one nucleotide of the sequence; (iii) quantification of signal from sequence-based approaches is based on counting sequence tags rather than relative measures between samples: the result is unlimited fully-quantitative dynamic range of signal; (iv) because all next-generation platforms have the same data output it is hoped that the reproducibility of experimentation, and simplicity of bioinformatics analysis, will be

much improved over a variety of microarray platforms with unique oligonucleotide probes and hybridization conditions.

The analysis of sequencing data have revealed a set of DMRs between *PAX3-FOXO1* positive and *PAX3-FOXO1* negative RMS, many of them are associated to promotorial regions. The comparison with available datasets of gene expression on RMS biopsies and cell lines (Davicioni *et al.,* 2006; Whachtel *et al.,* 2004; De Pittà *et al*., 2006; Rapa *et al.,* 2012) demonstrated an inverse correlation between gene expression and promoter methylation status in 50-60% of analyzed genes. A subsequent gene ontology analysis allowed us to detect the enriched classes, thus, based on literature data, we selected some target genes for validation studies.

For *GADD45G* and *NELL1* genes we confirmed an epigenetic regulation by DNA methylation. Treatment with 5-aza-2'-deoxycytine determined a restoration of *GADD45G* and *NELL1* expression in *PAX3-FOXO1* positive cell lines. Moreover, a strong synergic effect was observed for *NELL1* using a combination of 5-aza-2'-deoxycytine and trichostatin A, suggesting a positive cross-talk between promoter methylation and histones acetylation in *NELL1* regulation. All these data was supported by a bisulfite sequencing of promotorial regions of *GADD45G* and *NELL1*, that demonstrated a hypermethylation in *PAX3-FOXO1* positive RMS cell lines versus *PAX3-FOXO1* negative ones. Moreover, we evaluated the gene expression level of *GADD45G* and *NELL1* in a new cohort of RMS samples and we demonstrated a statistically different expression of *GADD45G* and *NELL1* between *PAX3-FOXO1* positive RMS samples and *PAX3-FOXO1* negative ones, in both tumor biopsies and cell lines.

Growth arrest and DNA damage 45G (*GADD45G*) encodes a stress-responsive protein that is involved in DNA damage response and cell growth arrest through modulating a number of cellular proteins, including the proliferating cell nuclear antigen (PCNA), p21, Cdk1, cdc2/cyclin B1, p38 and c-Jun N-terminal kinase (JNK) (Ying *et al.*, 2005; Liebermann *et al.*, 2011; Cretu *et al.,* 2009; Vairapandi *et al.,* 2002). Moreover, *GADD45G* levels are remarkably downregulated in different types of solid tumors compared to their corresponding normal tissues (Zhang *et al.,* 2014). An aberrant inactivation, through proximal promoter methylation, plays an important role in esophageal squamous cell carcinoma (ESCC) carcinogenesis and this suggests a possible role as tumor suppressor (Guo *et al.,* 2013). *GADD45G* is thought to act as functional tumor suppressor also in gastric cardia adenocarcinoma (GCA). It is frequently inactivated, together with *GADD45A*, in a epigenetically way, in patients with GCA. Silencing of *GADD45A* and *GADD45G* is most

likely responsible for conferring a selective growth advantage during GCA evolution and outgrowth (Guo *et al.*, 2013).

The *NELL1* (*Neural epidermal growth factor-like 1*) gene encodes a protein kinase C-binding protein that contains six EGF-like domains and belongs to a new class of cell-signaling molecules controlling cell growth and differentiation (Matsuhashi *et al.*, 1995; Watanabe *et al.*, 1996; Kuroda and Tanizawa, 1999; Desai *et al.*, 2006). The precise role of *NELL1* in physiology and pathophysiology is not fully understood. Overexpression of *NELL1* increases osteoblast differentiation and reduces cell proliferation in transgenic mice, while downregulation of *NELL1* using *in vitro* approaches inhibits osteoblast differentiation and suggests that reduced levels of NELL1 protein leads to promote cell proliferation at the suture line (Zang *et al.*, 2002). It has also been reported that overexpression of *NELL1* promotes apoptosis in osteoblasts both *in vitro* and *in vivo* (Zhang *et al.*, 2003), and that this apoptotic activity may be associated with the *Fas* signaling pathway (Zhang *et al.*, 2006). In *NELL1* mutant mice, loss of *NELL1* expression was associated with reduced expression of genes encoding tumor necrosis factor receptor superfamily member 11b and extracellular matrix proteins (Desai *et al.*, 2006), which have also been implicated in human carcinogenesis (Ingber, 2002; Rowinsky, 2005). Furthermore, the high frequency (44%) of *NELL1* promoter hypermethylation in colon cancer suggests a potential role for *NELL1* inactivation in colon tumorigenesis (Mori *et al.*, 2006). *NELL1* has been found inactivated via promoter hypermethylation also in human esophageal adenocarcinoma (EAC) and is associated with a poor prognosis in early-stage EAC patients. Hypermethylation of *NELL1* could occur early in the genesis of EAC constituting a potentially useful early detection biomarker in this disease. (Jin *et al.*, 2007). Recently, it has been demonstrated that NELL1 protein expression is downregulated in human renal cell carcinoma (RCC), presumably by promoter hypermethylation, and that lack of NELL1 expression may contribute to RCC progression by altered regulation of cancer cell behaviour (Nakamura *et al.*, 2015). Moreover, up-regulation of *NELL1* in *PAX3-FOXO1* positive RMS cell lines was demonstrated also by previous microarray analysis performed on RMS cell lines, but a correlation with methylation was not investigated (Rapa *et al.*, 2012). Taken together, these findings suggest that *GADD45G* and *NELL1* function as a tumor suppressor gene in certain human cancers and that its expression is impaired by aberrant DNA methylation.

Herein, we provided the first evidence that *GADD45G* and *NELL1* may be involved also in RMS biology. Correlation analysis demonstrated as *GADD45G* and *NELL1* expression levels are able to discriminate in a strong statistically way not only *PAX3-FOXO1* positive and

negative RMS, but also metastatic and non-metastatic RMS. This is robustly supported by ROC analysis that showed as both these genes could be good predictors to distinguish *PAX3-FOXO1* positive and negative RMS. Furthermore, *GADD45G* and *NELL1* expression levels affected the progression free survival suggesting their association with a poor prognosis in RMS patients.

In conclusion, we observed a different methylation profile between *PAX3-FOXO1* positive and *PAX3-FOXO1* negative RMS, as well as among metastatic and non-metastatic RMS, using microarray platform. The RRBS sequencing has also demonstrated that *PAX3-FOXO1* positive and negative RMS have a different methylation pattern. With RRBS approach we found that *GADD45G* and *NELL1*, already described in the literature as tumor suppressors in a variety of cancers, are implicated in RMS biology. Moreover, high expression levels of *GADD45G* and *NELL1* are indicators of unfavorable prognosis. Our efforts are now focusing in the establish the biological role of *GADD45G* and *NELL1* in RMS, expanding the patients cohort and performing *in vitro* functional studies.

Thus, we demonstrated that *GADD45G* and *NELL1* could be novel potential biomarkers for diagnosis and prognosis of RMS patients. Futhermore, the DNA methylation changes in RMS could be interesting to apply new therapeutic strategies, probably in combination with histone deacetylase inhibitors drugs. Finally, a combined approach of standard chemotherapy with epigenetic drugs could be investigated to improve the RMS outcome.

BIBLIOGRAFY

Agrawal A, Murphy RF, Agrawal DK. (2007). DNA methylation in breast and colorectal cancers. *Mod Pathol*. **20(7)**:711-21.

Ahn EH, Mercado GE, Laé M, Ladanyi M. (2013). Identification of target genes of PAX3-FOXO1 in alveolar rhabdomyosarcoma. *Oncol Rep*. **30(2)**:968-78.

Akiyama Y1, Watkins N, Suzuki H, Jair KW, van Engeland M, Esteller M, Sakai H, Ren CY, Yuasa Y, Herman JG, Baylin SB. (2003). GATA-4 and GATA-5 transcription factor genes and potential downstream antitumor target genes are epigenetically silenced in colorectal and gastric cancer. *Mol Cell Biol*. **23(23)**:8429-39.

Antequera F, Bird A. (1993). CpG islands. *EXS.* **64**:169-85

Baylin SB, Ohm JE. (2006). Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction. *Nat Rev Cancer*. **6(2)**:107-16.

Baylin SB1, Herman JG. (2000). DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet*. **16(4)**:168-74.

Begum S, Emami N, Cheung A, Wilkins O, Der S, Hamel PA. (2005). Cell-type-specific regulation of distinct sets of gene targets by Pax3 and Pax3/FKHR. *Oncogene.* **24(11):**1860-72.

Bender J. (2004). Chromatin-based silencing mechanisms. *Curr Opin Plant Biol.*

Bennicelli JL, Edwards RH, Barr FG. (1996). Mechanism for transcriptional gain of function resulting from chromosomal translocation in alveolar rhabdomyosarcoma. *Proc Natl Acad Sci U S A*. **93(11)**:5455-9

Bernasconi M, Remppis A, Fredericks WJ, Rauscher FJ 3rd, Schäfer BW. (1996). Induction of apoptosis in rhabdomyosarcoma cells through down-regulation of PAX proteins. *Proc Natl Acad Sci U S A*. **93(23)**:13164-9

Bestor TH. (1988). Cloning of a mammalian DNA methyltransferase. *Gene*. **74(1)**:9-12.

Bestor TH. (2000). Gene silencing as a threat to the success of gene therapy. *J Clin Invest*.**105(4)**:409-11.

Bickle TA, Krüger DH. (1993). Biology of DNA restriction. *Microbiol Rev.*

Bird A, Taggart M, Frommer M, Miller OJ, Macleod D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell.* **40(1)**:91-9.

Bird A. (1992). The essentials of DNA methylation. *Cell.* **70(1)**:5-8.

Bird A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* **16(1)**:6-21.

Bird AP, Wolffe (1999)AP. Methylation-induced repression--belts, braces, and chromatin. *Cell.* **99(5)**:451-4.

Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**: 185–193.

Bostick M, Kim JK, Estève PO, Clark A, Pradhan S, Jacobsen SE. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science.* **317(5845)**:1760-4.

Breneman JC, Lyden E, Pappo AS, Link MP, Anderson JR, Parham DM, Qualman SJ, Wharam MD, Donaldson SS, Maurer HM, Meyer WH, Baker KS, Paidas CN, Crist WM. (2003). Prognostic factors and clinical outcomes in children and adolescents with metastatic rhabdomyosarcoma-a report from the Intergroup Rhabdomyosarcoma Study IV. *J Clin Oncol*.**1;21(1)**:78-84.

Bridge JA, Liu J, Weibolt V et al. (2000). Novel genomic imbalances in embryonal rhabdomyosarcoma revealed by comparative genomic hybridization and fluorescence in situ hybridization: an intergroup rhabdomyosarcoma study. *Genes Chromosomes Cancer* **27(4)**,337–344.

Brock MV1, Hooker CM, Ota-Machida E, Han Y, Guo M, Ames S, Glöckner S, Piantadosi S, Gabrielson E, Pridham G, Pelosky K, Belinsky SA, Yang SC, Baylin SB, Herman JG. (2008). DNA methylation markers and early recurrence in stage I lung cancer. *N Engl J Med*. **358(11)**:1118-28.

Brown R1, Strathdee G. (2002). Epigenomics and epigenetic therapy of cancer. *Trends Mol Med*. **8(4 Suppl)**:S43-8.

C. D. M. Fletcher, J. A. Bridge, P. Hogendoorn, and F. Mertens, (2013). WHO Classification of Tumours of Soft Tissue and Bone*, IARC Press*.

Chaib H, Nebbioso A, Prebet T, Castellano R, Garbit S, Restouin A, Vey N, Altucci L, Collette Y. (2012). Anti-leukemia activity of chaetocin via death receptor-dependent apoptosis and dual modulation of the histone methyl-transferase SUV39H1. *Leukemia.* **26(4)**:662-74.

Cheng JC1, Matsen CB, Gonzales FA, Ye W, Greer S, Marquez VE, Jones PA, Selker EU. (2003). Inhibition of DNA methylation and reactivation of silenced genes by zebularine. *J Natl Cancer Inst.* **95(5)**:399-409.

Clark SJ, Melki J. (2002). DNA methylation and gene silencing in cancer: which is the guilty party. *Oncogene.* **21(35)**:5380-7.

Clark SJ1, Harrison J, Paul CL, Frommer M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Res*. **22(15)**:2990-7.

Colella S, Shen L, Baggerly KA, Issa JP, Krahe R. (2003). Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. *Biotechniques.* **35(1)**:146-50.

Cooper DN. (1983). Eukaryotic DNA methylation. *Hum Genet*. **64(4)**:315-33.

Cosma MP, Tanaka T, Nasmyth K. (1999). Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell.* **97(3)**:299-311

Costello JF, Frühwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X, Wright FA, Feramisco JD, Peltomäki P, Lang JC, Schuller DE, Yu L, Bloomfield CD, Caligiuri MA, Yates A, Nishikawa R, Su Huang H, Petrelli NJ, Zhang X, O'Dorisio MS, Held WA, Cavenee WK, Plass C. (2000). Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat Genet*. **24(2)**:132-8.

Costello JF, Plass C. (2001). Methylation matters. *J Med Genet.* **38(5)**:285-303.

Cretu A, Sha X, Tront J, Hoffman B and Liebermann DA.(2009) Stress sensor Gadd45 genes as therapeutic targets in cancer. *Cancer therapy*.; **7**:268-276.

Crist W, Gehan EA, Ragab AH, Dickman PS, Donaldson SS, Fryer C, Hammond D, Hays DM, Herrmann J, Heyn R, et al. (1995). The Third Intergroup Rhabdomyosarcoma Study. *J Clin Oncol.* **13(3)**:610-30.

Crist W, Gehan EA, Ragab AH, et al. (1995). The third Intergroup Rhabdomyosarcoma Study. *J Clin Oncol.* **13**:610–630.

Cross SH1, Bird AP. (1995). CpG islands and genes. *Curr Opin Genet Dev*. **5(3)**:309-14.

Dagher R, Helman L. (1999). Rhabdomyosarcoma: an overview.*Oncologist*.;**4(1)**:34-44

Daskalakis M, Nguyen TT, Nguyen C, Guldberg P, Köhler G, Wijermans P, Jones PA, Lübbert M. (2002). Demethylation of a hypermethylated P15/INK4B gene in patients with myelodysplastic syndrome by 5-Aza-2'-deoxycytidine (decitabine) treatment. *Blood*. **100(8)**:2957-64.

Davicioni E, Anderson MJ, Finckenstein FG et al. (2009). Molecular classification of rhabdomyosarcoma – genotypic and phenotypic determinants of diagnosis: a report from the Children's Oncology Group. *Am. J. Pathol*. **174(2)**,550–564

Davicioni E, Finckenstein FG, Shahbazian V, Buckley JD, Triche TJ, Anderson MJ. (2006) Identification of a PAX-FKHR gene expression signature that defines molecular classes and determines the prognosis of alveolar rhabdomyosarcomas. *Cancer Res.***66(14)**:6936-46.

De Pittà C1, Tombolan L, Albiero G, Sartori F, Romualdi C, Jurman G, Carli M, Furlanello C, Lanfranchi G, Rosolen A. (2006). Gene expression profiling identifies potential relevant genes in alveolar rhabdomyosarcoma pathogenesis and discriminates PAX3-FKHR positive and negative tumors. *Int J Cancer*. **118(11)**:2772-81.

DeBaun MR, Tucker MA. (1998). Risk of cancer during the first four years of life in children from The Beckwith-Wiedemann Syndrome Registry. *J Pediatr.* **132(3 Pt 1)**:398-400.

DeBlacam C, Byrne C, Hughes E, McIlroy M, Bane F, Hill AD, Young LS.(2011). HOXC11-SRC-1 regulation of S100beta in cutaneous melanoma: new targets for the kinase inhibitor dasatinib. Br J Cancer. **105(1)**:118-23.

Dehner LP, Jarzembowski JA, Hill DA. (2012) . Embryonal rhabdomyosarcoma of the uterine cervix: a report of 14 cases and a discussion of its unusual clinicopathological associations. *Mod Pathol.* **25(4)**:602-14.

Deng J, Liang H, Ying G, Dong Q, Zhang L, Yu J, Fan D, Hao X. (2014) Clinical significance of the methylated cytosine-phosphate-guanine sites of protocadherin-10 promoter for evaluating the prognosis of gastric cancer. *J Am Coll Surg*. **219(5)**:904-13.

Desai J, Shannon ME, Johnson MD, Ruff DW, Hughes LA, Kerley MK *et al*. (2006). Nell1-deficient mice have reduced expression of extracellular matrix proteins causing cranial and vertebral defects. *Hum Mol Genet* **15**: 1329–1341.

Doerfler W. (2006). The almost-forgotten fifth nucleotide in DNA: an introduction. *Curr Top Microbiol Immunol.* **301:3**-18.

Doros L, Yang J, Dehner L, Rossi CT, Skiver K, Jarzembowski JA, Messinger Y, Schultz KA, Williams G, André N, Hill DA. (2012). DICER1 mutations in embryonal rhabdomyosarcomas from children with and without familial PPB-tumor predisposition syndrome. *Pediatr Blood Cancer*. **59(3)**:558-60.

Echizen K, Nakada M, Hayashi T, Sabit H, Furuta T, Nakai M, Koyama-Nasu R, Nishimura Y, Taniue K, Morishita Y, Hirano S, Terai K, Todo T, Ino Y, Mukasa A, Takayanagi S, Ohtani R, Saito N, Akiyama T. (2014) PCDH10 is required for the tumorigenicity of glioblastoma cells. *Biochem Biophys Res Commun.***444(1)**:13-8

Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, Haefliger C, Horton R, Howe K, Jackson DK, Kunde J, Koenig C, Liddle J, Niblett D, Otto T, Pettett R, Seemann S, Thompson C, West T, Rogers J, Olek A, Berlin K, Beck S. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet*. **38(12)**:1378-85.

Eden A, Gaudet F, Waghmare A, Jaenisch R. (2003). Chromosomal instability and tumors promoted by DNA hypomethylation. *Science.* **300(5618)**:455.

Egas-Bejar D, Huh WW. (2014). Rhabdomyosarcoma in adolescent and young adult patients: current perspectives. *Adolesc Health Med Ther.* **17;5**:115-25.

Egger G, Liang G, Aparicio A, Jones PA. (2004). Epigenetics in human disease and prospects for epigenetic therapy *Nature*. **429(6990)**:457-63.

Ehrlich M. (2002). DNA methylation in cancer: too much, but also too little. *Oncogene*. **21(35)**:5400-13.

Esteller M. (2008). Epigenetics in cancer. *N Engl J Med.* **358(11)**:1148-59.

Esteller M1, Garcia-Foncillas J, Andion E, Goodman SN, Hidalgo OF, Vanaclocha V, Baylin SB, Herman JG. (2000). Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Engl J Med*. **343(19):**1350-4.

Fandy TE. (2009). Development of DNA methyltransferase inhibitors for the treatment of neoplastic diseases. *Curr Med Chem*. **16(17)**:2075-85

Fatemi M, Hermann A, Pradhan S, Jeltsch A. (2001). The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. *J Mol Biol.* **309(5)**:1189-99.

Feinberg AP, Ohlsson R, Henikoff S. (2006). The epigenetic progenitor origin of human cancer. *Nat Rev Genet*. **7(1)**:21-33.

Feinberg AP, Vogelstein B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*. **301(5895)**:89-92

Filion GJ, Zhenilo S, Salozhin S, Yamada D, Prokhortchouk E, Defossez PA. (2006). A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Mol Cell Biol.* **26(1)**:169-81.

Foulks JM, Parnell KM, Nix RN, Chau S, Swierczek K, Saunders M, Wright K, Hendrickson TF, Ho KK, McCullar MV, Kanner SB. (2012). Epigenetic drug discovery: targeting DNA methyltransferases. *J Biomol Screen*. **17(1)**:2-17.

Fraga MF, Herranz M, Espada J, Ballestar E, Paz MF, Ropero S, Erkek E, Bozdogan O, Peinado H, Niveleau A, Mao JH, Balmain A, Cano A, Esteller M. (2004). A mouse skin multistage carcinogenesis model reflects the aberrant DNA methylation patterns of human tumors. *Cancer Res*. **64(16)**:5527-34.

Frommer M1, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A*. **89(5)**:1827-31.

Gallego Melcón S, Sánchez de Toledo Codina J. (2007). Molecular biology of rhabdomyosarcoma. *Clin Transl Oncol.* **9(7)**:415-9.

Gardiner-Garden M1, Frommer M. (1987). CpG islands in vertebrate genomes. *J Mol Biol*. **196(2)**:261-82.

Genome-wide profiling of DNA methylation reveals transposon targets of CHROMOMETHYLASE3. *Curr Biol*. **12(1)**:65-8.

Gitan RS, Shi H, Chen CM, Yan PS, Huang TH. (2002). Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis. *Genome Res*. **12(1)**:158-64

Glass T, Cochrane DD, Rassekh SR, Goddard K, Hukin J. (2014) Growing teratoma syndrome in intracranial non-germinomatous germ cell tumors (iNGGCTs): a risk for secondary malignant transformation—a report of two cases. *Childs Nerv Syst*; **30(5)**:953-7.

Goffin J, Eisenhauer E. (2002). DNA methyltransferase inhibitors-state of the art. *Ann Oncol.* **13(11)**:1699-716.

Goyal R1, Reinhardt R, Jeltsch A. (2006). Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. *Nucleic Acids Res*. **34(4)**:1182-8.

Grigg G, Clark S. (1994). Sequencing 5-methylcytosine residues in genomic DNA. *Bioessays*. **6(6)**:431-6.

Gripp KW, Lin AE, Stabley DL, Nicholson L, Scott CI Jr, Doyle D, Aoki Y, Matsubara Y, Zackai EH, Lapunzina P, Gonzalez-Meneses A, Holbrook J, Agresta CA, Gonzalez IL, Sol-Church K (2006). . HRAS mutation analysis in Costello syndrome: genotype and phenotype correlation. *Am J Med Genet A* **140(1)**:1-7.

Gu H, Bock C, Mikkelsen TS, et al. (2010). Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nat Methods* **7**:133–6.

Guo W, Dong Z, Guo Y, Chen Z, Kuang G, Yang Z. (2013). Methylation-mediated repression of GADD45A and GADD45G expression in gastric cardia adenocarcinoma. *Int J Cancer* **133(9)**:2043-53

Guo W, Zhu T, Dong Z, Cui L, Zhang M, Kuang G. (2013). Decreased expression and aberrant methylation of Gadd45G is associated with tumor progression and poor prognosis in esophageal squamous cell carcinoma. *Clin Exp Metastasis*.**30(8):**977-92

Harada H, Miyamoto K, Yamashita Y, Taniyama K, Mihara K, Nishimura M, Okada M .(2015). Prognostic signature of protocadherin 10 methylation in curatively resected pathological stage I non-small-cell lung cancer. *Cancer Med*. **4(10)**:1536-46.

Haruki S, Imoto I, Kozaki K, Matsui T, Kawachi H, Komatsu S, Muramatsu T, Shimada Y, Kawano T, Inazawa J.(2010). Frequent silencing of protocadherin 17, a candidate tumour suppressor for esophageal squamous cell carcinoma. *Carcinogenesis.* **31(6)**:1027-36

Hashimoto K1, Kokubun S, Itoi E, Roach HI. (2007). Improved quantification of DNA methylation using methylation-sensitive restriction enzymes and real-time PCR. *Epigenetics*. **2(2)**:86-91.

Hasle H. (2009). Malignant diseases in Noonan syndrome and related disorders. *Horm Res.* **2**:8-14.

Hatada I, Hayashizaki Y, Hirotsune S, Komatsubara H, Mukai T. (1991). A genomic scanning method for higher organisms using restriction sites as landmarks. *Proc Natl Acad Sci U S A*. **88(21)**:9523-7.

Hegi ME, Diserens AC, Godard S, Dietrich PY, Regli L, Ostermann S, Otten P, Van Melle G, de Tribolet N, Stupp R. (2004). Clinical trial substantiates the predictive value of O-6-methylguanine-DNA methyltransferase promoter methylation in glioblastoma patients treated with temozolomide. *Clin Cancer Res.* **10(6)**:1871-4.

Helman LJ, Meltzer P. (2003). Mechanisms of sarcoma development. *Nat. Rev. Cancer* **3(9)**,685–694.

Hendrich B1, Bird A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol.* **18(11)**:6538-47.

Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB. (1996). Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci U S A*. **93(18)**:9821-6.

Herman JG1, Baylin SB. (2003). Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med.* **349(21)**:2042-54.

Hermann A1, Schmitt S, Jeltsch A. (2003). The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity. *J Biol Chem*. **278(34)**:31717-21.

Hingorani P, Missiaglia E, Shipley J, Anderson JR, Triche TJ, Delorenzi M, Gastier-Foster J, Wing M, Hawkins DS, Skapek SX. (2015). Clinical Application of Prognostic Gene Expression Signature in Fusion Gene-Negative Rhabdomyosarcoma: A Report from the Children's Oncology Group. *Clin Cancer Res*.**21(20)**:4733-9

Hirohashi S. (1998). Inactivation of the E-cadherin-mediated cell adhesion system in human cancers. *Am J Pathol.* **53(2)**:333-9.

Holliday R, Pugh JE. (1975). DNA modification mechanisms and gene activity during development. *Science*. **187(4173)**:226-32.

Ingber DE. (2002). Cancer as a disease of epithelial-mesenchymal interactions and extracellular matrix regulation. *Differentiation* **70**: 547–560.

Ji J, Eng C, Hemminki K. (2007). Familial risk for soft tissue tumors: a nation-wide epidemiological study from Sweden. *J Cancer Res Clin Oncol*. **134(5)**:617-24.

Jin Z, Mori Y, Yang J, Sato F, Ito T, Cheng Y, Paun B, Hamilton JP, Kan T, Olaru A, David S, Agarwal R, Abraham JM, Beer D, Montgomery E, Meltzer SJ. (2007) Hypermethylation of the nel-like 1 gene is a common and early event and is associated with poor prognosis in early-stage esophageal adenocarcinoma. *Oncogene.* **26(43)**:6332-40.

Jones PA, Laird PW. (1999). Cancer epigenetics comes of age. *Nat Genet.* **21(2)**:163-7.

Jones PA. (1999). The DNA methylation paradox. *Trends Genet*. **15(1)**:34-7.

Jones PA1, Baylin SB. (2002). The fundamental role of epigenetic events in cancer. *Nat Rev Genet.* **3(6)**:415-28.

Kai-Hung Wang, Cuei-Jyuan Lin, Chou-Jen Liu, Dai-Wei Liu, Rui-Lan Huang,Dah-Ching Ding2,, Ching-Feng Weng & Tang-Yuan Chu. (2014). Global methylation silencing of clustered proto-cadherin genes in cervical cancer: serving as diagnostic markers comparable to HPV. *Cancer Medicine*. **4(1)**:43–55

Kareta MS, Botello ZM, Ennis JJ, Chou C, Chédin F. (2006). Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. *J Biol Chem.* **281(36)**:25893-902.

Keshet I, Schlesinger Y, Farkash S, Rand E, Hecht M, Segal E, Pikarski E, Young RA, Niveleau A, Cedar H, Simon I. (2006). Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nat Genet*. **38(2)**:149-53.

Khulan B, Thompson RF, Ye K, Fazzari MJ, Suzuki M, Stasiek E, Figueroa ME, Glass JL, Chen Q, Montagna C, Hatchwell E, Selzer RR, Richmond TA, Green RD, Melnick A, Greally JM. (2006). Comparative isoschizomer profiling of cytosine methylation: the HELP assay. *Genome Res.* **16(8)**:1046-55.

Kikuchi K1, Tsuchiya K, Otabe O, Gotoh T, Tamura S, Katsumi Y, Yagyu S, Tsubai-Shimizu S, Miyachi M, Iehara T, Hosoi H. (2007). Effects of PAX3-FKHR on malignant phenotypes in alveolar rhabdomyosarcoma. Biochem Biophys Res Commun.**365(3)**:568-74.

Kim JK, Samaranayake M, Pradhan S. (2009). Epigenetic mechanisms in mammals. *Cell Mol Life Sci.* **66(4)**:596-612.

Kloten V, Becker B, Winner K, Schrauder MG, Fasching PA, Anzeneder T, Veeck J, Hartmann A, Knüchel R, Dahl E. (2013). Promoter hypermethylation of the tumor-suppressor genes ITIH5, DKK3, and RASSF1A as novel biomarkers for blood-based breast cancer screening. *Breast Cancer Res*. **15(1)**:R4.

Kratz CP, Rapisuwon S, Reed H, Hasle H, Rosenberg PS (2011). Cancer in Noonan, Costello, cardiofaciocutaneous and LEOPARD syndromes. *Am J Med Genet C Semin Med Genet.* **157C(2):**83-9.

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

Kuroda S, Tanizawa K. (1999). Involvement of epidermal growth factor-like domain of NELL proteins in the novel protein-protein interaction with protein kinase C. *Biochem Biophys Res Commun* **265**: 752–757.

Laird PW. (2010). Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet*. **11(3)**:191-203.

Lam PY, Sublett JE, Hollenbach AD, Roussel MF. (1999). The oncogenic potential of the Pax3-FKHR fusion protein requires the Pax3 homeodomain recognition helix but not the Pax3 paired-box DNA binding domain. *Mol Cell Biol.* **19(1)**:594-601

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, et al. (2001). Initial sequencing and analysis of the human genome. *International Human Genome Sequencing Consortium.Nature.* **409(6822)**:860-921.

Lauster R1, Trautner TA, Noyer-Weidner M. (1989). Cytosine-specific type II DNA methyltransferases. A conserved enzyme core with variable target-recognizing domains. *J Mol Biol.* **206(2)**:305-12.

Li FP, Fraumeni JF Jr (1969) . Rhabdomyosarcoma in children: epidemiologic study and identification of a familial cancer syndrome. *J Natl Cancer Inst.* **43(6)**:1365-73.

Li LC, Dahiya R. (2002). MethPrimer: designing primers for methylation PCRs. *Bioinformatics.* **18(11)**:1427-31.

Liebermann DA, Tront JS, Sha X, Mukherjee K, Mohamed-Hadley A and Hoffman B.(2011) Gadd45 stress sensors in malignancy and leukemia. *Critical reviews in oncogenesis*; **16**:129-140.

Lin YL, Ma JH, Luo XL, Guan TY, Li ZG. (2013) Clinical significance of protocadherin-8 (PCDH8) promoter methylation in bladder cancer. *J Int Med Res*.**41(1)**:48-54

Lin YL, Wang YL, Ma JG, Li WP. (2014) Clinical significance of protocadherin 8 (PCDH8) promoter methylation in non-muscle invasive bladder cancer. *J Exp Clin Cancer Res* **33**:68

Linardic CM. (2008). PAX3-FOXO1 fusion gene in rhabdomyosarcoma. *Cancer Lett.* **270(1):**10-8.

Liu YJ, Zhu Y, Yuan HX, Zhang JP, Guo JM, Lin ZM.(2015) Overexpression of HOXC11 homeobox gene in clear cell renal cell carcinoma induces cellular proliferation and is associated with poor prognosis.Tumour Biol.**36(4)**:2821-9.

Lizard-Nacol S, Mugneret F, Volk C, Turc-Carel C, Favrot M, Philip T. (1987). Translocation (2;13)(q37;q14) in alveolar rhabdomyosarcoma: a new case. Cancer Genet Cytogenet. **25(2):**373-4.

Lübbert M. (2000). DNA methylation inhibitors in the treatment of leukemias, myelodysplastic syndromes and hemoglobinopathies: clinical results and possible mechanisms of action. *Curr Top Microbiol Immunol*.**249**:135-64.

Marshall AD, Grosveld GC. (2012). Alveolar rhabdomyosarcoma - The molecular drivers of PAX3/7-FOXO1-induced tumorigenesis. *Skelet Muscle*. **2(1)**:25.

Matsuhashi S, Noji S, Koyama E, Myokai F, Ohuchi H, Taniguchi S et al. (1995). New gene, nel, encoding a M(r) 93K protein with EGF-like repeats is strongly expressed in neural tissues of early stage chick embryos. *Dev Dyn* **203**:212–222.

Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res*. **33(18)**:5868-77.

Mercado GE1, Xia SJ, Zhang C, Ahn EH, Gustafson DM, Laé M, Ladanyi M, Barr FG. (2008)Identification of PAX3-FKHR-regulated genes differentially expressed between alveolar and embryonal rhabdomyosarcoma: focus on MYCN as a biologically relevant target. *Genes Chromosomes Cancer.* **47(6)**:510-20.

Meyer WH, Spunt SL. (2004). Soft tissue sarcomas of childhood. *Cancer Treat Rev*. **30(3)**:269-80.

Meza JL, Anderson J, Pappo AS, Meyer WH. (2006). Analysis of prognostic factors in patients with nonmetastatic rhabdomyosarcoma treated on intergroup rhabdomyosarcoma studies III and IV: the Children's Oncology Group. *J Clin Oncol*. **24(24)**:3844-51

Michie AM, McCaig AM, Nakagawa R, Vukovic M. (2010). Death-associated protein kinase (DAPK) and signal transduction: regulation in cancer. *FEBS J.* **277(1)**:74-80.

Missiaglia E, Williamson D, Chisholm J, Wirapati P, Pierron G, Petel F, Concordet JP, Thway K, Oberlin O, Pritchard-Jones K, Delattre O, Delorenzi M, Shipley J. (2012) PAX3/FOXO1 fusion gene status is the key prognostic molecular marker in rhabdomyosarcoma and significantly improves current risk stratification. *J Clin Oncol*. **30(14)**:1670-7

Mori Y, Cai K, Cheng Y, Wang S, Paun B, Hamilton JP *et al*. (2006). A genome-wide search identifies epigenetic silencing of somatostatin, tachykinin-1, and 5 other genes in colon cancer. *Gastroenterology* **131**: 797–808

Moschovi M, Touliatou V, Papadopoulou A, Mayakou MA, Nikolaidou-Karpathiou P, Kitsiou-Tzeli S. (2007). Rhabdomyosarcoma in a patient with Noonan syndrome phenotype and review of the literature. *J Pediatr Hematol Oncol.* **29(5)**:341-4.

Naini S, Etheridge KT, Adam SJ, Qualman SJ, Bentley RC, Counter CM, Linardic CM. (2008). Defining the cooperative genetic changes that temporally drive alveolar rhabdomyosarcoma. *Cancer Res.* **68(23)**:9583-8.

Nakamura R, Oyama T, Tajiri R, Mizokami A, Namiki M, Nakamoto M, Ooi A. (2015) Expression and regulatory effects on cancer cell behavior of NELL1 and NELL2 in human renal cell carcinoma. *Cancer Sci.* **106(5)**:656-64.

Neri F1, Krepelova A, Incarnato D, Maldotti M, Parlato C, Galvagni F, Matarese F, Stunnenberg HG, Oliviero S. (2013). Dnmt3L antagonizes DNA methylation at bivalent promoters and favors DNA methylation at gene bodies in ESCs. *Cell*. **155(1)**:121-34.

Oberlin O, Rey A, Lyden E, et al. (2008). Prognostic factors in metastatic rhabdomyosarcomas: Results of a pooled analysis from united states and European cooperative groups. *J Clin Oncol*. **26**:2384–2389.

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

Okano M, Xie S, Li E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet.* **19(3)**:219-20.

Parham DM, Ellison DA. (2006). Rabdomyosarcomas in adults and children: an update. *Arch Pathol Lab Med*. **130(10)**:1454-65

Parham DM. (2001). Pathologic classification of rhabdomyosarcomas and correlations with molecular studies. *Mod Pathol.* **14(5)**:506-14.

Park MT, Lee SJ. (2003). Cell cycle and cancer. *J Biochem Mol Biol.* **36(1)**:60-5.

Paulino AC, Okcu MF.(2008) Rhabdomyosarcoma. *Curr Probl Cancer.* **32(1)**:7-34.

Pfaffl MW, Horgan GW, Dempfle L. (2002). Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.* **30(9)**:e36.

Pfaffl MW. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. **29(9)**:e45.

Prokhortchouk A1, Hendrich B, Jørgensen H, Ruzov A, Wilm M, Georgiev G, Bird A, Prokhortchouk E. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev*. **15(13)**:1613-8.

Prokhortchouk E1, Hendrich B. (2002). Methyl-CpG binding proteins and cancer: are MeCpGs more important than MBDs. *Oncogene.* **21(35)**:5394-9.

Rapa E, Hill SK, Morten KJ, Potter M, Mitchell C. (2012) The over-expression of cell migratory genes in alveolar rhabdomyosarcoma could contribute to metastatic spread. *Clin Exp Metastasis*. **29(5)**:419-29.

Riggs AD. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet.***14(1)**:9-25.

RIOPELLE JL, THERIAULT JP.(1956). An unknown type of soft part sarcoma: alveolar rhabdomyosarcoma. *Ann Anat Pathol.* **1(1):**88-111.

Risso D, Massa MS, Chiogna M, Romualdi C (2009) A modified LOESS normalization applied to micro-RNA arrays: a comparative evaluation. *Bioinformatics* **25**: 2685–2691.

Robertson KD1. (2001). DNA methylation, methyltransferases, and cancer. *Oncogene.* **20(24)**:3139-55.

Robson EJ, He SJ, Eccles MR. (2006). A PANorama of PAX genes in cancer and development. *Nat Rev Cancer.* **6(1)**:52-62.

Rodriguez J1, Frigola J, Vendrell E, Risques RA, Fraga MF, Morales C, Moreno V, Esteller M, Capellà G, Ribas M, Peinado MA. (2006). Chromosomal instability correlates with genome-wide DNA demethylation in human primary colorectal cancers. *Cancer Res.* **66(17)**:8462-9468.

Rouillard JM, Erson AE, Kuick R, Asakawa J, Wimmer K, Muleris M, Petty EM, Hanash S. (2001). Virtual genome scan: a tool for restriction landmark-based scanning of the human genome. *Genome Res.* **11(8)**:1453-9.

Rowinsky EK. (2005). Targeted induction of apoptosis in cancer management: the emerging role of tumor necrosis factor-related apoptosis-inducing ligand receptor activating agents. *J Clin Oncol* **23**: 9394–9407

Rüter B, Wijermans PW, Lübbert M. (2004). DNA methylation as a therapeutic target in hematologic disorders: recent results in older patients with myelodysplasia and acute myeloid leukemia. *Int J Hematol.* **80(2)**:128-35.

Samuel DP, Tsokos M, DeBaun MR.(1999). . Hemihypertrophy and a poorly differentiated embryonal rhabdomyosarcoma of the pelvis. *Med Pediatr Oncol.* **32(1)**:38-43.

Scheidler S1, Fredericks WJ, Rauscher FJ 3rd, Barr FG, Vogt PK.(1996). The hybrid PAX3-FKHR fusion protein of alveolar rhabdomyosarcoma transforms fibroblasts in culture. *Proc Natl Acad Sci U S A.***93(18)**:9805-9.

Schmidt B1, Liebenberg V, Dietrich D, Schlegel T, Kneip C, Seegebarth A, Flemming N, Seemann S, Distler J, Lewin J, Tetzner R, Weickmann S, Wille U, Liloglou T, Raji O, Walshaw M, Fleischhacker M, Witt C, Field JK. (2010). SHOX2 DNA methylation is a biomarker for the diagnosis of lung cancer based on bronchial aspirates. *BMC Cancer.* **10**:600.

Schumacher A, Kapranov P, Kaminsky Z, Flanagan J, Assadzadeh A, Yau P, Virtanen C, Winegarden N, Cheng J, Gingeras T, Petronis A. (2006). Microarray-based DNA methylation profiling: technology and applications. *Nucleic Acids Res.* **34(2)**:528-42.

Singal R, Ginder GD. (1999). DNA methylation. *Blood.* **93(12)**:4059-70.

Singal R, Wang SZ, Sargent T, Zhu SZ, Ginder GD. (2002). Methylation of promoter proximal-transcribed sequences of an embryonic globin gene inhibits transcription in primary erythroid cells and promotes formation of a cell type-specific methyl cytosine binding complex. *J Biol Chem.* **277(3)**:1897-905.

Smith AC, Choufani S, Ferreira JC, Weksberg R. (2007). Growth regulation, imprinted genes, and chromosome 11p15.5. *Pediatr. Res.* **61**(5 Pt 2),43R–47R.

Smith ZD, Gu H, Bock C, Gnirke A, Meissner A. (2009). High-throughput bisulfite sequencing in mammalian genomes. *Methods.* **48(3)**:226-32.

Song L1, Li Y2. (2015). A Specific Circulating Biomarker for Colorectal Cancer. *Adv Clin Chem.* **72:**171-204.

Song SH, Han SW, Bang YJ. (2011). Epigenetic-based therapies in cancer: progress to date. *Drugs.* **71(18)**:2391-403.

Stout AP. (1946). Rhabdomyosarcoma of the Skeletal Muscles. *Ann Surg*. **123(3)**:447-72.

Sugimura T1, Ushijima T. (2000). Genetic and epigenetic alterations in carcinogenesis. *Mutat Res.* **462(2-3)**:235-46.

Sultan, I. Qaddoumi, S. Yaser, C. Rodriguez-Galindo, and A. Ferrari. (2009). Comparing adult and pediatric rhabdomyosarcoma in the surveillance, epidemiology and end results program, 1973 to 2005: an analysis of 2,600 patients. *Journal of Clinical Oncology.* **27(20)**, 3391–3397.

Szyf M. (2003). Targeting DNA methylation in cancer. *Ageing Res Rev.* **2(3)**:299-328.

Takeshi Yagi .(2008). Clustered protocadherin family. *Develop. Growth Differ*. **50**, S131

Tarasova GV, Nayakshina TN, Degtyarev SK. (2008). Substrate specificity of new methyl-directed DNA endonuclease GlaI. *BMC Mol Biol*. **9**:7.

Tate PH, Bird AP. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. *Curr Opin Genet Dev*.

Thiagalingam S, Cheng KH, Lee HJ, Mineva N, Thiagalingam A, Ponte JF. (2003). Histone deacetylases: unique players in shaping the epigenetic histone code. *Ann N Y Acad Sci*. **983**:84-100.

Tombolan L, Zampini M, Casara S, Boldrin E, Zin A, Bisogno G, Rosolen A, De Pittà C, Lanfranchi G. (2015). MicroRNA-27a Contributes to Rhabdomyosarcoma Cell Proliferation by Suppressing RARA and RXRA. *PLoS One.* **10(4)**:e0125171.

Tompa R, McCallum CM, Delrow J, Henikoff JG, van Steensel B, Henikoff S. (2002).

Tost J, Herman JG. (2009). Welcome to Epigenomics. *Epigenomics.* **1(1)**:1-3.

Treat B. Johnson, Robert D. Coghill. (1925). researches on pyrimidines. c111. the discovery of 5-methyl-cytosine in tuberculinic acid, the nucleic acid of the tubercle bacillus1. *J. Am. Chem. Soc*. **47 (11)**, pp 2838–2844

Vairapandi M, Balliet AG, Hoffman B and Liebermann DA.(2002) GADD45b and GADD45g are cdc2/cyclinB1 kinase inhibitors with a role in S and G2/M cell cycle checkpoiinduced by genotoxic stress. *Journal of cellular physiology.* **192**:327-338.

Vaissière T1, Sawan C, Herceg Z. (2008). Epigenetic interplay between histone modifications and DNA methylation in gene silencing Mutation Research **659:** 40–48

Veeck J, Esteller M. Breast cancer epigenetics: from DNA methylation to microRNAs. *J Mammary Gland Biol Neoplasia*. **5(1)**:5-17.

Visser M, Sijmons C, Bras J et al. (1997). Allelotype of pediatric rhabdomyosarcoma. *Oncogene **15(11)***,1309–1314.

Wachtel M, Dettling M, Koscielniak E, Stegmaier S, Treuner J, Simon-Klingenstein K, Bühlmann P, Niggli FK, Schäfer BW. (2004). Gene expression signatures identify rhabdomyosarcoma subtypes and detect a novel t(2;2)(q35;p23) translocation fusing PAX3 to NCOA1. *Cancer Res*. **64(16)**:5539-45.

Wang L, Xie PG, Lin YL, Ma JG, Li WP. (2014). Aberrant methylation of PCDH10 predicts worse biochemical recurrence-free survival in patients with prostate cancer after radical prostatectomy. *Med Sci Monit*. **20**:1363-8.

Watanabe TK, Katagiri T, Suzuki M, Shimizu F, Fujiwara T, Kanemoto N *et al*. (1996). Cloning and characterization of two novel human cDNAs (NELL1 and NELL2) encoding proteins with six EGF-like repeats. *Genomics* **38**: 273–276.

Watt F, Molloy PL. (1988)**.** Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev* **2(9)**:1136-43

Weinhaeusel A, Thiele S, Hofner M, Hiort O, Noehammer C. (2008). PCR-based analysis of differentially methylated regions of GNAS enables convenient diagnostic testing of pseudohypoparathyroidism type Ib. *Clin Chem*. **54(9)**:1537-45.

Wilson IM, Davies JJ, Weber M, Brown CJ, Alvarez CE, MacAulay C, Schübeler D, Lam WL. (2006). Epigenomics: mapping the methylome. *Cell Cycle*. **5(2)**:155-8.

Wu Ct, Morris JR. (2001). Genes, genetics, and epigenetics: a correspondence. *Science*. **293(5532)**:1103–1105

Wu, Q. and Maniatis, T. (1999). A striking organization of a large family of human neural cadherin-like cell adhesion genes. *Cell* **97**, 779-790.

Yang L, Takimoto T, Fujimoto J. (2014). Prognostic model for predicting overall survival in children and adolescents with rhabdomyosarcoma. *BMC Cancer*. **14**:654

Ying J, Srivastava G, Hsieh WS, Gao Z, Murray P, Liao SK, Ambinder R and Tao Q. (2005) The stress-responsive gene GADD45G is a functional tumor suppressor, with its response to environmental stresses frequently disrupted epigenetically in multiple tumors. *Clinical cancer research : an official journal of the American Association for Cancer Research*. **11**:6442-6449.

Zhang L, Yang Z and Liu Y.(2014;) GADD45 proteins: roles in cellular senescence and tumor development. *Experimental biology and medicine*. **239**:773-778.

Zhang X, Carpenter D, Bokui N, Soo C, Miao S, Truong T *et al*. (2003). Overexpression of Nell-1, a craniosynostosis-associated gene, induces apoptosis in osteoblasts during craniofacial development. *J Bone Miner Res* **18**: 2126–2134.

Zhang X, Cowan CM, Jiang X, Soo C, Miao S, Carpenter D *et al*. (2006). Nell-1 induces acrania-like cranioskeletal deformities during mouse embryonic development. *Lab Invest* **86**: 633–644.

Zhang X, Kuroda S, Carpenter D, Nishimura I, Soo C, Moats R *et al*. (2002). Craniosynostosis in transgenic mice overexpressing Nell-1. *J Clin Invest* **110**: 861–870.

RINGRAZIAMENTI

Ringrazio il Prof. Giuseppe Basso per avermi dato l'opportunità di svolgere questo lavoro di tesi, che mi ha permesso di acquisire nuove importanti competenze scientifiche e una maggiore autonomia nella gestione di un progetto di ricerca.

Un grazie particolare va al Dott.ssa Angelica Zin per la sua collaborazione e il valido supporto e alla Dott.ssa Lucia Tombolan per aver condiviso con me gioie e dolori di questo progetto, per i consigli, la supervisione e la pazienza durante la stesura della tesi.

Ringrazio calorosamente tutte le mie carissime colleghe Marica, Lara, Elena, Elisa, Simona e Federica, che mi hanno permesso di trascorrere dei momenti bellissimi e mi sono state accanto nei momenti di tristezza e di pazzia.

Un ringraziamento speciale va al Dott. Paolo Bonvini, per tutti gli insegnamenti che ha saputo darmi, per aver condiviso con me la passione per la ricerca o per meglio citarlo "il sacro fuoco della scienza", per essere stato un grande mentore, ma soprattutto per essere stato ed essere tuttora una persona su cui posso sempre contare.

Grazie infinitamente alla mia famiglia per il forte sostegno e la stima che mi hanno saputo donare e a Carlo che si è preso cura di me con tanta dedizione e mi ha sopportato nei momenti più difficili.

E infine un grazie a due persone che hanno iniziato con me questo percorso, ma che purtroppo non ci sono più: il Prof. Angelo Rosolen che mi aveva accolto nel suo laboratorio e Matteo che a forza di dirmi "la vita è una sola" mi ha insegnato veramente a viverla.