UNIVERSITÀ DEGLI STUDI DELL'INSUBRIA

SCUOLA DI DOTTORATO IN BIOTECNOLOGIE, BIOSCIENZE E TECNOLOGIE CHIRURGICHE

Curriculum BIOLOGIA CELLULARE E MOLECOLARE

XXX CICLO

RUNX2 associated long non-coding RNA characterization

Caratterizzazione di long non-coding RNA associati a RUNX2

Docente guida: Prof. Douglas M. Noonan
Tutor: Dott.ssa Alessia Ciarrocchi

Tesi di Dottorato di:
Dott.ssa Teresa Rossi

Matricola n. 725802

Dip. Biotecnologie e Scienze della Vita - Università degli Studi dell'Insubria
Anno accademico 2016-2017

ABSTRACT

RUNX2 is a lineage-specific transcription factor (TF) known to promote cancer progression. However, the molecular mechanisms that control RUNX2 expression in cancer remain widely unknown. Long non-coding RNAs (lncRNAs) are a novel class of transcripts that do not code for proteins and are often engaged in gene expression regulation.

Using the ENCODE annotation data, we identified a previously uncharacterized family of lncRNAs within the RUNX2 locus, that we named RAIN (RUNX2 Associated Intergenic Non-coding RNA). We showed that RAIN comprises 4 major variants that share a common central region but differ at the 5'- and 3'-ends. The longest isoform (l-RAIN) is nuclear and strongly associated with chromatin, suggesting a role of RAIN in gene expression regulation. Expression analysis in cancer cell lines and patient samples demonstrated that RAIN correlates with RUNX2. Furthermore, RAIN silencing resulted in a significant RUNX2 repression demonstrating that this lncRNA is required for the expression of this TF in cancer. We showed that RAIN promotes RUNX2 expression at least through two distinct mechanisms. Interacting with WDR5 and directing its recruitment to the RUNX2-P2 promoter, RAIN modifies its transcriptional activation status, bursting transcription initiation. In parallel, RAIN sequesters NELFe preventing the binding of the NELF complex to the RUNX2 P2 promoter and restraining its inhibitory function on nascent transcripts elongation. Finally, we investigated the RAIN associated transcriptional profile in thyroid cancer showing that beside RUNX2, this lncRNA controls a panel of cancer associated TFs. Overall, our data characterize the function of a novel lncRNA and identify an additional layer in the complex of RUNX2 regulation in cancer.

INTRODUCTION

RUNX family

The Runt-related transcription factors (RUNX) belong to a family of metazoan transcription factors essential during development. The *Runt* gene was first identified in *Drosophila* melanogaster [1] as a transcription factor important for the development of the limbs, eye and antennae. In mammals, there are three proteins belonging to this family: RUNX1, also known as CBFA2 (Core-Binding Factor subunit α-2), AML1 (Acute Myeloid Leukemia 1) and Pebp2αb (Polyomavirus Enhancer Binding Protein 2 subunit αb); RUNX2, also known as CBFA1, AML3 and Pebp2αa; RUNX3, also known as CBFA3, AML2 and Pebp2αc. In human, these genes localized on 21q22.12, 6p21.1 and 1p36.11, respectively [2-4].

The three RUNX genes share a common gene organization and a common protein structure likely since they arise from single gene duplications and functional diversification. In particular, they share a highly conserved Runt domain. This is a DNA binding domain of 128 aminoacids that recognizes a specific DNA sequence PyGPyGGTPy (Py=pyrimidin) and is essential for RUNX heterodimerization with the transcriptional co-activator CBFβ (Core-Binding Factor β)/PEBP2β (Polyomavirus Enhancer Binding Protein 2 β) [5]. All RUNX factors have two alternative transcriptional starting sites, within two different promoters: the distal P1 promoter and the proximal P2 promoter. These promoters are selectively activated during development and give rise to two alternative proteins with different N-terminal. In addition, the RUNX factors share a carboxyl terminus (VWRPY) and present several activation domains (AD) and inhibitory domains (ID) that can interact with other transcription factors either with activatory or inhibitory functions [6;7] (Fig.1) (reviewed in [8]).

Figure 1 The RUNX family structure RUNX genes have two promoters (P1 and P2), a common RUNT homology domain, activation and inhibition domain (AD/ID) and a VWRPY (Valine-Tryptophan-Arginine-Proline-Tyrosine) domain.

For example, the RUNX2 AD/ID domain has been shown to interact with transcription promoting factors including YAP (Yes-Associated Protein) [9], HES-1 (Enhancer of Split-1) [10], MOZ (Monocytic leukemia Zinc finger) and MORF (MOZ-Related Factor) [11], or with repressor factors like HDAC6 (Histone Deacetylase 6) [12] and TLE (Transducin Like Enhancer Of Split 1) [13].

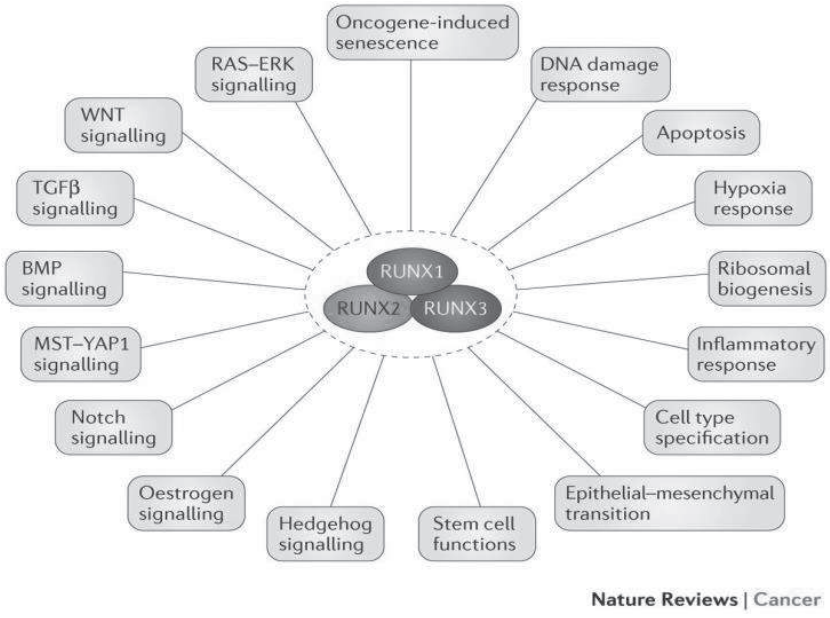

Figure 2 The RUNX family interaction landscape The RUNX family can interact with different targets belonging to several pathways; these can be functionally redundant and can impact on distinct transcriptional programs to regulate cell development, differentiation and proliferation. These pathways can be deregulated in cancer, promoting tumor aggressiveness and metastatization (Image from [23]).

In some cases, changes in one of the RUNX factor may alter the levels of the others. For example, in B cells RUNX1 and RUNX3 are inversely correlated [2]. As well an inverse correlation between RUNX2 and RUNX1 has been observed during skeletal development [24], while in breast cancer RUNX1 is inversely related to RUNX2 and RUNX3 [25].

RUNX2

The human RUNX2 gene stretches 227.766 nucleotides on chromosome 6p21.1 and the most represented splicing isoform consists of 8 exons. As all RUNX genes, it presents two major isoforms, starting from the two different promoters: the RUNX2 I-type (also called mesenchymal) is transcribed from the proximal P2 promoter and it originates a 507 aminoacids protein, while the RUNX2 II-type (also known as osteoblastic) starts from P1 and is translated in a 521 aminoacids protein [26]. These promoters are separately activated during different developmental processes and are able to generate two different proteins with different amino-termini: MASNS (Methionine-Alanine-Serine-Asparagine-Serine) and MRIPV (Methionine-Arginine-Isoleucine-Proline-Valine), respectively. In addition, both RUNX2 isoforms present a QA-rich (glutamine-arginine) domain, a NLS (Nuclear Localization Signal) and a NMTS (Nuclear Matrix Targeting Signal) [27-28] (Fig.3).

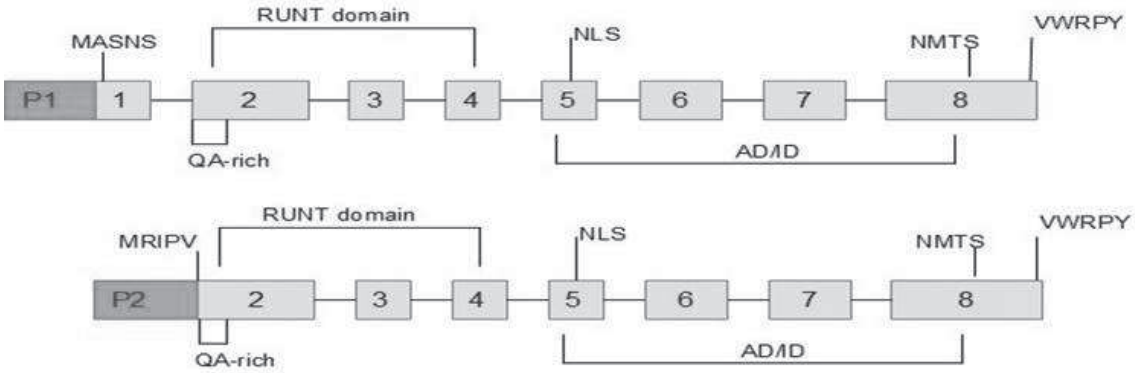

Figure 3 RUNX2 structure The two RUNX2 isoforms share common domains to the other RUNX family members (RUNT domain, AD/ID and VWRPY motif) but present also peculiar domains. P1-isoform II present a MASNS motif at N-terminus, while the P2-isoform I have one less exon than RUNX2-II and have a MRIPV motif at N-terminus. The two RUNX2 isoforms also share common domains: QA-rich motif, NLS and NMTS

The expression of the two RUNX2 isoforms is finely regulated and highly dependent on the activity of the two promoters. The P1-derived isoform is mainly expressed throughout the entire bone morphogenetic process, from osteoblast precursors to mature osteoblasts and terminal differentiated chondrocytes, while the P2-derived isoform is more widely expressed. Its expression is enriched in early precursor of chondrocytes and osteoblasts [29] but also in non-osseous tissues, such as thyroid, breast, prostate and lung. Within the RUNX2 promoter sequence, there are several RUNX consensus binding sites implying that *runx* proteins are able to cross-regulate themselves and the other RUNX paralogues [8; 30-32].

RUNX2 expression is tightly regulated by different signaling pathways.

A critical role in maintaining bone mass and in promoting osteoblast differentiation is played by the WNT (Wingless-type MMTV integration site)/LRP5 (low-density lipoprotein receptor related protein5)/β-catenin pathway [33-34]. Activation of the canonical WNT signaling lead to multiple events that induce TCF1 (T-cell factor 1) expression and the translocation of β-

catenin into the nucleus where it forms a complex with TCF1 on the RUNX2 promoter for its induction [35].

BMP2 (Bone Morphogenetic Protein 2) induces osteoblast differentiation and bone formation through ligation with its receptor and resulting Smad1/5/8 phosphorylation and generation of a nuclear complex with Smad4 that is able to activate RUNX2 gene [36]. A similar signaling cascade is activated by TGFβ (Transforming Growth Factor β) [37]. On the other hand, TNF (Tumor Necrosis Factor) have a contrary effect on RUNX2 acting on MAPK (Mitogen-Activated Protein Kinase)/p38 signaling cascade leading to inhibition of osteoblast differentiation and to bone mass loosening [38-39].

RUNX2 is also regulated through post-transcriptional modifications, such as acetylation, sumoylation, phosphorylation and ubiquitination. RUNX2 phosphorylation is usually mediated by ERK (Extracellular signal–Regulated Kinases)/MAPK cascades in the nucleus and can lead to positive [40] or negative [41-42] regulation. Even ubiquitination of RUNX2 is able to regulate its activity both positively [43] and negatively [44].

Finally, RUNX2 is regulated post-transcriptionally by both miRNA (microRNA) and small non-coding RNAs. Different studies have linked diverse miRNAs to RUNX2 activity in normal and in tumor cells; for example, mir-30a [45] and miR-103a [46] inhibit osteolysis through RUNX2 down-regulation; miR-204/211 regulates RUNX2 promoting adipogenesis and inhibiting osteogenesis of mesenchymal progenitor cells [47].

RUNX2 and cancer

Several studies, including our [48], have linked the over-expression of RUNX2 to tumor development and progression. Isoform I, encoded by the P2 promoter,

is by far the most prominent (and in epithelial derived cancer the solely) RUNX2 expressed isoform, being associated with development of osteosarcoma [49-50], prostate cancer [51], melanoma [52], ovarian cancer [53] and thyroid cancer [48]. Furthermore, many scientific evidence associate RUNX2 expression to bone metastatization in breast and thyroid cancer through TGFβ [54-55] and WNT pathways [35], in addition to estrogen signaling [56-57].

Cancer cells that metastasize to bone are able to activate a genetic pathway similar to the bone cells one, this phenomenon is called “osteomimcry” (reviewed in [58]). The ability of RUNX2 to promote bone metastasis is associated to the induction of bone-related genes (BRGs [59]) leading to bone-like phenotype of cancer cells. Bellahcène [60] and Kang [61] have previously demonstrated that breast cancer metastases present a specific gene signature, with the over-expression of BSP (Bone Sialoprotein), ALP (Alkaline Phosphatase), Col1A1 (Collagen 1α1) and OPN (Osteopontin) and other genes.

RUNX2 is also the master-regulator of several genes associated to matrix degradation and cells motility [62], such as MMP-13 (metalloproteinase-13) and OPN [63]. Furthermore, RUNX2 is able to promote tumor angiogenesis by regulating factors such as VEGF (Vascular Endothelial Growth Factor) [64] and MMP-9 [65]. Thanks to its ability to regulate all these target genes, RUNX2 has been associated with tumor progression. Furthermore, RUNX2 over-expression has been linked to epithelial-mesenchymal transition (EMT) program, especially in breast and thyroid cancer [59; 66-67], which further underlines the contribution of RUNX2 to the acquisition of aggressive features and tumor progression. A previous study, in our lab, demonstrated that the expression of CDH6 is under the control of RUNX2 and correlates with EMT and invasion potential [55].

The oncogenic potential of RUNX2 has been also associated with the inhibition of p53 activity. Indeed, p53 is able to arrest cell cycle progression in G1/S and/or G2/M, if there is a repairable DNA damage, or to activate cells apoptosis if a severe DNA damage occurs [68-69]. Several studies have determined that RUNX2 inhibits apoptosis through Bcl-2 (B-cell lymphoma 2) induction [70] and neutralizing p53 [71] and p21 collaborating with HDAC6 [12;72].

Because of its role in cancer promotion, RUNX2 is a promising target for anti-cancer strategies. Indeed, we have recently shown that the cytotoxic effects of epigenetic drugs like HDAC and BET (Bromodomain and Extraterminal Domain) inhibitors (HDACi and BETi, respectively) [73-74] are associated with a profound reduction of RUNX2 expression. Thus, understanding the molecular mechanisms that drive RUNX2 expression in cancer is important not only to get insights into the processes that support cancer progression but also to develop better strategies to counteract the activity of this transcription factor in cancer.

Long non-coding RNAs (lncRNAs)

In recent years, increasing evidence indicate that the non-coding genome plays fundamental roles in the regulation of coding-genes. In particular, in 2003 the US National Human Genome Research Institute (NHGRI) launched the ENCODE (Encyclopedia of DNA Elements) project, which involves research groups worldwide. This project aims to characterize all functional elements in the human genome; in 2007 they published the first results of their analyses [75]. One of the most exciting and surprising observation has been the wide transcriptional activity of the non-coding genome. They identified many non-coding transcripts, comprising new regulatory elements and new transcription starting sites, overlapping protein-coding region and “silent” DNA region.

Long non-coding RNAs (lncRNAs) are molecules longer than 200 nucleotides that do not encode for proteins [76-78]. These transcripts share the same transcriptional biogenesis as the mRNAs, being transcribed by the RNA polymerase II (RNA-Pol II) and containing exons. They also have 5' terminal methylguanosine cap and are frequently spliced and polyadenylated. By contrast, lncRNAs lack or have limited open reading frames (ORFs), are less expressed than mRNAs and display a higher tissue-specific expression pattern [79-80]. Furthermore, lncRNAs are poorly conserved during the evolution, even if they may present conserved secondary structures or short domains [81]. As for proteins, the identification of structurally conserved domains could represent a useful tool for the functional annotation and classification of these new molecules. However, differently from proteins, this seems to be a very difficult challenge for lncRNAs determined primarily by the high sequence heterogeneity and by the still limited information on their biological function. Several bioinformatic tools have been recently developed to identify potential domains able to mediate the functional interaction of lncRNAs with specific proteins. Some of them are derived from mRNA analyses, as MEMERIS (Multiple Em for Motif Elucitadion in RNA's Including secondary Structures) [137], a tool that integrate information from sequence motif and secondary structure to unveil RNA binding proteins interaction, and GraphProt [138], a tool to unveil binding sites of RNA-interacting protein. Only a few software are specific for lncRNAs; one of them is based on the analyses of CLIP-seq data combined with RNA-seq and GWAS (Genome-Wide Association Study) data [139]. Nevertheless, the application of these tools is still limited and will surely be implemented when we will reach a more consistent knowledge on the lncRNAs' domains functionality. Anyway, implementing these tools to further characterized lncRNAs would be very useful and it would be of great interest to make these tools easier for not-bioinformatic researcher.

LncRNAs can be transcribed from different functional elements in the genome. Actually, diverse non-coding transcripts originate from enhancers, promoters or intron regions. Otherwise, lncRNA can be named on the basis of their genomic localization; in particular, they can be intergenic, it means they are between protein-coding genes or gene-associated. Moreover, in this case, lncRNAs can be transcribed either in sense or anti-sense relatively to their associated coding gene (Fig.4).

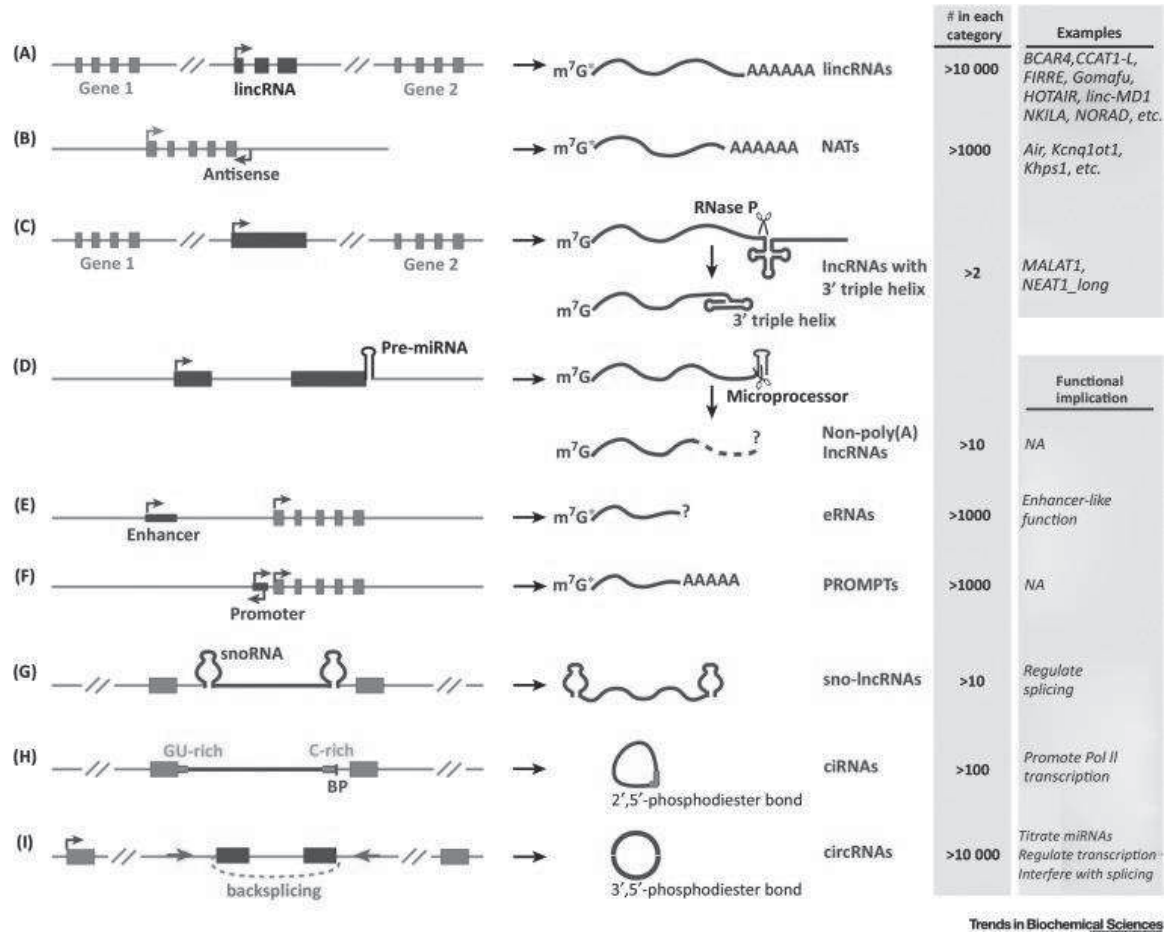

Figure 4 The multiplicity of lncRNAs in mammalian genome LncRNAs divided on the basis of their transcription site, the sense of the transcription and the post-transcriptional processes. Abbreviation: lincRNA (Large Intergenic Non-coding RNA), NAT (Natural Antisense Transcript), eRNA (enhancer RNA), PROMPT (Promoter Upstream Transcript), sno-lncRNA (lncRNA with Small Nucleolar RNA ends), ciRNA (Circular Intronic RNA), circRNA (Circular RNA) (Image from [82]).

LncRNAs are finely regulated at various levels: localization, chromatin state and post-transcriptional regulation works together to determine the cell-, tissue-developmental-, disease- state (reviewed in [83] and [84]). It is also known that lncRNAs are controlled at different levels of their genesis, maturation and degradation. Analysis of histone modification patterns have been largely used for the identification of active lncRNA-transcription sites. Similar to protein-coding genes, actively transcribed lncRNA loci are enriched in H3K4me3, H3K9ac, H3K27ac. LncRNAs are subjected to special post-transcriptional processing different from those of mRNA and similar to the one of tRNA (transfer RNA): RNase P is able to cleave the 3'-termini of some lncRNAs, such as MALAT1 (Metastasis-Associated Lung Adenocarcinoma Transcript 1) and NEAT1 (Nuclear-Enriched Abundant Transcript 1) to obtain mature lncRNA and to increase their stability [85]. One more mechanism include the stabilization of non-coding transcripts through the transcription of protein-introns and the formation of sno-lncRNA: a lncRNA transcript, lacking 5'cap and polyadenylated, flanked by two snoRNAs (small nucleolar RNA).

Another mechanism of lncRNA post-transcriptional regulation, is the circularization of some RNAs (circRNAs) that can also have sponge-like features to retain miRNAs, as CDR1as (Cerebellar Degeneration-Related protein 1 Antisense) which is able to retain more than 70 miRNAs [86].

Classification of lncRNAs

The great effort in mapping functional elements within the genome lead to a massive annotation of novel lncRNAs, the majority of which are still functionally uncharacterized.

Now a day, the number of identified lncRNA exceeds 30.000. Their wide number and high expression specificity, qualify lncRNAs as promising biomarkers in different diseases. Many annotation databases have been recently developed representing precious tools for the study of lncRNA biology. Beside the ENCODE project, which results were pivotal for the comprehension of genomic function, other databases relevant in the field of lncRNAs are FANTOM (Functional Annotation of the mammalian genome), GTEx (Genotype-Tissue Expression) and GENCODE (Encyclopædia of genes and gene variants). The FANTOM project, started in 2000, allocates functional annotations to the full-length cDNAs first in mouse [87] and later in mammalian [88-89]. This consortium intent is to identify and characterize the non-coding genome elements in different cells types, revealing what genome portions are actively transcribed during the development phases and cells differentiation. In 2017, FANTOM5 (fifth phase of the project) also generated a comprehensive atlas of more than 27.000 lncRNAs, with independent cell-type-specific expression profiles, using a CAGE (5' Cap Analysis of Gene Expression) approach [90]. Moreover, also the GENCODE project keep on studying the human genome to integrate and expand human annotation from the ENCODE project [91], adding information on lncRNA expression, structure and function [79].

The massive amount of novel ncRNAs and their great variety make difficult their functional characterization, while their heterogeneity complicates their possible classification based on common features. Nevertheless, first attempts of classification for these molecules have been suggested. To simplify, non-coding RNAs have been firstly categorized by their size: short non-coding RNA (less than 200 base pairs in length), including transcripts as snoRNAs, tRNA, miRNA, siRNA (short/small interfering RNA); and long non-coding RNA

(more than 200 base pairs in length), that includes, for example, lincRNAs, pseudogenes, ciRNAs and many others. Furthermore, lncRNAs have been stratified based on their genomic localization. These classifications are continuously updated, the categorization is not accepted worldwide, and the borders are flaky (Fig.5).

Category	Abbreviation	Refs	Specific examples
Classification based on transcript length			
Long noncoding RNA	lncRNA	[38,39]	
Long-intergenic noncoding RNA; large intervening noncoding RNA, long-intervening noncoding RNA	lincRNA	[18]	ANRIL [117], H19 [147], HOTAIR [18], HOTTIP [148], lincRNA-p21 [149], XIST [150], Paupar [151]
Very long intergenic noncoding RNA	vlincRNA	[29]	HELLP transcript [42], Vlinc_21, vlinc_185, vlinc_377, vlinc_500 [29]
macroRNA		[28,152]	Airn, Gtl2lt, KCNQOT1, Lncat, Nespas (reviewed in [152]), STAiR1 [28]
Promoter-associated long RNA	PALR	[38]	
Classification based on association with annotated protein-coding genes			
Intronic ncRNA; stable intronic sequence RNA; totally intronic RNA, partially intronic RNA	sisRNA, TIN, PIN	[49,50,54] additional references in the text	
Circular intronic RNAs	ciRNAs	[55]	
Sense ncRNA		[44]	
Natural antisense ncRNA	asRNA, NAT	[57]	BACE1-AS [153], aHIF [154], Tsix [155]
Mirror antisense		[44,67]	Globin antisense [67]
Exonic circular RNAs	ecircRNAs	[62]	cANRIL [118]
Chimeric RNAs, trans-spliced RNAs, exon juxtaposition		[44,63–66]	
Stand-alone ncRNAs made from 3′UTRs	uaRNA	[60]	
Chromatin-interlinking RNA	ciRNA	[68]	
Transcription start site-associated RNAs	TSSa-RNAs	[156]	
Classification based on association with other DNA elements of known function			
Enhancer-associated RNA	eRNA	[157]	
Promoter-associated long RNA	PALR	[38]	
Upstream antisense RNA	uaRNA	[158]	
PROMoter uPstream Transcript	PROMPT	[89]	
Telomeric repeat-containing RNA	TERRA	[159]	
Classification based on protein-coding RNA resemblance			
mRNA-like noncoding RNAs	mlncRNAs	[18]	
Long-intergenic noncoding RNA; large intervening noncoding RNA, long-intervening noncoding RNA	lincRNA	[18]	ANRIL [117], H19 [147], HOTAIR [18], HOTTIP [148], lincRNA-p21 [149], XIST [150]
Classification based on association with repeats			
C0T-1 repeat RNA		[160]	
Long interspersed nuclear element	LINE1/2	[161]	
Transcribed endogenous retroviruses		[81]	
Expressed Satellite Repeats		[162]	
Non-coding RNA driven by promoters within repeats	vlincRNAs, NASTs	[29,76]	Vlinc_21, vlinc_185, vlinc_377, vlinc_500 [29]
Polypurine-repeat-containing RNA	GRC-RNA	[163]	
Transcribed pseudogenes		[83]	PTENP1 and KRASP1 [86]
Classification based on association with a biochemical pathway or stability			
Nrd1-unterminated transcript	NUT	[164]	
miRNA primary transcripts		[165]	H19 [166]
piRNA primary transcripts		[167]	
Cryptic unstable transcript	CUT	[88]	
PROMoter uPstream Transcript	PROMPT	[89]	
Xrn1-sensitive unstable transcript	XUT	[91]	
Stable Uncharacterized Transcript, Stable Unannotated Transcript	SUT	[92]	
Classification based on sequence and structure conservation			
Transcribed-ultraconserved regions	T-UCR	[95]	UCR106 [95]
Hypoxia-induced noncoding ultraconserved transcript	HINCUT	[100]	
Long-intergenic noncoding RNA; large intervening noncoding RNA, long-intervening noncoding RNA	lincRNA	[18]	HOTAIR [18], HOTTIP [148]
RNA-Z regions		[97]	
EvoFold regions		[98]	

Category	Abbreviation	Refs	Specific examples
Classification based on expression in different biological states			
Long stress-induced noncoding transcript	LSINCT	[101]	
Hypoxia-induced noncoding ultraconserved transcript	HINCUT	[100]	
Non-Annotated Stem Transcript	NAST	[76]	
Classification based on association with subcellular structures			
Chromatin-associated RNA	CAR	[102]	
Chromatin-interlinking RNA	ciRNA	[68]	
Nuclear bodies associated RNAs		[168]	
PRC2 associated RNAs		[19,103]	
Classification based on function			
Long noncoding RNAs with enhancer-like function; ncRNA-activating	ncRNA-a	[108]	ncRNA-a7 [108]
miRNA primary transcripts		[165]	H19 [166]
piRNA primary transcripts		[167]	
Competing endogenous RNA	ceRNA	[109]	PTENP1 and KRASP1 [86]

Figure 5 LncRNAs category Examples of classifications hypothesized on different characteristics (Image from: [92]).

Enhancer-Associated lncRNAs

For many decades, gene expression regulation has been considered as a mono-dimensional process in which each gene was controlled by the activity of the nearest promoter. Systematic functional analysis of non-coding genome has revealed that gene expression requires a continuous and widespread regulatory landscape involving a specific genomic architecture and the hierarchical interactions of multiple interspersed regulatory elements.

Many factors collaborate to regulate the gene expression, so that genes are expressed in the right place in the right moment [93]. Among all these, there are the enhancers (ENHs), DNA sequences containing multiple binding sites for transcription factors, RNA-PolII (RNA-Polymerase II) and co-factors. ENHs are able to activate transcription independently from their distance from the promoter and the strand orientation; they can also act on genes located on different chromosomes [94-95]. These genome elements can recruit transcription factors and bring them in contact to the gene promoter enhancing transcription through the formation of chromatin loops.

In recent years, it has been observed that enhancers are also transcribed into non-coding RNAs. Whether these molecules contribute to the ENHs function is

still under debate. ENHs can be transcribed into two distinct classes of non-coding transcripts: eRNA short RNA coded in both orientation or lncRNA called as ENH-associated lncRNA. In 2010 Kim TK and colleagues [96] and Ørom UA and colleagues [97] used genome-wide analysis and GENCODE to define features of ENH-Associated ncRNA. They showed that transcribed enhancers have peculiar chromatin marks, like high levels of RNA-PolII, binding of CBP (CREB-Binding Protein) and p300, H3K4Me1 high and H3K4Me3 low. Furthermore, they found that these lncRNAs were able to regulate *in cis* neighboring protein-coding genes as well as control genes *in trans*, that is the control of genes located far in the linear sequence of the genome.

Currently, the main hypothesis about the role of these lncRNAs is that they serve as a chromatin hub interacting with other factors, such as histone modification complex and TFs. It has been demonstrated that lncRNAs are able to interact with WDR5 (WD Repeat Domain 5-a subunit of methyltransferase complex) to regulate genes activation [98-99].

Further, some of these lncRNAs exhibit also the role to stabilize the engagement between promoter and enhancers, interacting with Mediator complex [100] and cohesion [101].

Moreover, ENH-associated lncRNAs are able to assist the recruitment of TFs and to maintain them at their regulatory sites. For example, YY1 (Yin-Yang 1) is a transcription factor able to bind both promoter-/enhancers-associated elements and RNA transcribed from those, suggesting that these nascent RNAs can stabilize the engagement of this TF [102]. CCAT-1L (Colorectal Cancer Associated Transcript 1-long isoform) is an ENH-associated lncRNA, positively associated to MYC transcription, that interacts with CTCF (CCCTC-

binding factor) modulating its binding to chromatin leading to correct looping of the locus [103].

Finally, recently evidence have proposed that eRNA may control also transcription elongation by sequestering and inhibiting the Negative Elongation Factor (NELF) complex. Acting as decoy for NELFe, the RNA-binding subunit of the NELF complex, eRNA restrains the binding of the NELF complex downstream to the gene TSS (Transcription Starting Site) relieving RNA-PolII pausing and activating transcript elongation. However, whether this is a common property for lncRNA is still to be defined [104].

Localization and function of lncRNAs

LncRNAs are ubiquitously distributed in the cell compartments: they have been found to localize in cytosol, nuclear fraction or associated with chromatin; they can also shuttle between the nucleus to the cytosol. However, lncRNAs are more enriched in the nucleus rather than in the cytosol, differing from mRNA, that are more abundant in the cytosol [105-106]. LncRNAs and mRNAs differ also for the mechanism of degradation. Being mainly nuclear, lncRNAs are exposed to nuclear exosome and in minor degree to cytosolic nonsense-mediated decay (NMD). Instead, mRNAs only head towards the ribosomes in the cytoplasmic compartment and are degraded by decapping and 5'-to-3' exonuclease activity.

LncRNAs can be cleaved to form other short RNAs, as miRNAs and siRNAs [107], or tRNAs [108] that are able to shuttle to the cytosol.

The different localizations affect lncRNAs function. Furthermore, lncRNAs present binding-domain for DNA, RNA and proteins and the binding with

respective targets lead also to conformational changes. It has been demonstrated that presence of specific RNA motif lead to different localization: BORG (BMP2-OP1-Responsive Gene) is a lncRNA that present a pentamer sequence AGCCC and T or A at position -8 and G or C at -3 specific for nuclear localization [109].

LncRNA nuclear-localized/ chromatin-associated are often gene regulators. Indeed, lncRNAs are physically related to their genomic locus and make them able to exert their activity without been previously exported to the cytoplasm for post-transcriptional modification. In fact, lncRNAs are able to recruit histone modification complex to induce or inhibit specific genes both in *cis* or in *trans* [110]. For example, KCNQ1OT1 (Potassium voltage-gated Channel subfamily Q member 1 opposite strand/antisense transcript 1) is able to interact with histone modification complexes: G9a (also known as EHMT2-euchromatic histone lysine methyltransferase 2) and PRC2 (Polycomb Repressive Complex 2), both presenting methyltranferase activity, to mediate specific silencing of gene during fetal development [111]. HOTAIR (HOX Antisense Intergenic RNA) is another lncRNA that is able to interact with two different histone modification complexes, PRC2 and LSD1 (Lys-Specific Demethylase 1), in two different domains: PRC2 with a domain located in 5’ and LSD1 with a 3'domain [112].

LncRNAs are also implicated in the organization of nucleus and subnuclear compartments, such as speckles and paraspeckles. In particular, speckles are nuclear bodies that contain pre-mRNA splicing factor; instead, paraspeckles have a relevant role in the modulation of mRNA and protein levels because they are able to sequester them into nuclear bodies. MALAT1 is a lncRNA localized in the speckles that indirectly interacts with pre-mRNAs through serine/arginine

(SR) RNA splicing proteins; its down-regulation reduce the recruitment of SR proteins [113] and affect alternative splicing [114].

NEAT1 is a lncRNA fundamental for the architecture of paraspeckles; its depletion lead to the disassemble of these structures [115].

LncRNAs in cancer

A potential function for lncRNAs in human diseases has been proposed. Among these, lncRNAs associated to cancer is one of the best studied branch: more than 4900 papers about this, can be found in PubMed. This is due to the different roles and the multiple interactions that lncRNAs exhibit in cells, and due to the wide expression of these transcripts.

As previously described, lncRNAs act as fine regulators of gene transcription and chromatin accessibility. So, rearrangements and activating/inhibitory mismatch could lead to aberrant expression and function of onco-suppressor and oncogenic genes. Basically, deregulation of cell cycle, chromatin and epigenetic state, changes in RNA/DNA/proteins interactions and in their activity, could induce neoplastic transformation leading to carcinogenesis.

Many studies demonstrated that some lncRNAs are associated to specific cancer, while others are associated with several tumors originating from different tissues (fig.6).

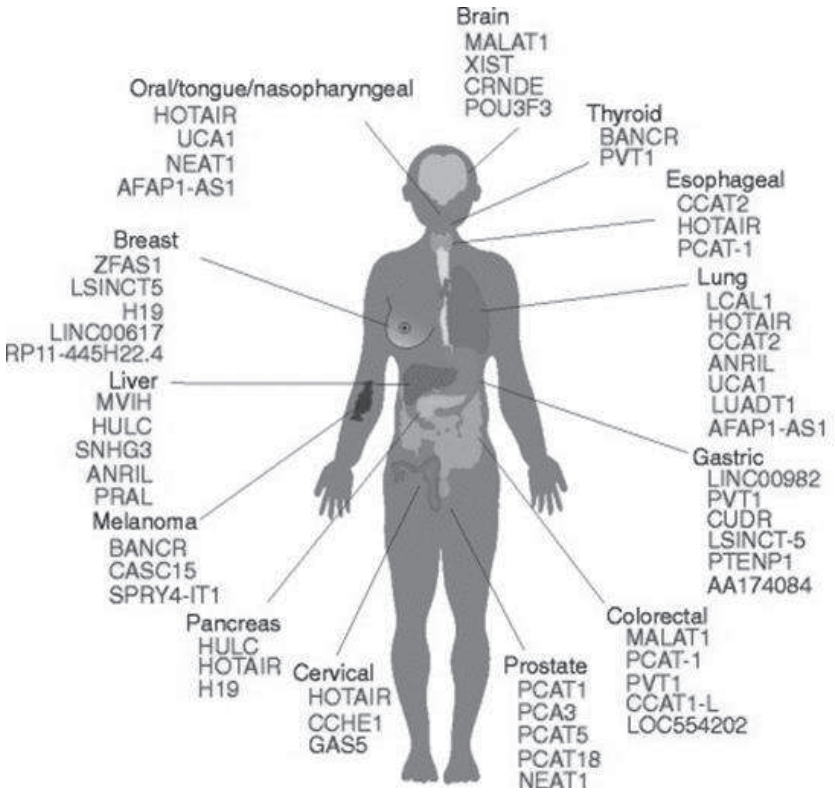

Figure 6 Examples of lncRNAs cancer-associated Red represents lncRNAs up-regulated in cancer compared to normal tissue, while blue represents lncRNA down-regulated in cancer (Image from [116]).

Examples of the first group are PCGEM1 (Prostate-specific transcript 1) [117], PCA3 (Prostate Cancer gene 3) [118] and PRNCR1 (Prostate cancer Non-Coding RNA 1) [119] that are expressed only in prostate cancer; thanks to their specificity they have been proposed as markers for prostate cancer progression.

MALAT1 belongs to the second group, lncRNAs that has been found to be associated with different tissues: lung, both early-stage of NSCLC (Non-Small Cell Lung Cancer) [120] and adenorcinoma [121], bones [122], colon [123] and other cancer sub-types.

HOTAIR is another lncRNA associated with different tumor types, e.g. liver [124], gastric [125] and cervical [126] carcinoma, and with metastasis onset [127].

Furthermore, some lncRNAs are associated with metastasis or more aggressive cancer, but not with normal tissue or low-grade cancer. For example, HULC (Highly Upregulated in Liver Cancer) is a lncRNA highly expressed in liver metastasis of CRC (Colorectal Cancer) and in hepatocarcinoma, but not in primary CRC or in non-liver metastasis [128-129].

Due to their relevance in the biological and clinical field, lncRNAs have been proposed as diagnostic/prognostic biomarkers (reviewed in [130] and [116]), as PCGEM1 and PCA3. In addition, considering their importance in tumor onset and development, some lncRNAs have been proposed as therapeutic targets for the design of new therapies. Up to now there are four clinical trials that are enrolling patient in studies from Phase 1 to Phase 3, that set to use lncRNA as biomarker of drug response and to progression of disease.

RAIN: a novel RUNX2 Associated Intergenic non-coding RNA

We have previously demonstrated that the major RUNX2 isoform expressed in cancer cells is the isoform I, transcribed form the proximal-promoter P2 and that its overexpression promotes aggressiveness and metastatic potential of cancer cells [48]. However, we also showed that the P2 promoter is an indolent region and does not contains the elements required for the high levels of expression of RUNX2 in cancer.

To unveil the molecular mechanisms that lead to deregulation of RUNX2 in cancer, we recently identified three previously uncharacterized RUNX2 ENHs downstream to the P2 promoter: ENH3, ENH11 and ENH13 [74]. Being aware of the ability of active ENHs of being transcribed into lncRNA we searched the ENCODE annotation data to discover RUNX2 associated lncRNA. Several

potential transcripts were described downstream of the RUNX2 locus, overlapping with the regions of the RUNX2 ENH11-ENH13 (Fig.7).

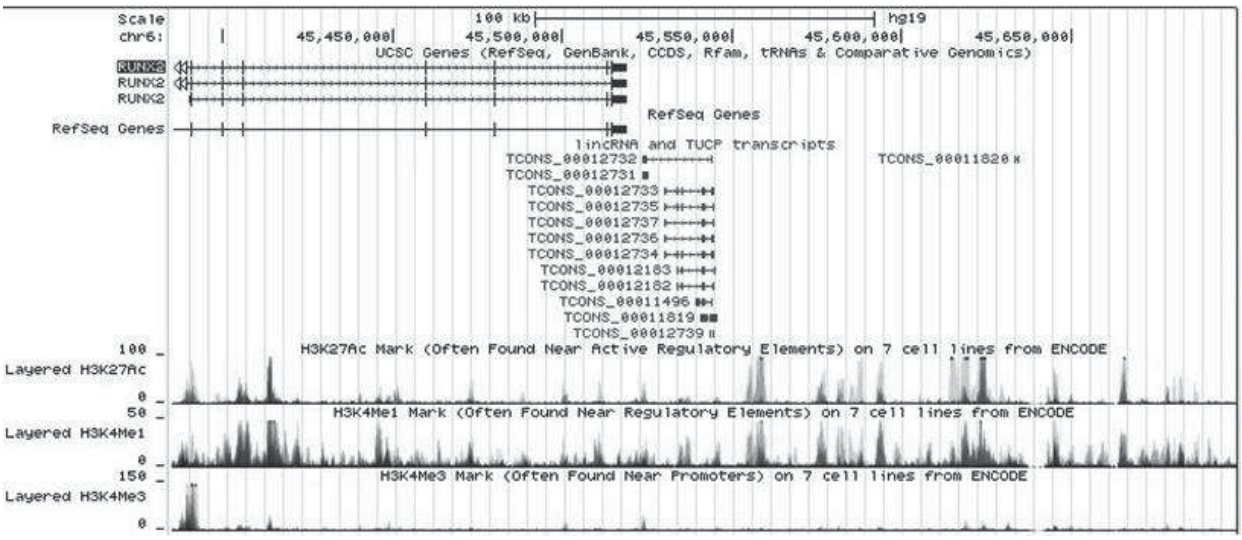

Figure 7 Genome Browser image with ENCODE data Focusing on ENCODE data downstream of RUNX2 locus, there are several predicted non-coding transcripts within the enhancer region that we have characterized.

However only one of these predicted molecules (TCONS00011820) was expressed in thyroid and mammary cancer cells. The annotate transcript was short and formed by two exons. Using a 5'- and 3'-RACE approach, we mapped the full length of this transcript in TPC1 cells. We found 4 major transcripts that presented a widely variable central region, two different starting sites, located in correspondence of the ENH10 and ENH11, and two different 3' end, a short and a long one. The long isoform is 3010bp longer than the short isoform. We named this lncRNA RAIN (RUNX2-Associated Intergenic Non-coding RNA) (Fig.8).

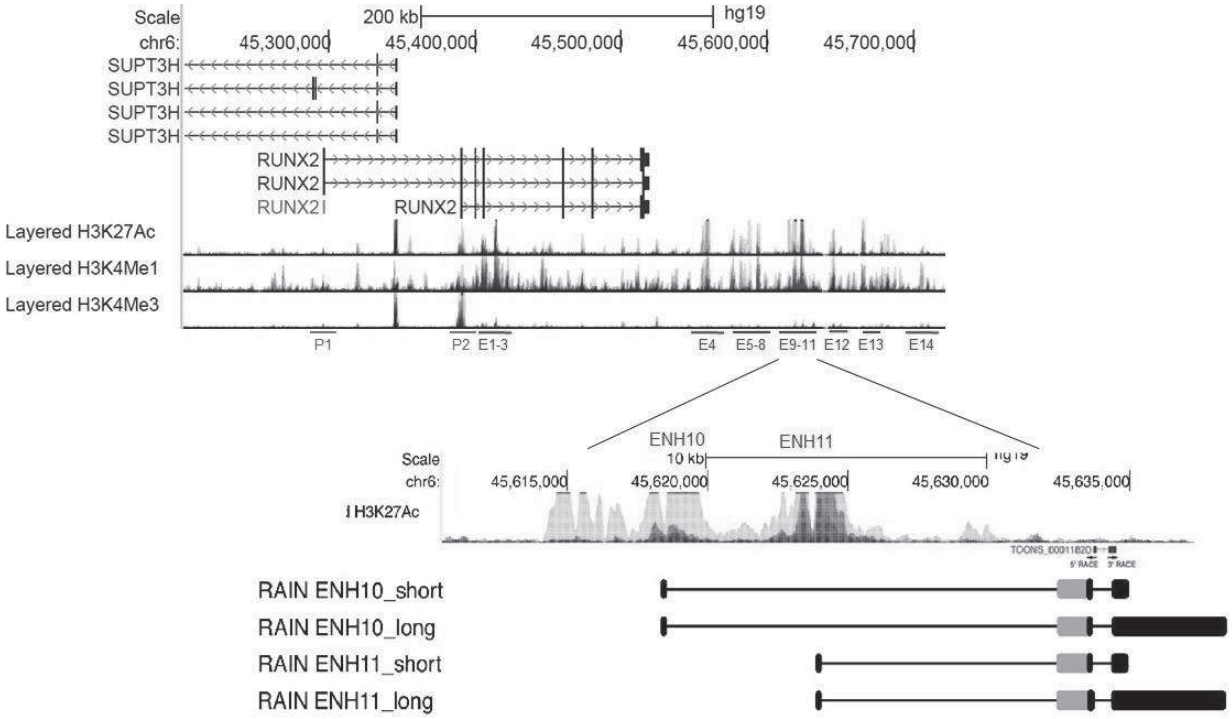

Figure 8 RAIN full transcript characterization RACE approach using TCONS00011820 as template to find the full transcript. RAIN family is composed of four members with a common central region (grey) and two different 5'- and two different 3'-ends. 5'-ends are within previously identified ENH 10 and 11; 3'-ends differ for about 3000bp of length.

RUNX2 and RAIN are co-regulated

We have previously demonstrated that RUNX2 enhancers are binding sites for different transcription factors, and the master regulator of ENH3, ENH11 and ENH13 is c-Jun [74]. Furthermore, RAIN's TSS are within RUNX2 ENH regions, in particular, correspond to ENH10 and ENH11. So, we wanted to determine if RUNX2 and RAIN can be regulated by the same elements.

We used siRNA approach and the use of a dominant negative (DN) c-Jun plasmid to interfere with this TF.

With both systems, we observed that RUNX2 and RAIN expression was down-regulated (Fig.9 a-c).

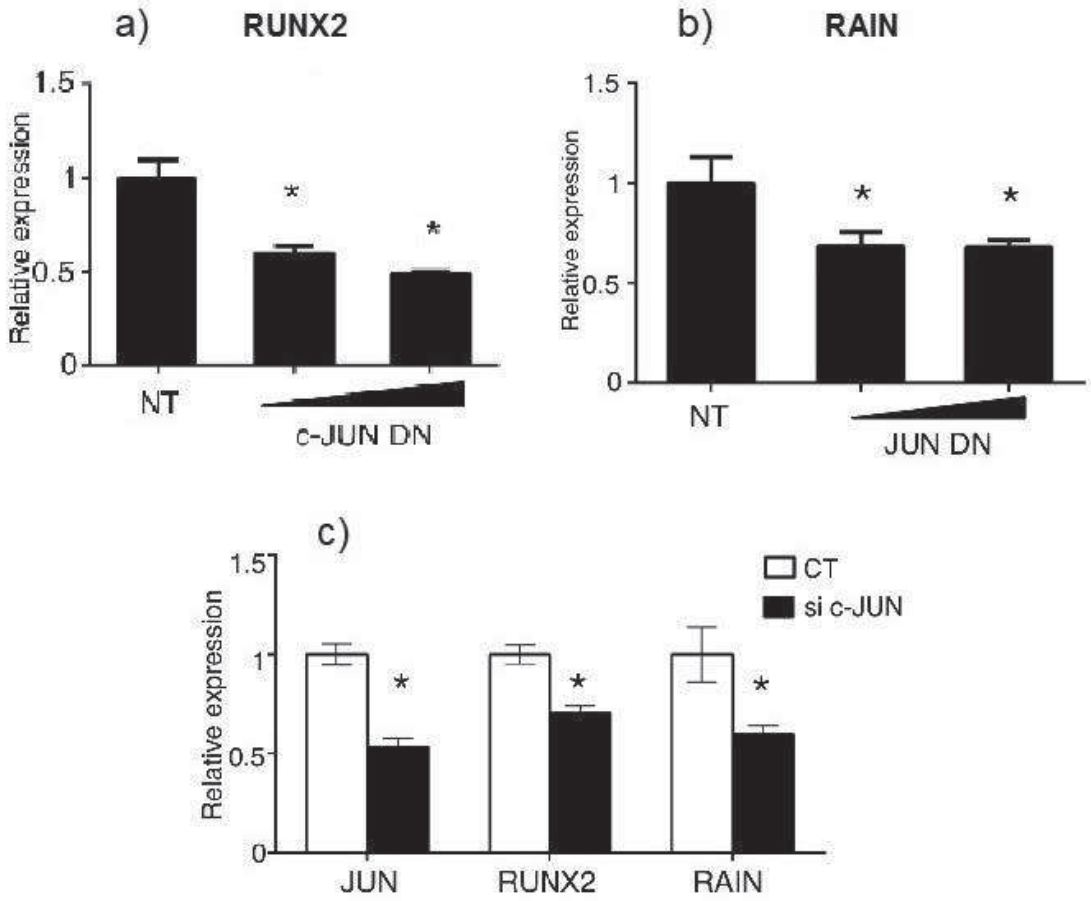

Figure 9 RUNX2 and RAIN expression quantification after c-Jun downregulation a-b) TPC1 cell line was transfected with increasing concentration (100ng and 20ng) of c-Jun dominant negative (DN) plasmid and RNA was extracted. RUNX2 (a) and RAIN (common) (b) expression was quantified by qRT-PCR, c) TPC1 cell line was reverse transfected with 25nM of siRNA against c-Jun and after 24h RNA was extracted and the expression of RUNX2 and RAIN was quantified by qRT-PCR. * p-value<0.05

BET proteins are a family of protein that interact with acetylated histones to recruit histone acetylation complex to enhance the protein-coding gene transcription. In particular, BRD4 (Bromodomain 4) present a major role in control of distal enhancer regions, especially in cancer. A recent study, has demonstrated that BET-inhibitor drugs, such as JQ1, are able to antagonize the synthesis of non-coding eRNAs [131]. We wanted to confirm this hypothesis on our ENH-associated lncRNA, even because we have previously demonstrated that JQ1 treatment lead to repression of RUNX2 expression [74]. We treated

TPC1, BCPAP, MCF7 and MDA-MB 231 cells with 1μM of JQ1 and we extracted RNA. The quantification of the expression of RAIN showed that the JQ1 treatment induce a down-regulation of RAIN (Fig.10).

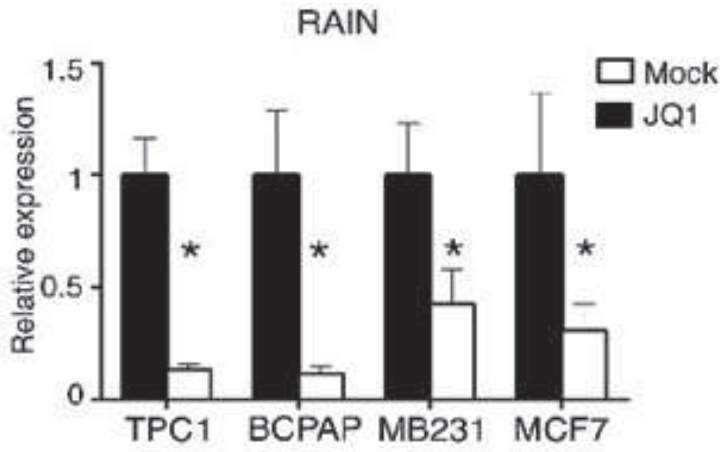

Figure 10 RAIN expression after JQ1 treatment TPC1, BCPAP. MDA-MB231 and MCF7 cell lines were treated with 1μM of JQ1. After 24h RNA was extracted and the expression of RAIN (common) was quantified by qRT-PCR. * p-value<0.05

All these observations suggest a possible relevant function of RAIN in controlling RUNX2 expression in cancer.

AIM OF THE PROJECT

We have recently discovered RAIN (RUNX2 Associated Intergenic Non-coding RNA) a new ENH-associated lncRNA within the RUNX2 locus; we have, also, observed that RAIN and RUNX2 are co-regulated.

The aim of this project was to characterize the function of RAIN, its interplay with RUNX2 and its potential relevance in cancer. We performed our analysis on a panel of different cancer cell lines, focusing on thyroid and breast tumor cell lines. In fact, in these tumor types, RUNX2 has been shown to be a marker of aggressiveness and its overexpression has been correlated with progression and metastasization.

First, we characterized the functional interaction between RAIN and RUNX2 promoter and enhancers with the intent of defining the effect and the mechanism of action of this lncRNA on RUNX2 expression regulation in the context of cancer cells. Finally, using RNA-Sequencing approach, we investigated the possibility that RAIN have other targets beside RUNX2 in thyroid cancer cells.

MATERIAL AND METHODS

Cells culture

Thyroid cancer cell lines (BCPAP, TPC1, WRO, 8505C, CAL62) were obtained from Prof. Massimo Santoro (University of Naples, Naples, Italy); FTC-133 (Thyroid follicular carcinoma cell line) were purchased from Sigma-Aldrich; SW579 (Thyroid papillary carcinoma cell line) were purchased from ATCC. All cancer cell lines were cultured at 37°C/5% CO_2 in DMEM (Life Technologies), supplemented with 10% FBS (Life Techinologies) and 1% penicillin-streptomycin (Life Technologies).

Breast cancer cell line MDA-MB-231 was obtained from Dr. Adriana Albini (Scientific and Technology Pole, IRCCS MultiMedica, Milan), MCF7 was obtained from Dr. Massimo Broggini (IRCCS-Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy), ZR-75-1 were obtained from Prof. Bertolini (IEO, Milan, Italy) and HCC1428 were purchased from ATCC. Breast cancer cell lines were cultured at 37°C/5% CO_2 in RPMI (ZR-75-1, HCC1428) or DMEM (MDA-MB-231, MCF7) supplemented with 10% FBS and 1% penicillin-streptomycin.

NCI-H1299, A549, NCI-H1650, NCI-H1975 (lung adenocarcinoma cell lines) were purchased from ATCC and cultured at 37°C/5% CO_2 in RPMI (Life Technologies) supplemented with 10% FBS and 1% penicillin-streptomycin.

LNCap, PC-3, DU145 (prostate cancer cell lines) were obtained from ATCC and were cultured at 37°C/5% CO_2 in RPMI (PC3 and LNCaP) or DMEM (DU 145) supplemented with 10% FBS and 1% penicillin-streptomycin.

Patient samples

52 fresh frozen patient samples, comprising 26 couples of matching normal and tumor tissues, were obtained from the Research Tissue Biobank of Arcispedale Santa Maria Nuova-IRCCS of Reggio Emilia after written informed consent obtained from all the patients involved in this project. The project was approved

by the local Ethical Committee (protocol no.: 2014/0014425 of 06/05/2014).

RNA isolation and qPCR Assays

Cell lines RNA samples were extracted using Trizol (Ambion) protocol and DNase (Roche) digestion was performed during purification of RNA samples.

Patient tissues RNA samples were extracted using Trizol and further purified with RNeasy extraction kit (Qiagen) following the RNA clean up protocol and performing on-column DNase (Qiagen) digestion.

Subsequently, cDNA was prepared by reverse transcription using the iScript cDNA kit (Bio-Rad); quantitative Reverse Transcription-PCR (qRT-PCR) was conducted using Sso Fast EvaGreen Super Mix (Bio-Rad) in the CFX96 Real Time PCR Detection System (Bio-Rad).

The Real-Time protocol used is: 98°C for 2'; [98°C for 2", 59°C for 5"] repeated for 40 times, followed by melting curve production: from 70°C to 95°C with an increment of 0.5°C, and 1" of plate read.

Primers sequences are listed in the table below.

Protein extraction and western blot

Total proteins were extracted with Passive Lysis Buffer (Promega) supplemented with a cocktail of protease inhibitors (Roche). Protein amount was quantified by Bradford (Bio-Rad) reagent. 15µg of proteins were loaded on any Kd SDS-PAGE gel (Bio-Rad). The proteins were transferred to nitrocellulose filters using the Trans-Blot Turbo Transfer System (Bio-Rad); after blocking with 5% milk (Bio-Rad)/PBST (PBS, with 0.1% Tween-20 (Sigma-Aldrich)), membranes were stained with primary antibodies over night at 4°C, while secondary antibody staining was performed for 1 hour at room temperature. Primary antibodies used were mouse anti-α-Tubulin (Santa Cruz, sc-8035), mouse anti-RNA Polimerase II (Abcam, ab817), anti-NELFe (F-9,

SC377052 SCBT) and anti-β-actin (A1978 Sigma-Aldrich) while the secondary antibody was Mouse IgG HRP-Linked Whole Ab (GE Healthcare, NXA931).

RACE (Rapid Amplification of cDNA Ends)

RAIN full transcripts were determined performing 5′ RACE and 3′ RACE through the SMARTer RACE 5′/3′ (Clontech) kit following the producer's instructions. Briefly, 1 μg of TPC1 cells' DNAse-treated RNA was retro-transcribed to generate 5′ RACE-ready and 3′ RACE-ready cDNA. 5′-ends and 3′-ends of RAIN transcripts were amplified using specific primers and a touchdown PCR program. Amplified fragments were extracted from agarose gel, cloned into the pRACE plasmid and sequenced.

Cells fractioning

Cells were fractionated to obtain cytosol, nucleus and chromatin fractions. Briefly, cells were harvested and washed twice with PBS (Sigma-Aldrich). Cells were resuspended in PBS and divided in three Eppendorf tubes and spinned at 3000rpm for 5' at 4°C; PBS was removed and cells were resuspended in Lysis Buffer (Promega) and protease inhibitor cocktail (bimake.com) to obtain total protein lysate, Trizol to obtain total RNA, or Cytosol Buffer (10mM HEPES pH7.9, 1.5mM $MgCl_2$, 10mM KCl, 0.5% NP-40, RNase inhibitor, protease inhibitor) for further processing. Cells were incubated with Cytosol Buffer for the appropriate time for each cell line (8' MCF7 and TPC1; 4' MDA-MB 231; 2' BCPAP) and spinned for 2' at 4°C. The supernatant was collected in two different Eppendorf tubes (cytosolic protein and RNA), centrifuged for 15' at 3000rpm to eliminate nuclear debris and transferred in two new Eppendorf tubes. Trizol was added to the RNA sample. The pellet was washed three times with Wash Buffer (10mM HEPES pH7.9, 1.5mM $MgCl_2$, 10mM KCl, RNase inhibitor, protease inhibitor) at 3000rpm for

2’ at 4°C. Then, the pellet was resuspended in Nuclear Buffer (20mM HEPES pH7.9, 25% glycerol, 0.42M NaCl, 1.5mM $MgCl_2$, 0.2mM EDTA (Sigma-Aldrich), RNase inhibitor, protease inhibitor) on ice for 30’ with frequent vortexing; after incubation, pellet was spinned at 14500rpm for 10’ at 4°C and supernatant was divided in two Eppendorf tubes to obtain nuclear protein and RNA. Trizol was added to extract nuclear RNA. The chromatin-containing pellet was resuspended in Trizol and further processed.
At least three biological replicates were conducted for each cell line.

Small interfering RNA (siRNA) and gapmeRs transfections

Select siRNA interference oligos against RAIN (common locus), NELFe, c-Jun and negative control oligos (Ambion) were used for transfections. GapmeRs against RAIN (targeting specifically the long isoforms) and negative control (Exiqon), both comprising LNA nucleotides, were used to specifically silence the RAIN long isoforms. Transfections were performed using the RNAiMax Lipofectamine (Thermo Fisher Reagent) reagent using the reverse transfection following the manufacturer’s protocol. Briefly, RNAiMax, Opti-MEM (Life Technologies) and 25nM of siRNA or gapmeRs were prepared and incubated for 20’; cells were harvested, resuspended in medium without antibiotics, and added to transfection reagents in culture plates. Next day, medium was replaced with complete fresh medium. Cells were harvested and analyzed 24 or 48 hours after transfection. GapmeRs and siRNA oligos sequences are indicated in the table below.

Plasmid vectors and transfection

c-Jun DN (Dominant Negative) expressing vector was kindly gifted by Dr. Mirko Marabese (IRCCS-Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy).

Cells were transfected with increasing concentration of c-Jun or empty vector using Lipofectamine 2000 (Thermo Scientific) following procedure's protocol. Briefly, TPC1 were plated at 70% of confluence in 24-well plate, next day increasing dilutions (100ng and 20ng) of plasmid vector or empty vector, were diluted in Opti-MEM medium and then lipofectamine 2000 was added. After 20' of incubation, reagents were added to the cells' medium (complete medium without antibiotics). After 48h hours cells were detached and further analyzed.

Cell proliferation assay

24 hours after cells transfection, $2x10^3$ cells for each cell line were seeded in triplicate in a 96-well plate in regular growth medium. Viable cells were counted every 24 hours for 4 days using trypan blue staining and Burker chamber. Three biological replicates were conducted for each cell line.

Scratch wound healing assay

24 hours after reverse transfection, each cell line was seeded at 70% confluence in a six-well plate. Next day, scratches were applied after cell adhesion by using a pipette tip. Healing areas were photographed at different time-point (0, 6 hours, 12 hours and 24 hours) and measured using ImageJ software. Three biological replicates were conducted for each cell line.

Analysis of mRNA stability

Actinomycin D (Sigma-Aldrich) was used to inhibit nascent RNA synthesis. $6x10^4$ cells/well (MDA-MB 231) or $5x10^4$ cells/well (TPC1) were seeded in each well of a 24-well plate and were treated with 5 μg/ml actinomycin D or DMSO (Sigma-Aldrich). Samples were collected at the time of treatment and after 30', 2 hours, 4 hours, 8 hours, 12 hours and 24 hours and RNA was isolated.

Fluorescence in situ hybridization

For *in situ* hybridization, $1.5x10^5$ cells were plated on a coverslip in each well of a 6-well plate; next day, cells were washed twice shortly with PBS. Cells were fixed for 10' with 4% PFA (paraformaldehyde, Santa Cruz) at room temperature. Cells were washed three times for 5' with PBS and rinsed once with ice cold 70% EtOH (Carlo Erba); then, ice cold 70% EtOH was added and cells were kept at -20°C for at least one night. On the next day, cells were rehydrated by washing twice with PBS for 5' each. Cells were permeabilized with permeabilizing solution (0.5% triton X-100 (Sigma-Aldrich) in PBS, RNase inhibitor) for 10' at room temperature with gentle swirl. Cells were washed three times for 5' with PBS and twice for 5' with 2xSSC,0.05% tween20.

Coverslips were incubated with blocking solution (1% BSA (Sigma-Aldrich), 2xSSC, 0.05% Tween 20) for 30' at room temperature in a humid chamber. Probes containing LNA nucleotides (Exiqon) were diluted to 50nM in hybridization buffer (50% deionized formamide (Carlo Erba), 2xSSC, 50mM Sodium phosphate pH 7 (Sigma-Aldrich), 10% dextran sulphate (MW>500,000 Alfa Aesar)) and added to each coverslip. Then, coverslips were incubated in HYBrite Genetic Analysis System (Abbott Laboratories) for 2' at 80°C followed by 1 hour at 57°C (negative control) or 54°C (positive control and RAIN probes) to denature nucleic acids, as indicated in the datasheets.

After incubation, coverslips were washed twice for 30' at 37ºC, with a pre-warmed wash solution (50% formamide, 1x SSC, 0.025% Tween 20, pH 7.0); then, coverslips were washed twice for 5' at 37ºC, and once at room temperature, with 2x SSC, 0.05% Tween 20.

Anti-digoxigenin (Abcam ab420) diluted 1:1000 (1µg/ml) in TNB buffer (100mM Tris-HCl pH 7.5, 150mM NaCl (Sigma-Aldrich), 0.5% BSA), was incubated in a dark humid chamber, at room temperature, for 30'. Then,

coverslips were washed three times for 5' with 2x SSC, 0.05% Tween 20 at room temperature.

Anti-mouse immunoglobulin Alexa488 (Thermo Fisher Reagents A11001) diluted 1:1000 (1µg/ml) in TNB, was incubated in a dark humid chamber, at room temperature, for 30'. Then, coverslips were washed twice for 5' with 2x SSC, 0.05% Tween 20 at room temperature.

Coverslips were stained with DAPI (D9542 Sigma-Aldrich) 1:1000 (1µg/ml) in TNB for 5' at room temperature; then the coverslips were washed shortly with PBS and mounted using SlowFade Gold antifade (Invitrogen).

Three biological replicates were conducted.

Probes sequences are listed in the table below.

ChIRP (Chromatin Isolation by RNA Purification)

ChIRP was performed following the protocol of Chu et colleagues [132] with minor adaptations. Eight biotin 3'-end TEG probes (Eurofins Genomics), matching the 3'long locus of RAIN, were used for the experiment; eight 3'- end biotin TEG probes against LacZ transcript (Eurofins Genomics) were used as negative control. Probes sequences are listed in the table below.

Briefly, $4x10^7$ cells were collected and divided in four different tubes and crosslinked for 10' at room temperature with 1% glutaraldehyde (Carlo Erba) in 10ml of final volume. Cross-linking reaction was quenched with 1.25 M glycine (Sigma-Aldrich) at room temperature for 5' and cells pellet was washed twice with ice-cold PBS and flash-frozen at -80°C. Next, pellet was thawed and resuspended in Lysis Buffer (50mM TRIS-Cl pH7.0, 10mM EDTA, 1% SDS (Sigma-Aldrich)) supplemented with RNase inhibitors and protease inhibitors. Cells were sonicated for 4 times: each time was composed of 10 cycles 30" ON-30" OFF in water sonicator Bioruptor Pico (Diagenode). To avoid different rates of sonication in the different tubes, lysates were pooled together every 10

cycles and redistributed into original tubes to ensure homogeneity. After sonication, lysates were centrifuged for 10 minutes 13000rpm at 4°C to clarify the lysate and 2% RNA and 2% DNA input were taken. Then, 1ml of chromatin was supplemented with 2ml of hybridization buffer, RNase inhibitor, proteinase inhibitor and 1.5µl of 100pmol/µl probes. The mix were incubated for 4 hours at 37°C with shaking. After incubation, 100µl of magnetic beads (C1 magnetic beads, Invitrogen) were added and incubated for 30' at 37°C with shaking. After that, beads were washed five times with wash buffer using Magna GrIP magnetic strip (Millipore) to separate beads from supernatant. At last wash, well resuspended beads were divided into 100µl for RNA isolation (10% of volume) and 900µl for DNA isolation (90% of volume).

RNA purification: Proteinase K buffer (100mM NaCl, 10mM TrisCl pH 7.0, 1mM EDTA, 0.5%SDS, 5% proteinase K (Promega 20mg/ml)) was added to 100µl of bead samples and RNA input for 45' at 50°C with end-to-end shaking; then, after centrifugation, samples were boiled for 10' at 95°C, chilled on ice and Trizol was added. Further, RNA extraction protocol with Trizol was followed and RNA was extracted with miRNeasy kit following procedure's protocol.

DNA isolation: bead samples and DNA input were supplemented with 10µl RNase A (10mg/ml) (Thermo Scientific) and 10U/µl RNase H (Thermo Scientific) per ml of DNA elution buffer (50mM $NaHCO_3$, 1% SDS) and incubated for 30' at 37°C with shaking. After, the supernatant derived from beads IPs was kept using Magna GrIP. A second round with RNAse A and H step was performed on beads and DNA input and supernatant derived from beads IPs was separated using Magna GrIP and collected with the previous. Collected supernatant and DNA input were incubated for 45' at 50°C with shaking with 15µl of proteinase K. All the DNA samples were transferred to phase-lock gel tubes (Eppendorf) and phenol:chloroform:isoamyl (Carlo Erba)

was added and shaken for 10'. Then, samples were spinned and the aqueous phase was kept and supplemented with glycogen (Thermo Scientific) and 100% EtOH and stored overnight at -20°C. Next day, samples were spinned and supernatant was let decant. 70% EtOH was added, vortexed and spinned down. Supernatant was removed, and pellet was air dried, then resuspended in 30µl of Elution Buffer; samples were used for further qPCR analysis.
Three biological replicates were conducted for each cell line.

ChIP (Chromatin Immunoprecipitation)

For Chromatin Immunoprecipitation, $4x10^6$ cells were reverse transfected and plated in a 150mm ($3.7x10^6$) petri dish and in a well of 6-well plate ($3x10^5$) (to control RUNX2 down-regulation before performing ChIP assay). 48 hours after reverse transfection, cells were cross-linked in PBS with 1% formaldehyde (Sigma-Aldrich) solution for 10' in gentle rotation. Subsequently, they were treated with 2.5 M glycine for 5' to quench the cross-link. Cells were washed twice with PBS and scraped. After that, cells were collected in a tube and centrifuged for 5' at 4 °C. Cells were lysed in Cell Lysis Buffer (10mM Tris-HCl pH 8.0, 85mM KCl, 0.5% NP-40) supplemented with Protease Inhibitor Cocktail for 10' at 4 °C. Nuclei were pelleted for 5' at 4 °C and pellet was incubated for 10' on ice in Nuclei Lysis Buffer (50mM Tris-HCl pH 8.0, 10mM EDTA, 10% SDS) supplemented with protease inhibitors. Nuclei were sonicated with 2 cycles 30" ON -30" OFF and cell debris were pelleted for 10' at 4°C. Lysate was diluted in ChIP dilution buffer (0.01% SDS, 1% Triton X-100, 1.2mM EDTA, 16.7mM Tris-HCl pH 8.0, 167mM NaCl) supplemented with protease inhibitors and 10 µL were kept as input. Diluted lysate was divided and in each tube was added a different antibody and incubated overnight at 4°C in gentle rotation. Next day, Dynabeads Protein G (Invitrogen) were added and incubated at 4°C for 2 hours with gentle rocking. Then, the

beads were washed with Low salt wash buffer (20mM Tris-HCl pH 8.0, 150mM NaCl, 0.1% SDS, 2mM EDTA, 1% Triton X-100) and with High salt wash buffer (20mM Tris-HCl pH 8.0, 500mM NaCl, 0.1% SDS, 2mM EDTA, 1% Triton X-100). Subsequently, the beads were washed once with a LiCl solution (10mM Tris-HCl pH 8.0, 250mM LiCl, 1% NP-40, 1mM EDTA) and twice with TE Buffer (10mM Tris-HCl pH 8.0, 1mM EDTA). Elution was performed using Elution Buffer (0.5M $NaHCO_3$, 10% SDS) added to each IP and incubated for 15' at room temperature with gentle rocking; elution was performed twice.

Reverse cross-link was performed overnight at 65°C adding 12µL NaCl 5M. Samples were treated with 2µL proteinase K (10 mg/mL), 12µL EDTA (0.5M) and 6µLTris pH 6.5 (1M) for 1 hour at 45°C.

DNA was isolated with PCR purification kit (Qiagen) following the manufacturer's protocol. Then, qPCR was performed.

Antibodies used for ChIP were against H3K4me3 (Abcam-ab8580), H3K27Ac (ab4729-Abcam), RNA polymerase II phospho-S5 (ab5408- Abcam), total H3 (ab180727, Abcam), WDR5 (A302-429A-bethyl) and NELFe (F-9, SC377052-SCBT) IgG (as negative control, IgG mouse sc-2025-SCBT; IgG rabbit 2729-Cell Signaling).

Three biological replicates were conducted.

RIP (RNA ImmunoPrecipitation)

RIP was performed modifying Hendrickson et al. [133] protocol.

Briefly, cells were collected and fixed in 0.1% formaldehyde (Sigma-Aldrich) solution for 10' with gentle rotation at room temperature. Cross-link was quenched with 0.125mM glycine for 5' with gentle swirl. Cells pellet was washed twice with PBS and resuspended in Nuclear Isolation Buffer (1.28M sucrose, 40mM Tris-HCl pH 7.5, 20mM $MgCl_2$, 4% Triton X-100)

supplemented with protease and RNase inhibitors and kept on ice for 20'. After centrifugation, nuclei were resuspended in RIP Buffer (150mM KCl, 25mM Tris-HCl pH7.5, 5mM EDTA, 0.5mM DTT (Sigma-Aldrich), 0.5% NP-40) supplemented with protease and RNase inhibitors and sonicated for one cycle 30"ON - 30"OFF. After sonication, nuclei debris were spinned and supernatant was kept. 10% of lysate were used for the input sample and $5x10^6$ cells were used for each IP with 6µg of NELF-e (F-9, SC377052 SCBT), 4µg of WDR5 (A302-429A, Bethyl) or BRD4 (A301-985A50, Bethyl) antibodies and the relative IgG control (mouse IgG SC2025 SCBT and rabbit IgG 2729S Cell Signaling). After overnight incubation with gentle rocking at 4°C, 20µl of Dynabeads protein G were added to each IP and incubated for 2 hours and 30' at 4°C in rotating wheel. IPs were subsequently washed twice with RIP wash buffer (150mM KCl, 25mM Tris pH7.5, 5mM EDTA, 0.5% NP-40, 0.5mM DTT) supplemented with protease and RNase inhibitors. Reverse cross-link was performed adding to each IPs and input samples, diluted to 1X the 3X reverse-crosslinking buffer (3X PBS (without Mg^{2+} or Ca^{2+}), 6% N-lauroyl sarcosine (Sigma-Aldrich), 30mM EDTA), 15mM DTT (added fresh), 10µl of proteinase K and RNase inhibitors for 1 hour at 42°C and 1 hour at 55°C. Supernatant was collected by Magna GrIP magnetic separation, Trizol was added and RNA was isolated as previously described.

At least three biological replicates were conducted for each cell line.

RNA Sequencing and bioinformatic analysis

For RNA-seq analysis, RNA was extracted using Trizol from cells pellet of TPC1 treated with gapmeRs against l-RAIN or control-Oligos and TPC1 treated with siRNA against RUNX2 or scramble. RNA quality and quantity were assessed by Bioanalyzer using Agilent RNA 6000 nano kit and by Nanodrop respectively. RNA-seq libraries were prepared using the TruSeq Stranded

mRNA Sample Preparation Kit (Illumina) starting from 1µg of RNA. Next generation sequencing was performed using NextSeq500 platform (Illumina). We loaded the pooled libraries in a 150 cycles High Output cartridge, in order to obtain a minimum of 20 million of sequencing reads for each sample replicate.

The bioinformatic data analysis included sequence adapters removal, that was performed by Trimmomatic, quality checks, performed using FastQC, and RNA sequences alignment, performed using STAR. After that, reads count and normalization were conducted applying Cufflink RNA-Seq workflow. Differential gene expression was calculated by Cuffdiff pipeline as fold-change (TPC1 treated with l-RAIN gapmeRs vs control-Oligos and TPC1 treated with RUNX2 siRNA vs scramble). Genes with a p-value < 0.05 were considered significantly deregulated. Next, the results of these two analyses were merged, in order to identify genes specifically deregulated by l-RAIN or commonly affected by l-RAIN and RUNX2 down-regulation. Bioinformatic data analyses were performed using R software (version 3.4.2). RNA-seq results investigation was conducted by Kegg pathways enrichment analysis.

Statistical analysis

Statistical analysis was performed using GraphPad Prism Software (GraphPad). Statistical significance was determined using the Student's t-test.

gapmeRs	
negative control	AACACGTCTATACGC
l-RAIN	CTATGATTAGAACGTC

siRNA	
negative control	Ambion cat.4390847
RAIN common	AAAGAAGUCAGUUAAAAUCAG
NELFe	Ambion cat.4392420 ID s15489
c-Jun	Ambion cat.4392420 ID s7660

ChIRP probes	
RAIN#1	AAGCCATAACAGCCCTAAAG
RAIN#2	TACACCATGTGAGTGACCAT
RAIN#3	GTTGTGACAGTGCTATTGAC
RAIN#4	CTTTGACCCACAGTACTACT
RAIN#5	AATGGCAATGCACACTGGTT
RAIN#6	TGCTACCAAGAGGAAGTCTA
RAIN#7	ATTGACCTTAAAGGGCCTAG
RAIN#8	CTTGGACCTTGGGATACTAA
LacZ#1	CCAGTGAATCCGTAATCATG
LacZ#2	GTAGCCAGCTTTCATCAACA
LacZ#3	AACGAGACGTCACGGAAAAT
LacZ#4	ACCATTTTCAATCCGCACCT
LacZ#5	AGACGATTCATTGGCACCAT
LacZ#6	ATTTAGCGAAACCGCCAAGA
LacZ#7	TTTACCTTGTGGAGCGACAT
LacZ#8	TAAGGTTTTCCCCTGATGCT

qRT-PCR primers	
RUNX2 F	GTGCCTAGGCGCATTTCA
RUNX2 R	GCTCTTCTTACTGAGAGTGGAAGG
RAIN common F	CTCAAAGCAAGTCGCCAAAG
RAIN common R	CCTGTGATCTGCCCTTTAGC
l-RAIN F	TCTTTCTTTAGGGCTGTTATGG
l-RAIN R	AGGAGGAACACTGGGGTCTC
l-RAIN RIP F	ACCAAAAGGACATCTGCACA
l-RAIN RIP R	ACCTCCTAACCTTGCACACA
Cyclophilin F	GACCCAACACAAATGGTTCC
cyclophilin R	TTTCACTTTGCCAAACACCA
GUSB F	TTGAGCAAGACTGATACCACCTG
GUSB R	TCTGGTCTGCCGTGAACAGT
XIST F	GGCCAAGCTCCAGCTAATCT
XIST R	CGTCAAAGGGAATGGATCAC
C-Jun F	TGACTGCAAAGATGGAAACG
C-Jun R	CAGGTCATGCTCTGTTTCA
NELFe F	AAGTCAGGAGCCATCAGTGC
NELFe R	CTGGAAAGTGGGGACTGGTC
WDR5 F	AGTGCCTGAAGACGCTCATC
WDR5 R	TGGCGGCCAGGATGTATTTG
CCNE2 F	TGCAGAGCTGTTGGATCTCTGTG
CCNE2 R	GGCCGAAGCAGCAAGTATACC
RUNX2 P2 F	ACCATGGTGGAGATCATCG
RUNX2 P2 R	GGCAGGGTCTTGTTGCAG
enh3 F	GCTGGGAAGATAGCCAAGAA
enh3 R	CCTTGCATCAGTTCCACAGA
enh11 F	CCCAAACCCCAAAGCAGAGA
enh11 R	CCCAAGTTCTCACCAGGCAT
enh13 F	GTGGAGTGGAGAGAGGAGAA
enh13 R	TGGCTTCATCTCACCCTCAG
ctrl- ChIP F	TCTCAAGGTGCCTGTCTGC
ctrl- ChIP R	TGAAGTTTGGCCTCTGGTCT
MALAT1 F	TGTTGGCACGAACACCTTCA

MALAT1 R	TGGCCTACTCAAGCTCTTCTG
KCNQ1ot1 F	GGCTACGCCACAGGTGAAA
KCNQ1ot1 R	GTCTGCTGGCTTGTGTGTTG
5.8S F	GGTGGATCACTCGGCTCGT
5.8S R	GCAAGTGCGTTCGAAGTGTC
GAPDH F	CAATTCCCCATCTCAGTCGT
GAPDH R	GCAGCAGGACACTAGGGAGT
- 1100 RUNX2 TSS F	CGCTCCTTCATCCTCTCGAC
- 1100 RUNX2 TSS R	AAAATGCTTCCGTGGCTGT
- 500 RUNX2 TSS F	CTCTCTGGTGTCTCGGCTTC
- 500 RUNX2 TSS R	CAGACTAGGGGCAATCTCGC
TSS RUNX2 F	TGGACTGCTGAACCCACAC
TSS RUNX2 R	TGAGTTTGCAGCTTGGAATG
+700 RUNX2 TSS F	ACCATGGTGGAGATCATCG
+700 RUNX2 TSS R	GGCAGGGTCTTGTTGCAG
+1300 RUNX2 TSS F	CTCTCACCCGCTTCCCTCA
+1300 RUNX2 TSS R	CCAGGACCGCTGAACTCTG

RESULTS

RUNX2 and RAIN expressions are correlated

Our preliminary evidence indicated that RAIN was co-regulated with RUNX2 in both thyroid and breast cancer cells. Thus, to explore a potential correlation between these two transcripts, we analyzed their expression in a panel of epithelial cancer cell lines, including lung-, prostate-, breast- and thyroid-derived cancer, which are the tumors in which RUNX2 has been shown to play relevant functions (Fig.11 a, b).

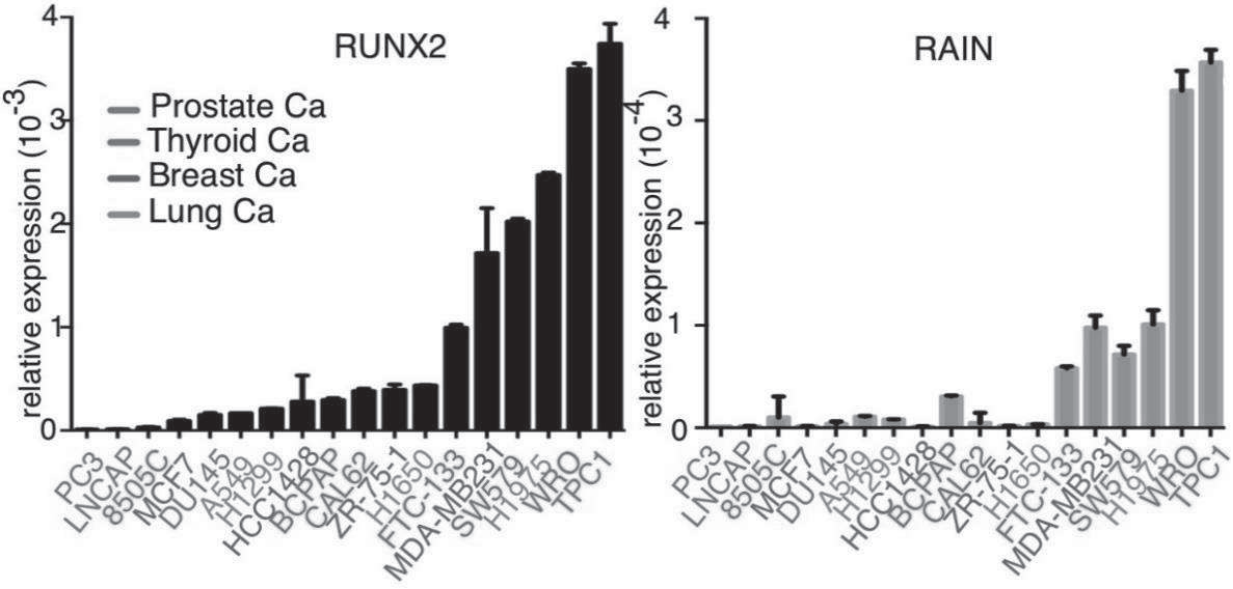

Figure 11 RUNX2 and RAIN expression in cancer cell lines Quantification of the expression of RUNX2 (a) and RAIN (b) in epithelial cancer cell lines: lung (green- H1299, H1975, H1650, A549), prostate (blue- PC3, LNCaP, DU145), thyroid (red- 8505C, BCPAP, TPC1, CAL62, FTC133, SW579, WRO) and breast (grey- MCF7, HCC1428, ZR-75-1, MDA-MB 231).

As shown in figure 11, RUNX2 and RAIN expression was significantly positive correlated, with a correlation coefficient R^2 =0.8752 and a p-value<0.0001 (Fig.12).

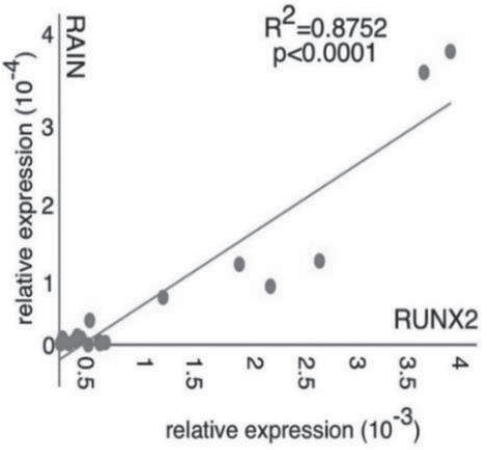

Figure 12 RUNX2 and RAIN correlation The graph shows the relative expression of RUNX2 and RAIN in a panel of cancer cell lines.

We have previously demonstrated that RUNX2 overexpression is associated with cancer development and aggressiveness in thyroid cancer. We extracted RNA from 26 thyroid cancer patients and we compared RUNX2 and RAIN expression in normal and tumor tissue. RUNX2 and RAIN were both overexpressed in cancer tissue as compared with normal thyroid. Next, we correlated the expression of RUNX2 and RAIN in tumor samples. We confirmed that also *in vivo* the expression of these transcripts was significantly correlated (fig 13 a, b).

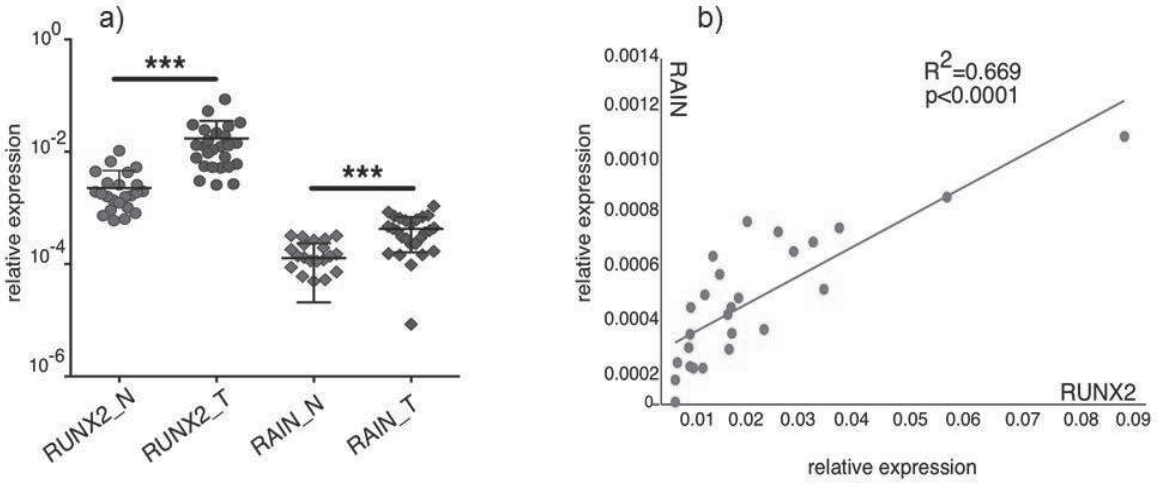

Figure 13 RUNX2 and RAIN expression in patient's samples We compared the expression of these transcripts in normal and tumor samples of the same patient (a). We also assessed the correlation of RUNX2 and RAIN in tumor samples and we obtained a significant positive correlation (b). *** p-value <0.001

Analysis of RAIN stability

ENH-associated lncRNAs are averagely stable transcripts, more stable than eRNA but less than the mRNAs to which they are associated. We assessed the stability of RAIN, along with RUNX2 and c-MYC mRNA stability, by treating TPC1 and MDA-MB 231 cell lines with 1µg/ml of actinomycin D to interfere with mRNA synthesis. We collected RNA at different time points: 0, 30', 2 hours, 4 hours, 8 hours, 12 hours, 24 hours. We compared the expression of each RNA at each time point with the expression at the corresponding time 0. We observed that RAIN is less stable than RUNX2 but more stable of c-MYC which is known to be rapidly degraded (Fig.14). We used KCNQ1ot1 and MALAT1, two lncRNAs, as control.

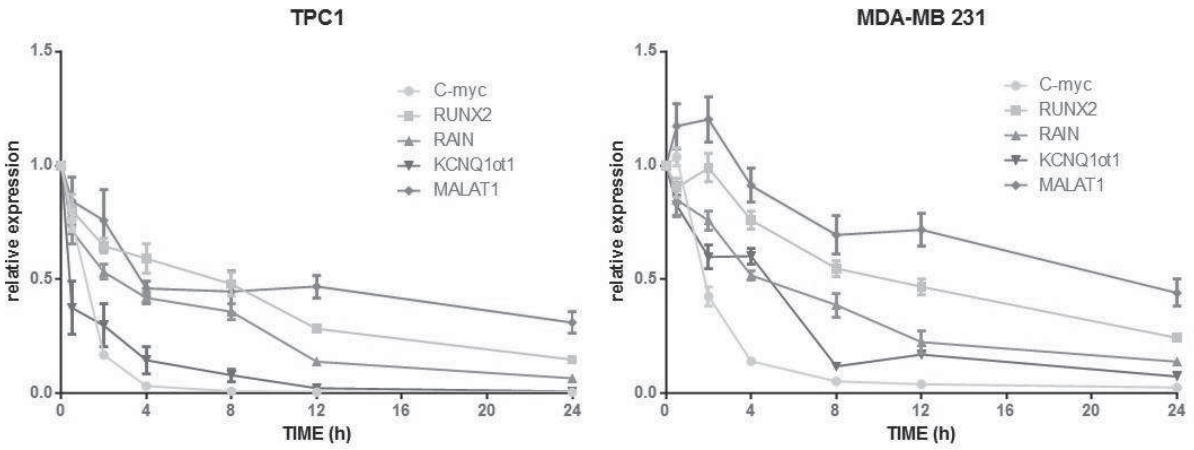

Figure 14 RAIN stability The RNA stability has been evaluated through qRT-PCR analysis, administrating actinomycin D to cultured cells and extracting RNA at different time points.

We also calculated the half-life rate of each RNA using linear regression analysis in GraphPad Prism Software. The half-life of RAIN was about 7 hours in TPC1 cells and 9 hours in MDA-MB 231, while RUNX2 half-life was 11 hours and 13 hours, respectively (Fig.15). These differences were probably due to the different replication rate of the analysed cell lines (TPC1 cells are more actively proliferating than MDA-MB 231) and to the lower expression of RAIN

and RUNX2 in MDA-MB 231 cells compared to the expression in TPC1 cells (Fig.11).

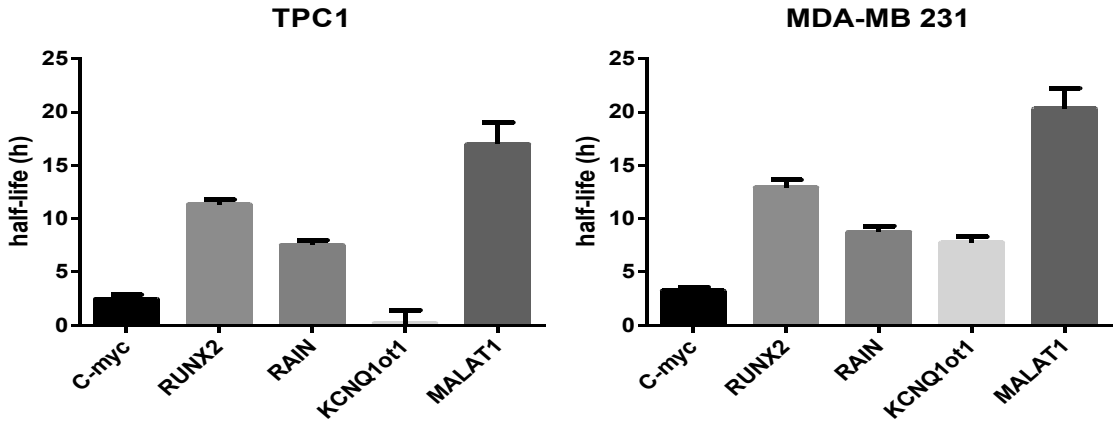

Figure 15 Half-life rate calculation 50% of RNA decay was calculated through linear regression analysis of relative expression of each RNA using GraphPad Prism software.

RAIN long isoforms are chromatin-associated

Because the different cellular localization of lncRNAs is associated with diverse function, we next investigated the subcellular localization of RAIN. To this end, we performed cell fractioning and separated cytosolic, nuclear and chromatin fractions. We performed western blot to ensure we have correctly separated the different phases. We controlled total RNA-PolII and α-Tubulin, as nuclear and cytosolic markers respectively. Then, we extracted RNA and analyzed RAIN distribution in the different fractions. We used two different primer pairs that recognized different regions of RAIN. One pair recognized the common central region and detected both the long and the short RAIN isoforms. The second pair was specific for the long isoform. The chromatin-associated lncRNA XIST was used as control.

Noticeably, the long RAIN isoform (l-RAIN) was largely enriched in the nuclear and chromatin fractions, while the short RAIN isoform distributed homogeneously in the cytoplasm and nucleus (fig.16).

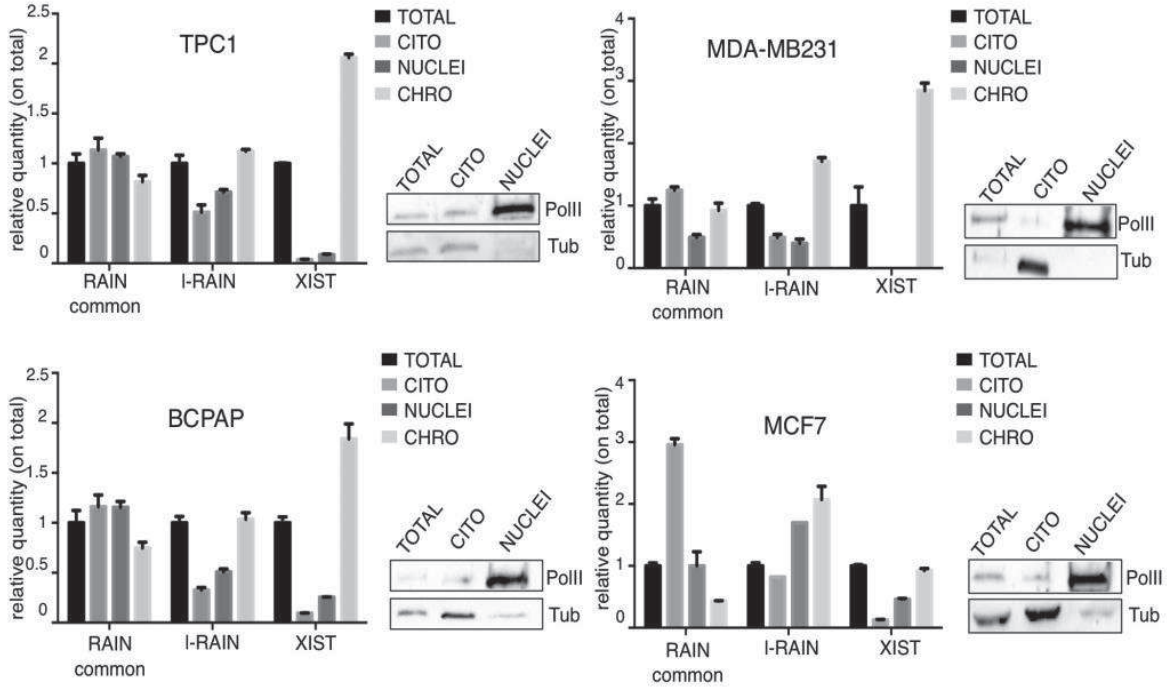

Figure 16 RAIN subcellular localization TPC1, MDA-MB231, BCPAP and MCF7 cell lines were fractionated, RNA and proteins were extracted in subcellular fractions and analyzed for the presence of RAIN. RAIN common means that qRT-PCR primers recognized the common central region, l-RAIN means that primers recognize only the long isoform.

We confirmed this different subcellular localization in TPC1 cells, through fluorescent *in situ* hybridization, using two different probes: one recognizing the common region and one only the long isoform (fig.17).

Figure 17 RAIN localization by *in situ* hybridization FISH analysis were performed to confirm the subcellular localization of RAIN in TPC1 cell line. RAIN-common means that probes recognize the common central region, while l-RAIN means that probes recognized only the long isoform.

Knockdown of RAIN impairs RUNX2 expression

We showed that RAIN is a chromatin-associated lncRNA and that its expression correlates with RUNX2. Thus, we reasoned that RAIN may take part to RUNX2 regulation. To test this hypothesis, we used two different approaches: siRNA (small interfering RNA) and gapmeRs (Fig.18).

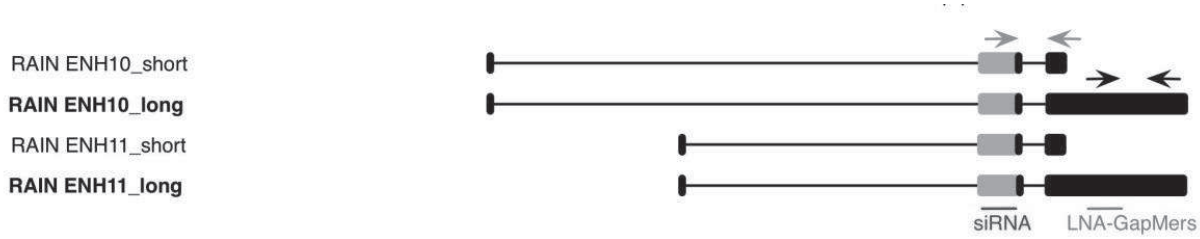

Figure 18 siRNA, gapmeRs and qRT-PCR primers localization siRNA probes were designed to recognize the common central region. While gapmeRs were designed to recognized only the long isoform because they are in the nucleus and gapmeRs are able to target nuclear lncRNA better than siRNA. Grey arrows indicate qRT-PCR primers that recognize common region (RAIN common), black arrows indicate the primers that recognize the long isoform (l-RAIN).

First, we used, siRNA to knockdown both short and long RAIN isoforms. The target RNA degradation mechanism of siRNA, is based on the perfect complementarity with target mRNA. SiRNA and target RNA coupled and are cleaved by RISC (RNA-induced silencing complex) in the cytosol.

TPC1 and MDA-MB 231 cells were transfected with siRNA targeting RAIN common region or with scramble oligos, as control, and RAIN and RUNX2 expression was assessed by qRT-PCR 24 hours after transfection. Noticeably, in both cell types RAIN silencing reduced RUNX2 expression. (Fig.19).

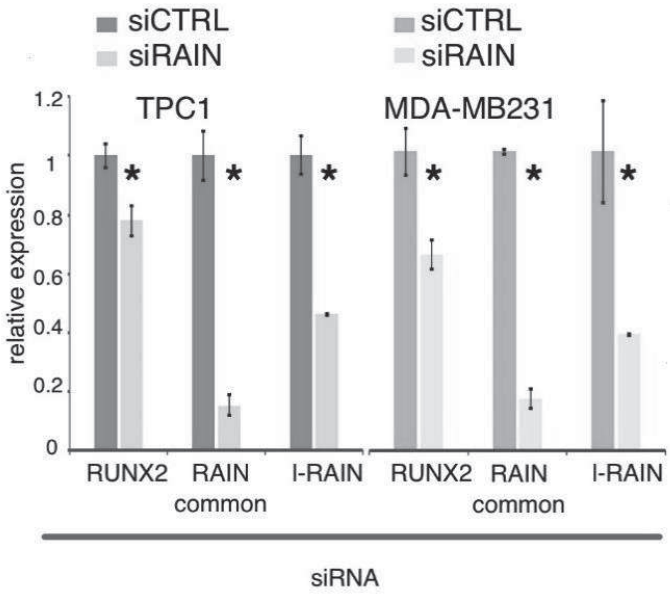

Figure 19 RAIN and RUNX2 expression after siRNA transfection TPC1 and MDA-MB 231 cell lines were reverse transfected with 25 nM of siRNA against the common region. After 24 hours of transfection, RNA was extracted and RAIN and RUNX2 expression was quantified by qRT-PCR. * p-value<0.05.

Next, we used gapmeRs, oligos with nucleotide LNA (locked nucleic acid) modification, to target the long isoforms of RAIN. These oligos recognized and paired to target RNA recruiting RNase H, an endonuclease present both in the cytosol and in the nucleus, that selectively degrade RNA of the DNA-RNA heteroduplex. Using nuclear RNAse to degrade target RNA, gapmeRs should be more efficient than siRNA to target the chromatin-associated lncRNA.

We transfected TPC1 and MDA-MB 231 with gapmeRs or scramble oligos and we observed a significant down-regulation of RAIN and RUNX2, with a more efficient downregulation of l-RAIN, that is the isoform associated with the chromatin (Fig.20). We also observed a consequent more effective downregulation of RUNX2.

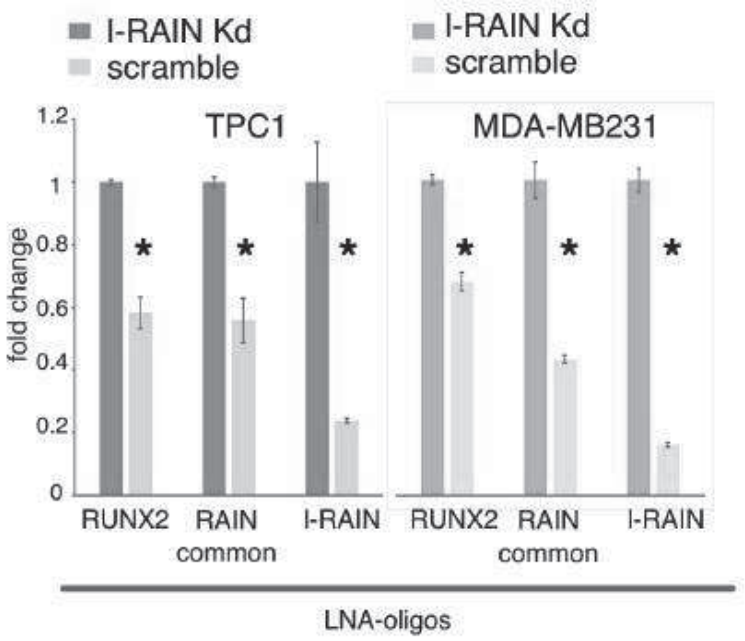

Figure 20 RAIN and RUNX2 expression after gapmeRs transfection TPC1 and MDA-MB 231 cell lines were reverse transfected with 25 nM of gapmeRs against the 3'-long region. After 24 hours of transfection RNA was extracted and RAIN and RUNX2 expression was quantified by qRT-PCR. * p-value<0.05

To ensure that was the RAIN interference that lead to RUNX2 down-regulation, we also transfected our cells with siRNA against RUNX2 and we did not observe any difference in RAIN expression.

l-RAIN interacts with RUNX2 locus

Next, we investigated the mechanism by which RAIN controls RUNX2 expression in cancer cells. First, we used ChIRP approach to define whether RAIN interacts with RUNX2 P2 promoter. To immunoprecipitate RAIN, we designed eight probes that mapped on the long 3'-end. As negative control, we designed eight probes that mapped on LacZ, which is not expressed in mammalian cells. We performed our experiments on both TPC1 and MDA-MB 231 cell lines. After IP, regions specifically bounded to target RNA were measured by qRT-PCR. In both cell lines, RAIN significantly interacted with the RUNX2 P2 promoter in a region spanning from the TSS and exon 1, 700bp

downstream of the transcription starting site (Fig.21). Furthermore, we used GAPDH and 5.8S promoter regions as negative control for RAIN interaction.

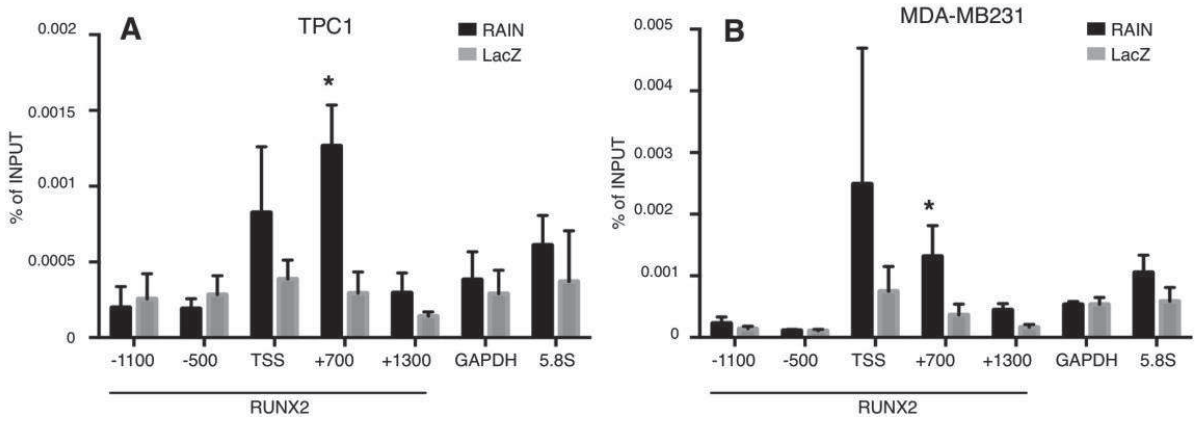

Figure 21 ChIRP analysis of RAIN and P2 interaction TPC1 (a) and MDA-MB 231 (b) cell lines were used for ChIRP approach to demonstrate that RAIN was able to interact with the genomic locus of RUNX2. L-RAIN interacted with P2 promoter of RUNX2-isoform I: we mapped the interaction site between TSS and 700bp downstream of the TSS. Promoters of GAPDH and 5.8S were used as negative control. * p-value<0.05.

The interaction was specific for RAIN probes, since RAIN was not immunoprecipitated with LacZ probes (Fig.22).

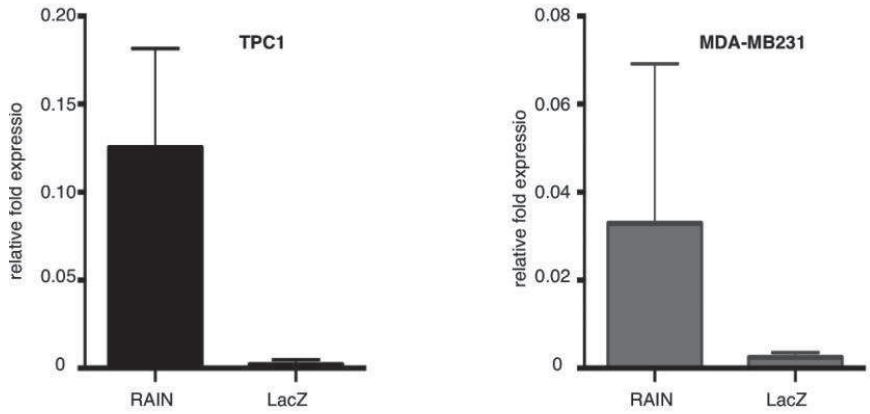

Figure 22 ChIRP control TPC1 and MDA-MB 231 RNA was extracted, during ChIRP protocol, to evaluate the correct procedure, to exclude that the ChIRP probes recognize and pull-down the wrong RNA. RAIN and LacZ IPs were analyzed for the presence of RAIN.

l-RAIN interacts with WDR5 and NELFe

LncRNAs have the ability to interact with histone modification complexes to regulate the transcription of protein-coding genes. WDR5 is a subunit of MLL1/MLL complex that mediate the trimethylation of Lys-4 of histone H3 (H3K4Me3) and gene activation.

We used a RIP approach to evaluate the interaction of l-RAIN with WDR5 on TPC1 and MDA-MB 231 cell lines. We observed that l-RAIN interacted with WDR5 in a significant manner, while it did not interact with BRD4. We used IgG as negative control (Fig.23).

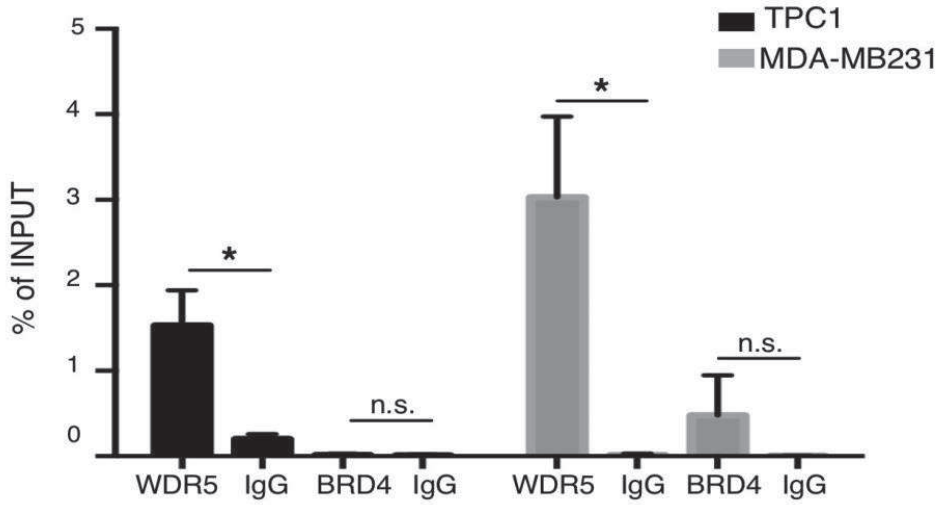

Figure 23 RIP approach to evaluate l-RAIN interaction TPC1 and MDA-MB 231 cell lines were used for RIP analysis to evaluate the potential interaction of l-RAIN with histone modification complex. IgG were used as negative control of assay. * p-value<0.05, n.s. means not-significative.

As previously mentioned, NELF is a protein complex, that inhibits elongation promoting RNA-PolII pausing 0-60 nucleotides downstream of TSS, consequently inhibiting the transcription. It has been demonstrated that NELF can interact with eRNA, but the interaction with lncRNA it has been only hypothesized. Thus, we used RIP to test if RAIN was able to associate also with NELFe in TPC1 and MDA-MB231 cells. NELFe is the RNA-binding subunit of the NELF complex and its activity is critical for NELF function. We showed

that IPs with NELFe antibody resulted in a significant enrichment of l-RAIN indicating that this lncRNA can also interact with the NELF complex (Fig.24).

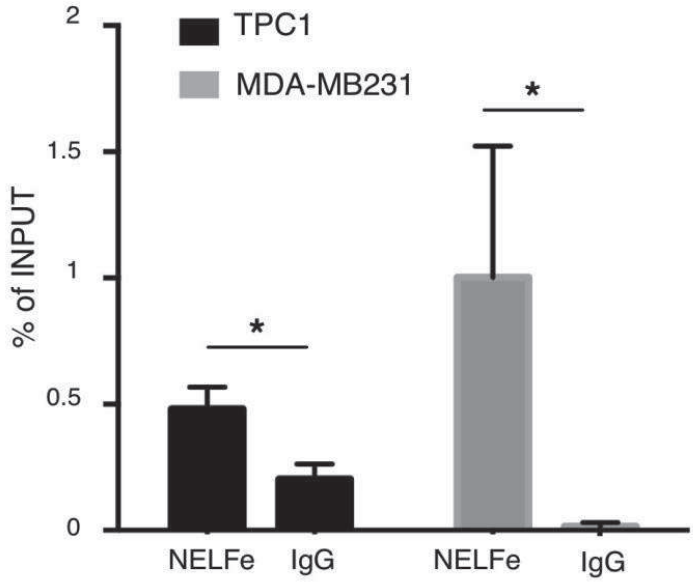

Figure 24 RIP approach to evaluate l-RAIN interaction TPC1 and MDA-MB 231 cell lines were used for RIP analysis to evaluate the potential interaction of l-RAIN with NELFe. IgG were used as negative control of assay. * p-value<0.05

Down-regulation of l-RAIN decreases the activation of P2

Based on our data, we hypothesized that l-RAIN positive activity on RUNX2 expression was mediated by the interaction with WDR5. To verify this hypothesis, we performed ChIP experiments on TPC1 cell line, after down-regulation of l-RAIN through gapmeRs transfection.

First, we investigated whether knockdown of l-RAIN affected the recruitment of WDR5 on the P2 promoter. As expected, WRD5 binding on the P2 promoter was significantly reduced upon l-RAIN_Kd (knock down) confirming that RAIN was required for the recruitment of this factor on the P2 promoter. Then, we assessed whether the inhibition of WDR5 by l-RAIN_kd had consequences on the transcriptional activity of the P2 promoter by assessing both H3K27 acetylation enrichment and RNA-PolII phospho-5S binding. We examined this RNA-PolII modification because serine 5 phosphorylation is associated with

the initiation of transcription and it is associated with the transcription complex downstream of the TSS of the genes. As well, since WRD5 mediates trimethylation of H3K4, we also analyzed the amount of this modification on the P2 promoter 48 hours after l-RAIN_Kd, that is the time in which we observed the most efficient l-RAIN down-regulation. Noticeably, reduction in the levels of l-RAIN significantly reduced both the RNA-PolII phospho-5S binding and H3K27Ac levels on the P2 promoter; as well, the amount of H3K4me3 on the P2 promoter was reduced upon l-RAIN_Kd, while the overall amount of H3 was not modified (Fig.25).

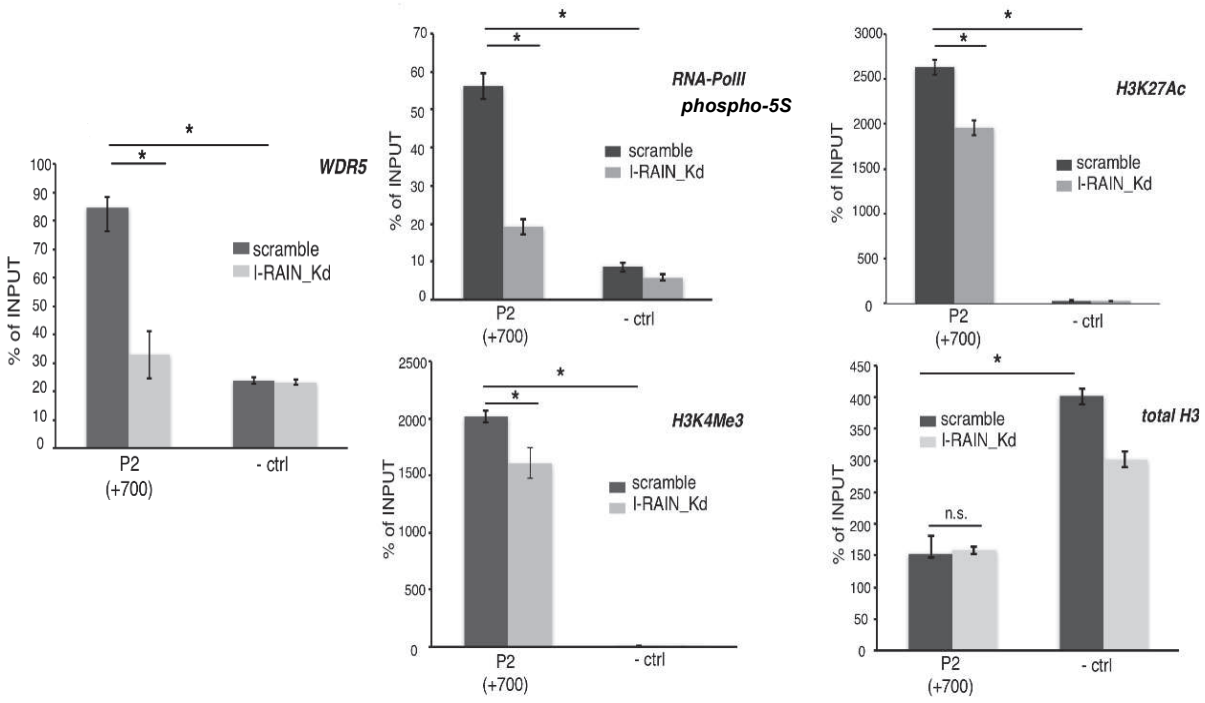

Figure 25 ChIP analysis on P2 of RUNX2 after l-RAIN down-regulation We performed ChIP analysis on TPC1 cell line 48 hours after reverse transfection with 25nM of gapmeRs against l-RAIN. We focused on P2 promoter (primers are located 700bp downstream of the TSS) and used an intronic region downstream of P2 as negative control. * p-value<0.05

By contrast, l-RAIN_kd have not relevant effect on the chromatin organization of RUNX2 ENH3, ENH11 and ENH13 (Fig.26).

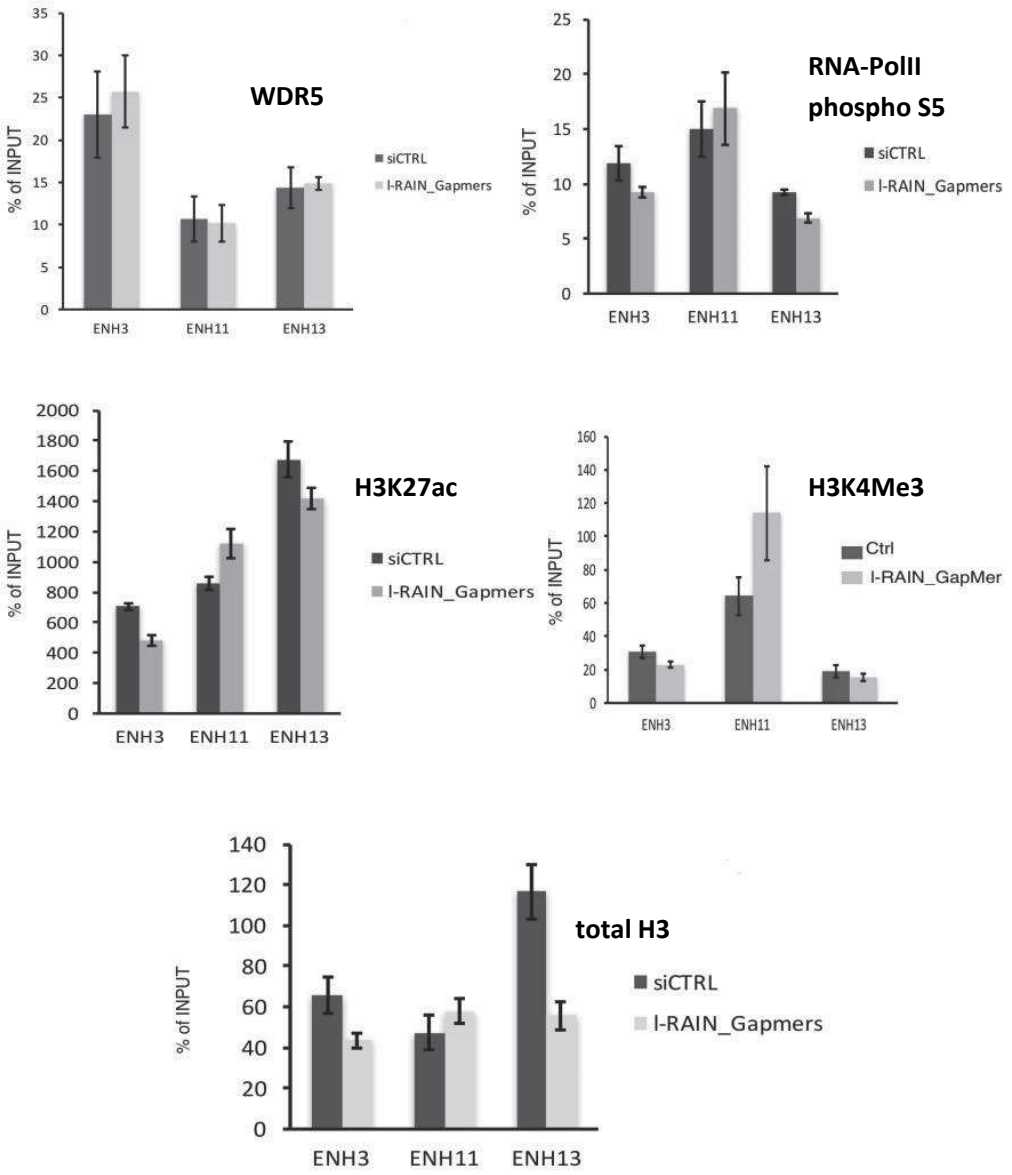

Figure 26 ChIP analysis on ENHs region of RUNX2 after l-RAIN down-regulation We performed ChIP analysis on TPC1 cell line after reverse transfection with 25nM of gapmeRs against l-RAIN. We focused on ENH3, ENH11 and ENH13 regions.

We used ChIP approach also to evaluate the effect of l-RAIN down-regulation on NELFe interaction with RUNX2 P2. We divided the P2 locus to obtain the precise region where NELFe interact with the promoter (Fig.27).

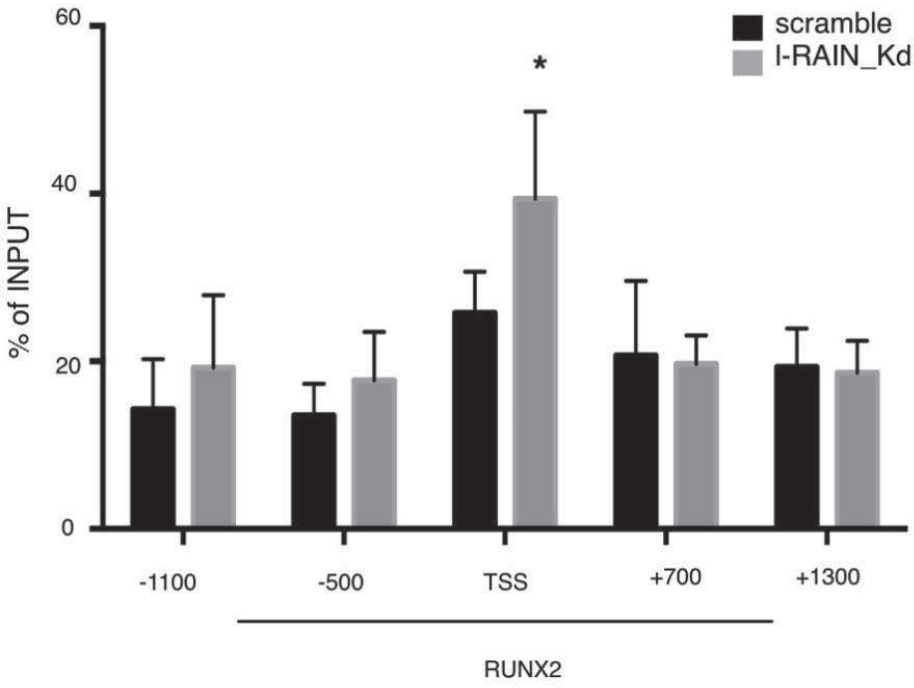

Figure 27 ChIP analysis on RUNX2 P2 to evaluate the interaction of NELFe TPC1 cell lines were reverse transfected with 25 nM of gapmeRs against l-RAIN and after 48hours we employ ChIP approach. P2 promoter region was subdivided to locate the region of interaction between NELFe and RUNX2. * p-value<0.05

We observed that NELFe interact with RUNX2 TSS and that the l-RAIN downregulation lead to an increasing interaction of NELFe with this region.

Down-regulation of NELFe leads to up-regulation of RUNX2

Since we showed that l-RAIN can also interact with NELFe, we explored the effect of NELFe silencing on RUNX2 expression. If l-RAIN binding to NELFe restraining the inhibitory effect of the NELF complex on transcription elongation, we may expect that NELFe silencing is associated with a positive effect on RUNX2 expression. We used siRNA to down-regulate NELFe in both TPC1 and MDA-MB231. We controlled the efficacy of down-regulation both by qRT-PCR and western blot analyses (Fig.28).

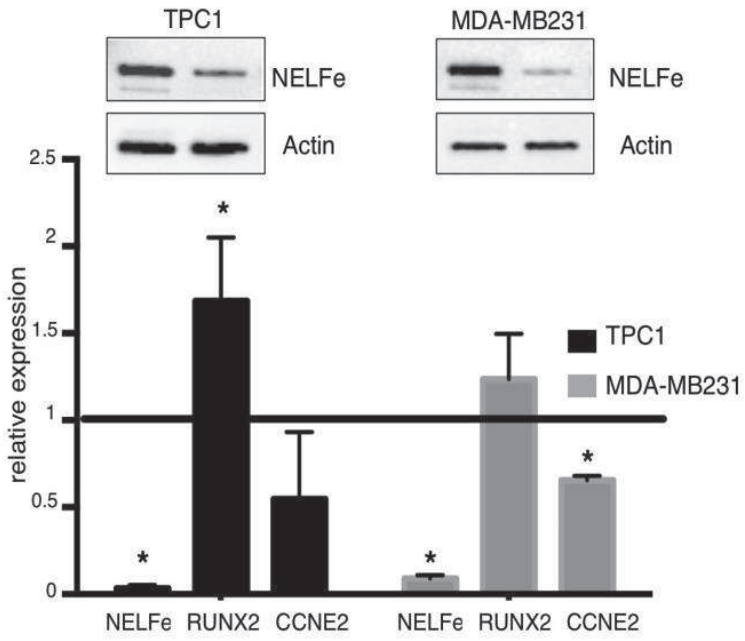

Figure 28 NELFe down-regulation We used siRNA approach to down-regulate NELFe in TPC1 and MDA-MB 231 cell line. We reverse transfected cells with 50nM of siRNA against NELFe and after 48 hours we extracted RNA and quantified RUNX2 and CCNE2 (positive control) expression. * p-value<0.05.

We observed that the down-regulation of NELFe leads to the upregulation of RUNX2 confirming the hypothesis that RAIN may act as decoy of NELF complex restraining its function. CCNE2 (Cyclin E2), which expression requires NELFe, was used as control of the functionality of the assay.

l-RAIN down-regulation affects cells' migration and proliferation

We demonstrated that l-RAIN is required for RUNX2 expression. Since RUNX2 over-expression in cancer cell is correlated with cancer development and progression, we hypothesized that RAIN may also affect aggressiveness of thyroid and breast cancer. In a previous work, we demonstrated that RUNX2 silencing lead to impairment of migration and invasiveness of thyroid cancer cells and that its overexpression increased these phenomena [48]. Thus, we analyzed the effect of RAIN silencing on proliferation and migration of TPC1 and MDA-MB231 cells.

We transfected cell lines with gapmeRs and count cells number every 24 hours, from 0 to 96 hours (Fig.29).

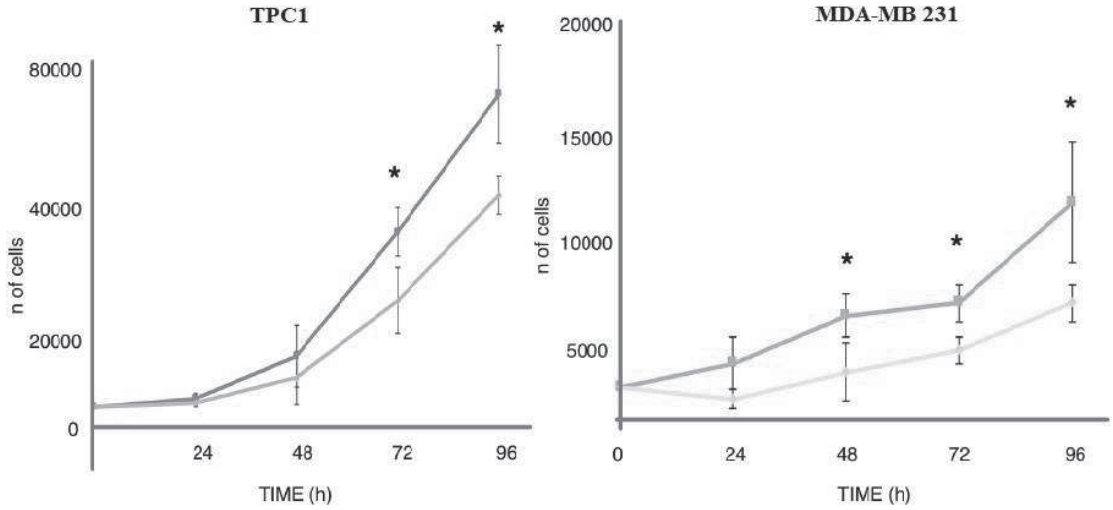

Figure 29 Proliferation assay after l-RAIN down-regulation TPC1 and MDA-MB 231 cell lines were reverse transfected with 25nM of gapmeRs against l-RAIN and 24 hours later, cells were plated in 96-well plate. Every 24h hours cells were detached and counted. *p-value<0.05

Transfected cells were also tested for their ability to wound healing in scratch test. Cells were photographed at 0, 6 hours, 12 hours and 24 hours (Fig.30).

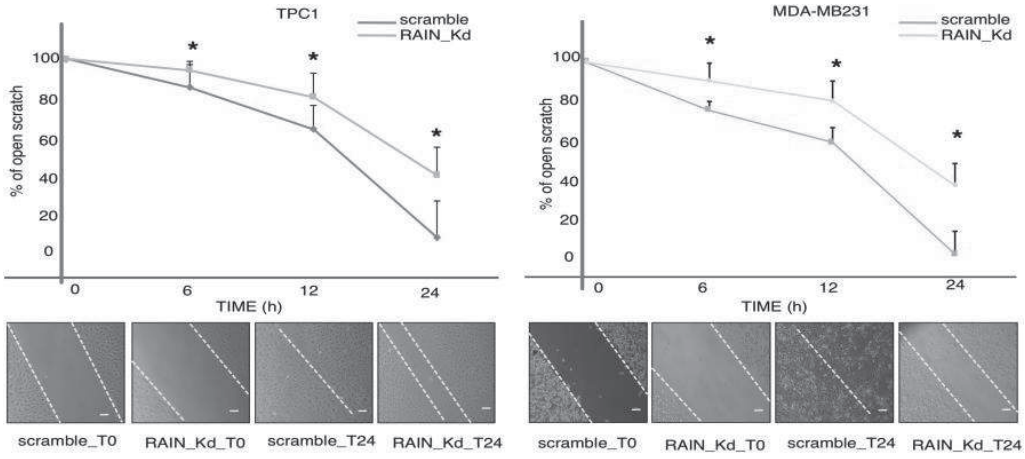

Figure 30 Wound healing assay after l-RAIN down-regulation TPC1 and MDA-MB 231 cell lines were reverse transfected with 25nM of gapmeRs against l-RAIN and 24 hours later, cells were plated in 6-well plate. After cells attachment, we scratched the well and photographed the scratches at different time point. We measured the scratches and compared them to T0. * p-value<0.05.

RUNX2 and RAIN down-regulation and analysis of their interplay in down-regulated cells

We have demonstrated that RUNX2 and RAIN are co-regulated and co-expressed in cancer, both *in vitro* and *in vivo*. We have also demonstrated that RAIN interacts with WDR5 and NELFe to regulate the expression of RUNX2. Following these results, we wanted to assess if RAIN is able to regulate other pathways not associated with RUNX2. We performed two RNA-sequencing experiments to define the gene expression profile of cells in which either RAIN or RUNX2 were silenced. Comparing the expression profile of cells silenced for RAIN with the profile of scramble transfected cells, we identified a list of 706 differentially expressed gene: 224 genes were up-regulated and 482 down-regulated. As well, differential analysis of the gene expression profile of TPC1 cells transfected with siRUNX2 or siCTRL identified 1754 gene of which 574 were up-regulated and 1180 down-regulated. Merging these lists, we observed that 163 of the 706 RAIN target genes were also affected by RUNX2 silencing suggesting that these genes could represent RAIN indirect-RUNX2 mediated targets (Fig.31).

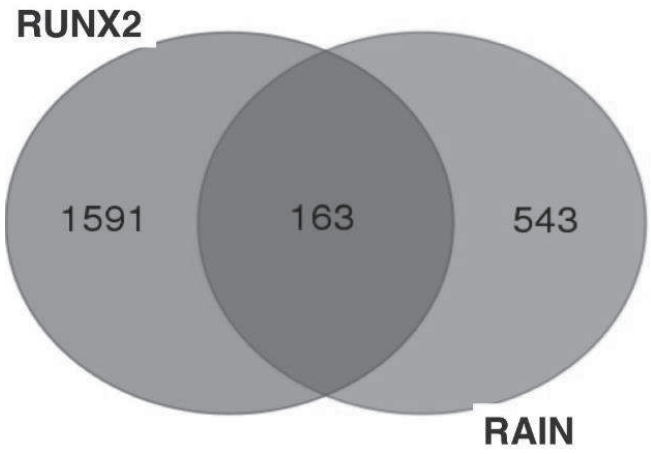

Figure 31 Diagrammatic representation of gene differential expressed in RNA-Seq analyses TPC1 cell line was reverse transfected with siRNA against RUNX2, gapmeRs against l-RAIN and control negative oligos.

We analyzed the list of RAIN-RUNX2 common genes using DAVID (Database for Annotation, Visualization and Integrated Discovery) searching for enriched pathways (KEGG pathway analysis) (Fig.32).

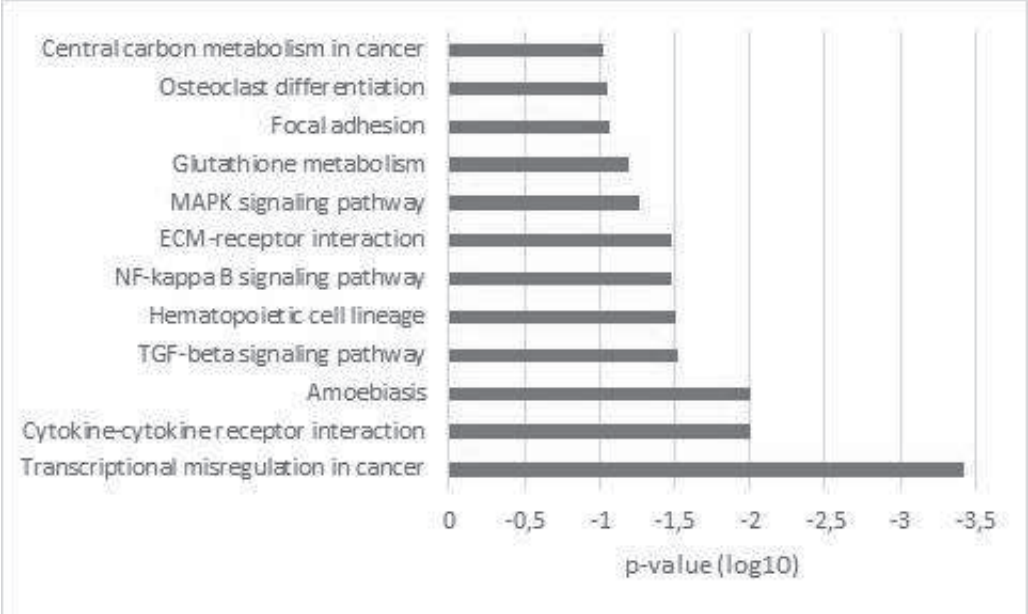

Figure 32 KEGG pathway analysis on functional annotation enrichment We submit our 163 differential-expressed genes on DAVID software and we focused on KEGG pathway functional annotation tool.

We observed that "transcriptional misregulation in cancer" was the most enriched pathway in common genes. Genes belonging of this pathway are: BCL2A1, CXCL8, HIST1H3E, IGFBP3, IL1R2, RUNX2, SPINT1 and TGFBR2. Other pathways enriched in this analysis are correlated to tumor aggressiveness and tumor microenvironment involvement, as cytokine-cytokine receptor interaction, MAPK, TGF-β and NF-kB pathways.

Among the 163 genes, 84 were down-regulated and 14 were up-regulated in both experiments. Focusing on genes down-regulated, we looked for the functional annotation on KEGG pathway and we obtained that the enriched pathways are "osteoclast differentiation" and "cytokine-cytokine receptor

interaction", strengthen the hypothesis that these are direct-RUNX2 mediated targets.

Next, we focused on those genes that resulted as RAIN specific (RUNX2 unrelated) targets. Of the 543 RAIN target genes 65% were down-regulated and 35% up-regulated, confirming an overall positive role of RAIN on gene expression. Of these genes 38 were on chromosome 6 and likely controlled *in cis* by this lncRNA. Among the 543 genes, there were also 9 long intergenic non-coding RNA (2 up-regulated and 9 down-regulated).

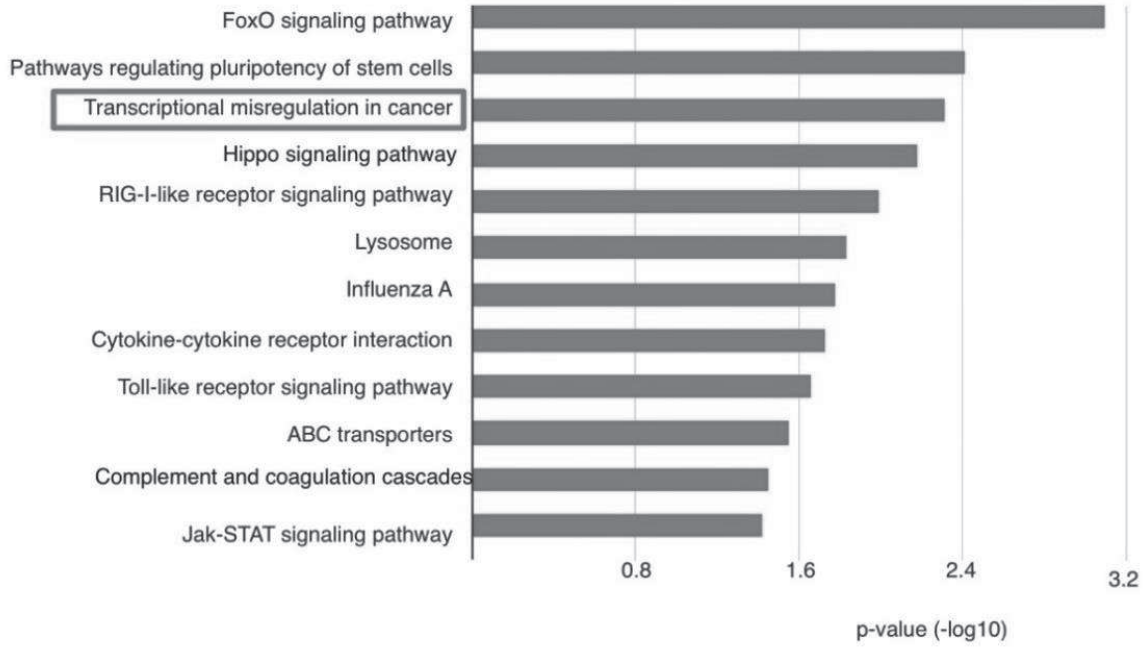

Figure 33 KEGG pathway analysis on functional annotation enrichment We submitted our 543 differential-expressed genes on DAVID software and we focused on KEGG pathway functional annotation tool.

DAVID and KEGG pathway analyses (Fig.33) reveled that RAIN-associated pathways were related to transcriptional misregulation in cancer and other cancer specific pathways, including the FOXO, Hippo and Jak-Stat pathways. We decided to focus on "transcriptional misregulation in cancer" because of the

high expression of RAIN in tumor samples. Genes belonging to this group are: MMP3, MYC, ETV6, AFF1, PLAT, EYA1, CDKN2C, BCL6, ID2, JUP, HPGD, PAX5 and RUNX2. In particular, some of these genes were down-regulated after l-RAIN_kd even more that RUNX2, suggesting a role of RAIN in direct regulation of other genes, beyond RUNX2 (Fig.34).

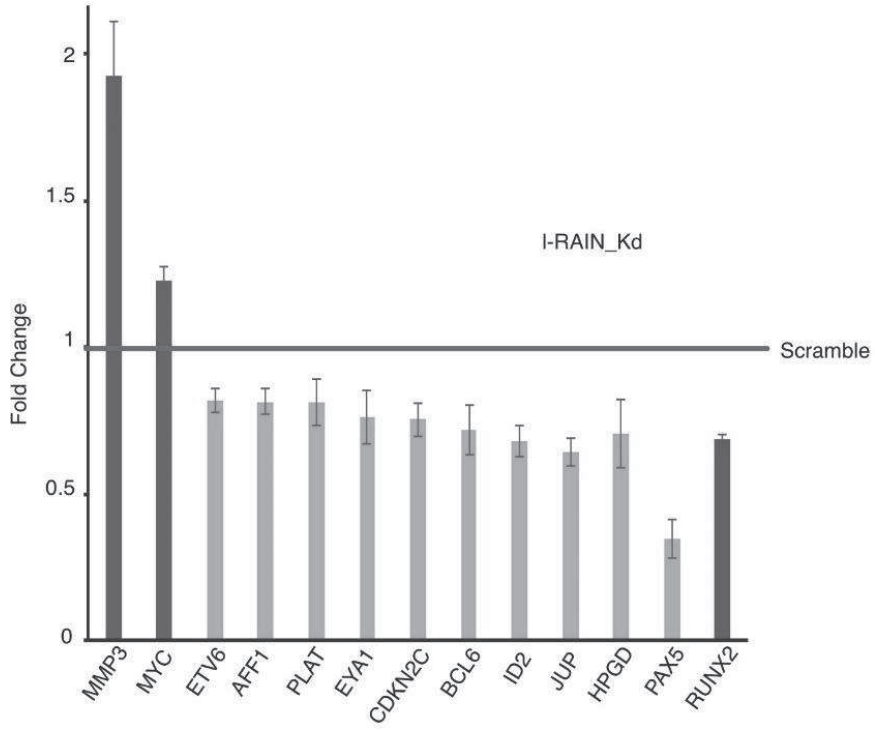

Figure 34 "Transcriptional misregulation in cancer" pathway enriched in l-RAIN_kd cells TPC1 cell line reverse transfected with gapmeRs against l-RAIN showed a differential expression of 13 genes. Relative expression is related to the expression of not-silenced TPC1 (scramble).

Further investigation and validation is needed to evaluate the impact of RAIN on these other target genes. We also would like to confirm the indirect-RUNX2 effects, down-regulating l-RAIN and expressing RUNX2 ectopically. However, these preliminary data suggest that the role of RAIN in cancer may be wider than the regulation of RUNX2 expression.

RAIN model of function

Thanks to our results, we hypothesized a model of action that can explain the mechanism of RAIN regulation on RUNX2.

RAIN is transcribed from RUNX2 ENH10 and ENH11 regions and it is able to interact with RUNX2 P2. Moreover, RAIN can recruit WDR5, bringing it in contact with the promoter of RUNX2-isoform I. WDR5 belong to the MLL complex that trimethylates H3K4 and enhance the transcription. Furthermore, RAIN can sequester NELFe, removing it to RNA-PolII. This detachment let the switch of RNA-PolII from paused to active state (Fig.35).

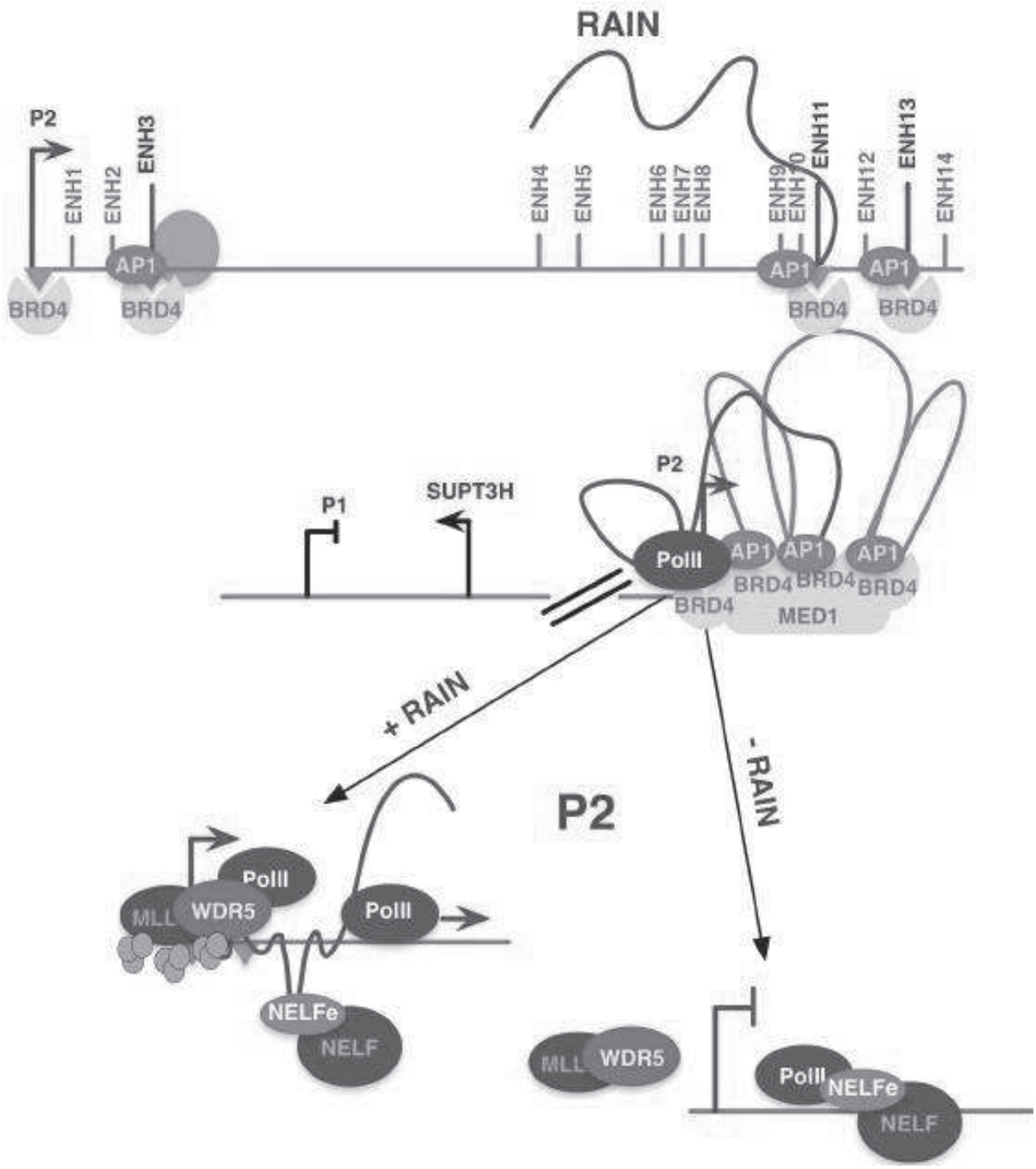

Figure 31 RAIN function model RAIN is transcribed from RUNX2 ENH10 and ENH11, then it collaborates to RUNX2 regulation interacting with WDR5 and NELFe enhancing RUNX2 transcription. Purple dots are H3K4 trimethylation, green triangles are H3K4 acetylation.

DISCUSSION

Long non-coding RNAs are non-coding elements which are gaining attention for their relevance in gene expression regulation. However, due to the limited information on the mechanisms of action of these molecules their role in gene regulation remain an open question. Some have hypothesized that the act of transcription of lncRNA rather than their sequence is required for the 3D architecture of genome, and for the topological organization of transcriptional domains [134]. The limited number of lncRNAs that have been functionally characterized have been shown to regulate the recruitment of chromatin remodeling complexes or transcription factors affecting transcription both in a positive or negative fashion.

Here we described RAIN, a novel family of lncRNAs, and we showed that not only its expression but also its specific functions are required to sustain RUNX2 expression in cancer. Some of the previously characterized chromatin-associated lncRNAs are able to interact with histone modification complex, in particular with members of Polycomb group (PcG), PRC1 and 2 (Polycomb Repressive Complex 1 and 2) or Trithorax group (TrxG). This two groups of proteins have opposite role in gene regulation. PcG having mainly a repressors function while TrxG activates transcription. Because RAIN expression is positively associated to RUNX2 expression, we reasoned that its function in controlling RUNX2 could be mediated by the interaction with members of the TrxG group. Some studies had previously demonstrated the functional interaction of WDR5 with other lncRNAs including BLACAT 2 (bladder cancer-associated transcript 2) [135], GClnc1 (gastric cancer–associated lncRNA 1) [98], HOTTIP (HOXA transcript at the distal tip) [136], HOXD-AS1 (HOXD antisense 1) [99]. Similarly, we demonstrated that RAIN interacts with WDR5 promoting its recruitment to the RUNX2 P2 promoter. The interaction between RAIN and WDR5 is functional since silencing of RAIN resulted in a marked reduction WDR5 binding on the RUNX2 promoter. As a

consequence, the levels of H3K4Me3, H3K27Ac and RNA-PolII binding on the RUNX2 promoter was significantly decreased upon RAIN knockdown with a consequential effect on RUNX2 transcription.

Furthermore, we also demonstrated that RAIN interacts with NELFe and restrains its binding within the RUNX2 promoter. The idea that ncRNAs could also affect elongation was recently proposed based on the evidences that eRNAs promote transcription of immediate early genes in neurons without affecting chromatin structure [104]. These authors demonstrated that sequestering NELFe, eRNAs restrain the binding and the inhibitory function of the NELF complex on target genes, promoting RNA-PolII progression and elongation. Similarly, we showed that silencing of RAIN increases binding of NELFe on the P2 promoter and silencing of NELFe results in increased RUNX2 expression. We observed that the decrease of RNA-PolII phospho-S5 binding was 700bp downstream of the TSS, while the NELFe binding was enriched on TSS of RUNX2 and that complied with the stop of initiation of transcription. To ensure that the RUNX2 downregulation is due to the NELFe binding, we should conduct more experiments to confirm the increase of RNA-PolII total on TSS and the decrease of RNA-PolII phospho-S2 (marker of late stage of transcription) at 3'end of the gene.

Before our work, the ability of interfering with the NELF complex activity was shown only for eRNAs. Thus, at the best of our knowledge, this is the first demonstration of a functional interaction between the NELF complex with an ENH-associated lncRNA. Furthermore, we can also affirm that this is the first evidence that lncRNA can play multiple function in the expression regulation of target genes thanks to the interaction with different functional complexes.

Finally, we provide evidence that RAIN not only affects RUNX2 expression but acts also on other targets and through this multiple transcriptional effect promote cancer cells phenotype. By performing RNA-Seq on thyroid cancer

cells, in which the expression of RAIN was down-regulated, we showed that RAIN controls a large panel of genes of which only a fraction (about 20%) were co-shared in regulation with RUNX2. By contrast, more than 500 genes were differentially regulated selectively after l-RAIN knock-down. Further experiments are needed to validate these results and to fully define the l-RAIN mechanism of action on these targets.

In conclusion, 1) we found and characterized a novel lncRNA family, 2) we functionally proved that RAIN is able to regulate RUNX2 expression through two different mechanisms (one of this never characterized before), and 3) we provided preliminary evidences that RAIN has additional, RUNX2-independent, downstream targets in cancer.

REFERENCES

1. Kagoshima H, Shigesada K, Satake M, Ito Y, Miyoshi H, Ohki M, Pepling M, Gergen P. The Runt domain identifies a new family of heteromeric transcriptional regulators. Trends Genet. 1993 Oct;9(10):338-41.PMID: 8273148
2. Levanon D, Negreanu V, Bernstein Y, Bar-Am I, Avivi L, Groner Y. AML1, AML2, and AML3, the human members of the runt domain gene-family: cDNA structure, expression, and chromosomal localization. Genomics. 1994 Sep 15;23(2):425-32. PMID: 7835892
3. Levanon, Brenner, Negreanua, Bettouna, Woolf, Eilam, Lotem, Gat, Otto, Specke, Groner. Spatial and temporal expression pattern of Runx3 (Aml2) and Runx1 (Aml1) indicates non-redundant functions during mouse embryogenesis. Mechanisms of Development 2001
4. Bangsow C, Rubins N, Glusman G, Bernstein Y, Negreanu V, Goldenberg D, Lotem J, Ben-Asher E, Lancet D, Levanon D, Groner Y. The RUNX3 gene--sequence, structure and regulated expression. Gene. 2001 Nov 28;279(2):221-32. PMID: 11733147
5. Wu M, Li C, Zhu G, Wang Y, Jules J, Lu Y, McConnell M, Wang YJ, Shao JZ, Li YP, Chen W. Deletion of core-binding factor β (Cbfβ) in mesenchymal progenitor cells provides new insights into Cbfβ/Runxs complex function in cartilage and bone development. Bone. 2014 Aug;65:49-59. doi: 10.1016/j.bone.2014.04.031. Epub 2014 May 4. PMID: 24798493
6. Kataoka, H., Ochi, M., Enomoto, K., and Yamaguchi, A. Intrinsic transcriptional activation-inhibition domains of the polyomavirus enhancer binding protein 2/core binding factor alpha subunit revealed in the presence of the beta subunit. Mol. Cell. Biol. 18, 2444–2454. (2000).
7. Tolkunova, E. N., Fujioka, M., Kobayashi, M., et al. Two domains unique to osteoblast-specific transcription factor Osf2/Cbfa1 contribute to its transactivation function and its inability to heterodimerize with Cbfbeta. Mol. Cell. Biol. 18, 4197–4208. (1998).
8. Levanon D, Groner Y. Structure and regulated expression of mammalian RUNX genes. Oncogene. 2004 May 24;23(24):4211-9. Review. PMID: 15156175
9. Yagi R, Chen LF, Shigesada K, Murakami Y, Ito Y. A WW domain-containing yes-associated protein (YAP) is a novel transcriptional co-activator. EMBO J. 1999 May 4;18(9):2551-62. PMID: 10228168

10. McLarren KW, Lo R, Grbavec D, Thirunavukkarasu K, Karsenty G, Stifani S. The mammalian basic helix loop helix protein HES-1 binds to and modulates the transactivating function of the runt-related factor Cbfa1. J Biol Chem. 2000 Jan 7;275(1):530-8. PMID: 10617648
11. Pelletier N, Champagne N, Stifani S, Yang XJ. MOZ and MORF histone acetyltransferases interact with the Runt-domain transcription factor Runx2. Oncogene. 2002 Apr 18;21(17):2729-40. PMID: 11965546
12. Westendorf JJ, Zaidi SK, Cascino JE, Kahler R, van Wijnen AJ, Lian JB, Yoshida M, Stein GS, Li X. Runx2 (Cbfa1, AML-3) interacts with histone deacetylase 6 and represses the p21(CIP1/WAF1) promoter. Mol Cell Biol. 2002 Nov;22(22):7982-92. PMID: 12391164
13. Levanon D, Goldstein RE, Bernstein Y, Tang H, Goldenberg D, Stifani S, Paroush Z, Groner Y. Transcriptional repression by AML1 and LEF-1 is mediated by the TLE/Groucho corepressors. Proc Natl Acad Sci U S A. 1998 Sep 29;95(20):11590-5. PMID: 9751710
14. Okuda T, van Deursen J, Hiebert SW, Grosveld G, Downing JR. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. Cell. 1996 Jan 26;84(2):321-30. PMID: 8565077
15. Wang Q, Stacy T, Binder M, Marin-Padilla M, Sharpe AH, Speck NA. Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. Proc Natl Acad Sci U S A. 1996 Apr 16;93(8):3444-9. PMID: 8622955
16. North T, Gu TL, Stacy T, Wang Q, Howard L, Binder M, Marín-Padilla M, Speck NA. Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. Development. 1999 Jun;126(11):2563-75. PMID: 10226014
17. Li QL, Ito K, Sakakura C, Fukamachi H, Inoue Ki, Chi XZ, Lee KY, Nomura S, Lee CW, Han SB, Kim HM, Kim WJ, Yamamoto H, Yamashita N, Yano T, Ikeda T, Itohara S, Inazawa J, Abe T, Hagiwara A, Yamagishi H, Ooe A, Kaneda A, Sugimura T, Ushijima T, Bae SC, Ito Y. Causal relationship between the loss of RUNX3 expression and gastric cancer. Cell. 2002 Apr 5;109(1):113-24. PMID: 11955451
18. Inoue K, Ozaki S, Shiga T, Ito K, Masuda T, Okado N, Iseda T, Kawaguchi S, Ogawa M, Bae SC, Yamashita N, Itohara S, Kudo N, Ito Y. Runx3 controls the axonal

projection of proprioceptive dorsal root ganglion neurons. Nat Neurosci. 2002 Oct;5(10):946-54. PMID: 12352981

19. Woolf E, Xiao C, Fainaru O, Lotem J, Rosen D, Negreanu V, Bernstein Y, Goldenberg D, Brenner O, Berke G, Levanon D, Groner Y. Runx3 and Runx1 are required for CD8 T cell development during thymopoiesis. Proc Natl Acad Sci U S A. 2003 Jun 24;100(13):7731-6. Epub 2003 Jun 9.
20. Komori T, Yagi H, Nomura S, Yamaguchi A, Sasaki K, Deguchi K, Shimizu Y, Bronson RT, Gao YH, Inada M, Sato M, Okamoto R, Kitamura Y, Yoshiki S, Kishimoto T. Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. Cell. 1997 May 30;89(5):755-64. PMID: 9182763
21. Endo T, Kobayashi T. Runx2 deficiency in mice causes decreased thyroglobulin expression and hypothyroidism. Mol Endocrinol. 2010 Jun;24(6):1267-73. doi: 10.1210/me.2010-0056. Epub 2010 Apr 7.
22. Ito Y. RUNX genes in development and cancer: regulation of viral gene expression and the discovery of RUNX family genes. Adv Cancer Res (2008) 99:33—76 3
23. Ito Y, Bae SC, Chuang LS. The RUNX family: developmental regulators in cancer. Nat Rev Cancer. 2015 Feb;15(2):81-95. doi: 10.1038/nrc3877. Epub 2015 Jan 16. Review. PMID: 25592647
24. Margaretha Van Der Deen, Jacqueline Akech, David Lapointe, Sneha Gupta, Daniel W. Young, Martin A. Montecino, Mario Galindo, Jane B. Lian, Janet L. Stein, Gary S. Genomic promoter occupancy of Runt-related transcription factor RUNX2 in osteosarcoma cells identifies genes involved in cell adhesion and motility. J Biol Chem. 2012 Feb 10;287(7):4503-17. doi: 10.1074/jbc.M111.287771. Epub 2011 Dec 9
25. Chimge NO, Frenkel B. The RUNX family in breast cancer: relationships with estrogen signaling. Oncogene. 2013 Apr 25;32(17):2121-30. doi: 10.1038/onc.2012.328. Epub 2012 Oct 8. PMID: 23045283
26. .Ya-lin Li, Zhou-sheng Xiao. Advances in Runx2 regulation and its isoforms. Medical Hypotheses (2007) 68, 169–175
27. Bae SC, Lee YH. Phosphorylation, acetylation and ubiquitination: the molecular basis of RUNX regulation. Gene. 2006 Jan 17;366(1):58-66. Epub 2005 Dec 1.

28. Schroeder TM, Jensen ED, Westendorf JJ. Runx2: a master organizer of gene transcription in developing and maturing osteoblasts. Birth Defects Res C Embryo Today. 2005 Sep;75(3):213-25.
29. Stock M, Otto F. Control of RUNX2 isoform expression: the role of promoters and enhancers. J Cell Biochem. 2005 Jun 1;95(3):506-17.
30. Otto F, Lübbert M, Stock M. Upstream and downstream targets of RUNX proteins. J Cell Biochem. 2003 May 1;89(1):9-18. Review. PMID: 12682904
31. Drissi H, Luc Q, Shakoori R, Chuva De Sousa Lopes S, Choi JY, Terry A, Hu M, Jones S, Neil JC, Lian JB, Stein JL, Van Wijnen AJ, Stein GS. Transcriptional autoregulation of the bone related CBFA1/RUNX2 gene. J Cell Physiol. 2000 Sep;184(3):341-50.
32. Drissi H, Pouliot A, Stein JL, van Wijnen AJ, Stein GS, Lian JB. Identification of novel protein/DNA interactions within the promoter of the bone-related transcription factor Runx2/Cbfa1. J Cell Biochem. 2002;86(2):403-12.
33. Hill TP, Später D, Taketo MM, Birchmeier W, Hartmann C Canonical Wnt/beta-catenin signaling prevents osteoblasts from differentiating into chondrocytes. Dev Cell. 2005 May;8(5):626-7
34. Day TF, Guo X, Garrett-Beal L, Yang Y. Wnt/beta-catenin signaling in mesenchymal progenitors controls osteoblast and chondrocyte differentiation during vertebrate skeletogenesis. Dev Cell. 2005 May;8(5):739-50.
35. .Canonical WNT signaling promotes osteogenesis by directly stimulating Runx2 gene expression. Gaur T, Lengner CJ, Hovhannisyan H, Bhat RA, Bodine PV, Komm BS, Javed A, van Wijnen AJ, Stein JL, Stein GS, Lian JB. J Biol Chem. 2005 Sep 30;280(39):33132-40. Epub 2005 Jul 25.
36. Gori F, Thomas T, Hicok KC, Spelsberg TC, Riggs BL. Differentiation of human marrow stromal precursor cells: bone morphogenetic protein-2 increases OSF2/CBFA1, enhances osteoblast commitment, and inhibits late adipocyte maturation. J Bone Miner Res. 1999 Sep;14(9):1522-35.
37. Lee KS1, Kim HJ, Li QL, Chi XZ, Ueta C, Komori T, Wozney JM, Kim EG, Choi JY, Ryoo HM, Bae SC. Runx2 is a common target of transforming growth factor beta1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific gene expression in the pluripotent mesenchymal precursor cell line C2C12. Mol Cell Biol. 2000 Dec;20(23):8783-92.

38. Kaneki H, Guo R, Chen D, Yao Z, Schwarz EM, Zhang YE, Boyce BF, Xing L. Tumor necrosis factor promotes Runx2 degradation through up-regulation of Smurf1 and Smurf2 in osteoblasts. J Biol Chem. 2006 Feb 17;281(7):4326-33. Epub 2005 Dec 22.
39. Guo R, Yamashita M, Zhang Q, Zhou Q, Chen D, et al. Ubiquitin ligase Smurf1 mediates tumor necrosis factor-induced systemic bone loss by promoting proteasomal degradation of bone morphogenetic signaling proteins. J Biol Chem. 2008 Aug 22;283(34):23084-92. doi: 10.1074/jbc.M709848200. Epub 2008 Jun 19.ù
40. Selvamurugan N, Shimizu E, Lee M, Liu T, Li H, Partridge NC. Identification and characterization of Runx2 phosphorylation sites involved in matrix metalloproteinase-13 promoter activation. FEBS Lett. 2009 Apr 2;583(7):1141-6. doi: 10.1016/j.febslet.2009.02.040. Epub 2009 Mar 3.
41. Wee HJ, Huang G, Shigesada K, Ito Y. Serine phosphorylation of RUNX2 with novel potential functions as negative regulatory mechanisms. EMBO Rep. 2002 Oct;3(10):967-74. Epub 2002 Sep 13.
42. Phillips JE, Gersbach CA, Wojtowicz AM. García AJ Glucocorticoid-induced osteogenesis is negatively regulated by Runx2/Cbfa1 serine phosphorylation. J Cell Sci. 2006 Feb 1;119(Pt 3):581-91
43. Zhu W, He X, Hua Y, Li Q, Wang J, Gan X. The E3 ubiquitin ligase WWP2 facilitates RUNX2 protein transactivation in a mono-ubiquitination manner during osteogenic differentiation. J. Biol Chem. 2017 Jul 7;292(27):11178-11188. doi: 10.1074/jbc.M116.772277. Epub 2017 May 12
44. Shen R, Chen M, Wang YJ, Kaneki H, Xing L, O'keefe RJ, Chen D Smad6 interacts with Runx2 and mediates Smad ubiquitin regulatory factor 1-induced Runx2 degradation. J. Biol Chem. 2006 Feb 10;281(6):3569-76. Epub 2005 Nov 18
45. Huanga, Jianga, Meng, Jing, Wei, et al. MiR-30a inhibits osteolysis by targeting RunX2 in giant cell tumor of bone. Biochemical and Biophysical Research Communications Volume 453, Issue 1, 10 October 2014, Pages 160-165
46. Zuo B, Zhu J, Li J, Wang C, Zhao X, Cai G, Li Z, Peng J, Wang P, Shen C, Huang Y, Xu J, Zhang X, Chen X microRNA-103a functions as a mechanosensitive microRNA to inhibit bone formation through targeting Runx2. J Bone Miner Res. 2015 Feb;30(2):330-45. doi: 10.1002/jbmr.2352.

47. Huang J, Zhao L, Xing L, Chen D. MicroRNA-204 regulates Runx2 protein expression and mesenchymal progenitor cell differentiation. Stem Cells. 2010 Feb;28(2):357-64. doi: 10.1002/stem.288.
48. Sancisi V, Borettini G, Maramotti S, Ragazzi M, Tamagnini I, Nicoli D, Piana S, Ciarrocchi A. Runx2 isoform I controls a panel of proinvasive genes driving aggressiveness of papillary thyroid carcinomas. J Clin Endocrinol Metab. 2012 Oct;97(10):E2006-15. doi: 10.1210/jc.2012-1903. Epub 2012 Jul 20.
49. Lau CC, Harris CP, Lu XY, Perlaky L, Gogineni S, Chintagumpala M, Hicks J, Johnson ME, Davino NA, Huvos AG, Meyers PA, Healy JH, Gorlick R, Rao PH Frequent amplification and rearrangement of chromosomal bands 6p12-p21 and 17p11.2 in osteosarcoma. Genes Chromosomes Cancer. 2004 Jan;39(1):11-21.
50. Martin JW, Zielenska M, Stein GS, van Wijnen AJ, Squire JA. The Role of RUNX2 in Osteosarcoma Oncogenesis. Sarcoma. 2011;2011:282745. doi: 10.1155/2011/282745. Epub 2010 Dec 9. PMID: 21197465
51. Akech J, Wixted JJ, Bedard K, van der Deen M, Hussain S, Guise TA, van Wijnen AJ, Stein et al. Runx2 association with progression of prostate cancer in patients: mechanisms mediating bone osteolysis and osteoblastic metastatic lesions. Oncogene. 2010 Feb 11;29(6):811-21. doi: 10.1038/onc.2009.389. Epub 2009 Nov 16.
52. Boregowda RK, Olabisi OO, Abushahba W, Jeong BS, Haenssen KK, Chen W, Chekmareva M, Lasfar A, Foran, et al. RUNX2 is overexpressed in melanoma cells and mediates their migration and invasion. Cancer Lett. 2014 Jun 28;348(1-2):61-70. doi: 10.1016/j.canlet.2014.03.011. Epub 2014 Mar 18.
53. Wang ZQ, Keita M, Bachvarova M, Gobeil S, Morin C, Plante M, Gregoire J, Renaud MC, et al. Inhibition of RUNX2 transcriptional activity blocks the proliferation, migration and invasion of epithelial ovarian carcinoma cells. PLoS One. 2013 Oct 4;8(10):e74384. doi: 10.1371/journal.pone.0074384. eCollection 2013.
54. Ikushima H, Miyazono K. TGFbeta signalling: a complex web in cancer progression. Nat Rev Cancer. 2010 Jun;10(6):415-24. doi: 10.1038/nrc2853. Review. PMID: 20495575
55. Sancisi V, Gandolfi G, Ragazzi M, Nicoli D, Tamagnini I, Piana S, Ciarrocchi A. Cadherin 6 is a new RUNX2 target in TGF-β signalling pathway. PLoS One. 2013 Sep 12;8(9):e75489. doi: 10.1371/journal.pone.0075489. eCollection 2013. PMID: 24069422

56. Goloudina AR, Tanoue K, Hammann A, Fourmaux E, Le Guezennec X, Bulavin DV et al. Regulation of breast cancer metastasis by Runx2 and estrogen signaling: the role of SNAI2. Breast Cancer Res 13(6):R127
57. Khalid O, Baniwal SK, Purcell DJ, Leclerc N, Gabet Y, Stallcup MR, Coetzee GA, Frenkel B. Modulation of Runx2 activity by estrogen receptor-alpha: implications for osteoporosis and breast cancer. Endocrinology. 2008 Dec;149(12):5984-95. doi: 10.1210/en.2008-0680. Epub 2008 Aug 28.
58. Rucci N, Teti A. Osteomimicry: How the Seed Grows in the Soil. Calcif Tissue Int. 2018 Feb;102(2):131-140. doi: 10.1007/s00223-017-0365-1. Epub 2017 Nov 16.
59. Tan CC, Li GX, Tan LD, Du X, Li XQ, He R, Wang QS, Feng et al. Breast cancer cells obtain an osteomimetic feature via epithelial-mesenchymal transition that have undergone BMP2/RUNX2 signaling pathway induction. Oncotarget. 2016 Nov 29;7(48):79688-79705. doi: 10.18632/oncotarget.12939
60. Bellahcène A, Bachelier R, Detry C, Lidereau R, Clézardin P, Castronovo V Transcriptome analysis reveals an osteoblast-like phenotype for human osteotropic breast cancer cells. Breast Cancer Res Treat. 2007 Jan;101(2):135-48. Epub 2006 Oct 7
61. Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, Cordón-Cardo C, Guise TA, Massagué J. A multigenic program mediating breast cancer metastasis to bone. Cancer Cell. 2003 Jun;3(6):537-49.
62. van der Deen M, Akech J, Lapointe D, Gupta S, Young DW, Montecino MA, et al.Genomic promoter occupancy of runt-related transcription factor RUNX2 in Osteosarcoma cells identifies genes involved in cell adhesion and motility. J Biol Chem. 2012 Feb 10;287(7):4503-17. doi: 10.1074/jbc.M111.287771. Epub 2011 Dec 9.
63. Inman and Shore. The Osteoblast Transcription Factor Runx2 Is Expressed in Mammary Epithelial Cells and Mediates osteopontin Expression. J Biol Chem. 2003 Dec 5;278(49):48684-9. Epub 2003 Sep 23
64. Zelzer E1, Glotzer DJ, Hartmann C, Thomas D, Fukai N, Soker S, Olsen BR. Tissue specific regulation of VEGF expression during bone development requires Cbfa1/Runx2. Mech Dev. 2001 Aug;106(1-2):97-106. PMID: 11472838

65. Pratap J, Javed A, Languino LR, van Wijnen AJ, Stein JL, Stein GS, Lian JB. The Runx2 osteogenic transcription factor regulates matrix metalloproteinase 9 in bone metastatic cancer cells and controls cell invasion. Cancer Res 63:2631–2637
66. Niu DF, Kondo T, Nakazawa T, Oishi N, Kawasaki T, Mochizuki K, Yamane T, Katoh R. Transcription factor Runx2 is a regulator of epithelial-mesenchymal transition and invasion in thyroid carcinomas. Lab Invest. 2012 Aug;92(8):1181-90. doi: 10.1038/labinvest.2012.84. Epub 2012 May 28.
67. Ciarrocchi A, Piana S, Valcavi R, Gardini G, Casali B. Inhibitor of DNA binding-1 induces mesenchymal features and promotes invasiveness in thyroid tumour cells. Eur J Cancer. 2011 Apr;47(6):934-45. doi: 10.1016/j.ejca.2010.11.009. Epub 2010 Dec 9. PMID: 21146400
68. Vousden KH, Lu X. Live or let die: the cell's response to p53. Nat Rev Cancer. 2002 Aug;2(8):594-604. PMID: 12154352
69. Sionov RV, Haupt Y. The cellular response to p53: the decision between life and death. Oncogene. 1999 Nov 1;18(45):6145-57.
70. Browne G, Nesbitt H, Ming L, Stein GS, Lian JB, McKeown SR, Worthington J. Bicalutamide-induced hypoxia potentiates RUNX2-mediated Bcl-2 expression resulting in apoptosis resistance. Br J Cancer. 2012 Nov 6;107(10):1714-21. doi: 10.1038/bjc.2012.455. Epub 2012 Oct 16. PMID: 23073173
71. Blyth K, Vaillant F, Hanlon L, Mackay N, Bell M, Jenkins A, Neil JC, Cameron ER. Runx2 and MYC collaborate in lymphoma development by suppressing apoptotic and growth arrest pathways in vivo. Cancer Res. 2006 Feb 15;66(4):2195-201.
72. Ozaki T, Wu D, Sugimoto H, Nagase H, Nakagawara A. Runt-related transcription factor 2 (RUNX2) inhibits p53-dependent apoptosis through the collaboration with HDAC6 in response to DNA damage. Cell Death Dis. 2013 Apr 25;4:e610. doi: 10.1038/cddis.2013.127. PMID: 23618908
73. Sancisi V, Gandolfi G, Ambrosetti DC, Ciarrocchi A. Histone Deacetylase Inhibitors Repress Tumoral Expression of the Proinvasive Factor RUNX2. Cancer Res. 2015 May 1;75(9):1868-82. doi: 10.1158/0008-5472.CAN-14-2087. Epub 2015 Mar 13. PMID: 25769725
74. Sancisi V, Manzotti G, Gugnoni M, Rossi T, Gandolfi G, Gobbi G, Torricelli F, Catellani F, Faria do Valle I, Remondini D, Castellani G, Ragazzi M, Piana S, Ciarrocchi A. RUNX2 expression in thyroid and breast cancer requires the cooperation

of three non-redundant enhancers under the control of BRD4 and c-JUN. Nucleic Acids Res. 2017 Nov 2;45(19):11249-11267. doi: 10.1093/nar/gkx802. PMID: 28981843

75. ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 2007 Jun 14;447(7146):799-816. PMID: 17571346
76. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet. 2009 Mar;10(3):155-9. doi: 10.1038/nrg2521. PMID: 19188922
77. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 2012;81:145-66. doi: 10.1146/annurev-biochem-051410-092902. PMID: 22663078
78. Mercer TR, Mattick JS. Structure and function of long noncoding RNAs in epigenetic regulation. Nat Struct Mol Biol. 2013 Mar;20(3):300-7. doi: 10.1038/nsmb.2480. PMID: 23463315
79. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al.. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 2012 Sep;22(9):1775-89. doi: 10.1101/gr.132159.111. PMID: 22955988
80. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 2011 Sep 15;25(18):1915-27. doi: 10.1101/gad.17446611. Epub 2011 Sep 2. PMID: 21890647
81. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. Cell. 2009 Feb 20;136(4):629-41. doi: 10.1016/j.cell.2009.02.006. PMID: 19239885
82. Chen LL. Linking Long Noncoding RNA Localization and Function. Trends Biochem Sci. 2016 Sep;41(9):761-72. doi: 10.1016/j.tibs.2016.07.003. Epub 2016 Aug 4. Review. PMID: 27499234
83. Flynn RA, Chang HY. Long noncoding RNAs in cell-fate programming and reprogramming. Cell Stem Cell. 2014 Jun 5;14(6):752-61. doi: 10.1016/j.stem.2014.05.014. PMID: 24905165
84. Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. Cell. 2013 Mar 14;152(6):1298-307. doi: 10.1016/j.cell.2013.02.012. PMID: 23498938

85. Marvin MC, Clauder-Münster S, Walker SC, Sarkeshik A, Yates JR 3rd, Steinmetz LM, Engelke DR. Accumulation of noncoding RNA due to an RNase P defect in Saccharomyces cerevisiae. RNA. 2011 Aug;17(8):1441-50. doi: 10.1261/rna.2737511. Epub 2011 Jun 10. PMID: 21665995
86. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, Loewer A, Ziebold U, Landthaler M, Kocks C, le Noble F, Rajewsky N. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature. 2013 Mar 21;495(7441):333-8. doi: 10.1038/nature11928. Epub 2013 Feb 27. PMID: 23446348
87. RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection. Nature. 2001 Feb 8;409(6821):685-90. PMID: 11217851
88. . FANTOM Consortium; RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). The transcriptional landscape of the mammalian genome. Science. 2005 Sep 2;309(5740):1559-63. Erratum in: Science. 2006 Mar 24;311(5768):1713. PMID: 16141072
89. RIKEN Genome Exploration Research Group; Genome Science Group (Genome Network Project Core Group); FANTOM Consortium. Antisense transcription in the mammalian transcriptome. Science. 2005 Sep 2;309(5740):1564-6. PMID: 16141073
90. Hon CC, Ramilowski JA, Harshbarger J, Bertin N, Rackham OJ, Gough J et al. An atlas of human long non-coding RNAs with accurate 5' ends. Nature. 2017 Mar 9;543(7644):199-204. doi: 10.1038/nature21374. Epub 2017 Mar 1. PMID: 28241135
91. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 2012 Sep;22(9):1760-74. doi: 10.1101/gr.135350.111. PMID: 22955987
92. St Laurent G, Wahlestedt C, Kapranov P. The Landscape of long noncoding RNA classification. Trends Genet. 2015 May;31(5):239-51. doi: 10.1016/j.tig.2015.03.007. Epub 2015 Apr 10. Review. PMID: 25869999
93. Maston GA, Evans SK, Green MR Transcriptional regulatory elements in the human genome. Annu Rev Genomics Hum Genet. 2006;7:29-59. PMID: 16719718
94. Geyer PK, Green MM, Corces VG. Tissue-specific transcriptional enhancers may act in trans on the gene located in the homologous chromosome: the molecular basis of transvection in Drosophila. EMBO J. 1990 Jul;9(7):2247-56. PMID: 2162766

95. Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R. Interchromosomal interactions and olfactory receptor choice. Cell. 2006 Jul 28;126(2):403-13. PMID: 16873069
96. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, et al. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010 May 13;465(7295):182-7. doi: 10.1038/nature09033. Epub 2010 Apr 14. PMID: 20393465
97. Ørom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, et al. Long noncoding RNAs with enhancer-like function in human cells. Cell. 2010 Oct 1;143(1):46-58. doi: 10.1016/j.cell.2010.09.001. PMID: 20887892
98. Sun TT, He J, Liang Q, Ren LL, Yan TT, Yu TC, Tang JY, et al. LncRNA GClnc1 Promotes Gastric Carcinogenesis and May Act as a Modular Scaffold of WDR5 and KAT2A Complexes to Specify the Histone Modification Pattern. Cancer Discov. 2016 Jul;6(7):784-801. doi: 10.1158/2159-8290.CD-15-0921. Epub 2016 May 4. PMID: 27147598
99. Gu P, Chen X, Xie R, Han J, Xie W, Wang B, Dong W, et al. lncRNA HOXD-AS1 Regulates Proliferation and Chemo-Resistance of Castration-Resistant Prostate Cancer via Recruiting WDR5. Mol Ther. 2017 Aug 2;25(8):1959-1973. doi: 10.1016/j.ymthe.2017.04.016. Epub 2017 May 6. PMID: 28487115
100. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhattar R. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature. 2013 Feb 28;494(7438):497-501. doi: 10.1038/nature11884. Epub 2013 Feb 17. PMID: 23417068
101. Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature. 2013 Jun 27;498(7455):516-20. doi: 10.1038/nature12210. Epub 2013 Jun 2. PMID: 23728302
102. Sigova AA, Abraham BJ, Ji X, Molinie B, Hannett NM, Guo YE, et al. Transcription factor trapping by RNA in gene regulatory elements. Science. 2015 Nov 20;350(6263):978-81. doi: 10.1126/science.aad3346. Epub 2015 Oct 29. PMID: 26516199
103. Xiang JF, Yin QF, Chen T, Zhang Y, Zhang XO, Wu Z, Zhang S, et al. Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. Cell Res. 2014 May;24(5):513-31. doi: 10.1038/cr.2014.35. Epub 2014 Mar 25. PMID: 24662484

104. Schaukowitch K, Joo JY, Liu X, Watts JK, Martinez C, Kim TK. Enhancer RNA facilitates NELF release from immediate early genes. Mol Cell. 2014 Oct 2;56(1):29-42. doi: 10.1016/j.molcel.2014.08.023. Epub 2014 Sep 25. PMID: 25263592

105. van Heesch S, van Iterson M, Jacobi J, Boymans S, Essers PB, de Bruijn E, et al. Extensive localization of long noncoding RNAs to the cytosol and mono- and polyribosomal complexes. Genome Biol. 2014 Jan 7;15(1):R6. doi: 10.1186/gb-2014-15-1-r6. PMID: 24393600

106. Cabili MN, Dunagin MC, McClanahan PD, Biaesch A, Padovan-Merhar O, Regev A, et al. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. Genome Biol. 2015 Jan 29;16:20. doi: 10.1186/s13059-015-0586-4. PMID: 25630241

107. Affymetrix ENCODE Transcriptome Project; Cold Spring Harbor Laboratory ENCODE Transcriptome Project. Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. Nature. 2009 Feb 19;457(7232):1028-32. doi: 10.1038/nature07759. Epub 2009 Jan 25. PMID: 19169241

108. Wilusz JE, Freier SM, Spector DL. 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. Cell. 2008 Nov 28;135(5):919-32. doi: 10.1016/j.cell.2008.10.012. PMID: 19041754

109. Zhang B, Gunawardane L, Niazi F, Jahanbani F, Chen X, Valadkhan S A novel RNA motif mediates the strict nuclear localization of a long noncoding RNA. Mol Cell Biol. 2014 Jun;34(12):2318-29. doi: 10.1128/MCB.01673-13. Epub 2014 Apr 14. PMID: 24732794

110. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A. 2009 Jul 14;106(28):11667-72. doi: 10.1073/pnas.0904715106. Epub 2009 Jul 1. PMID: 19571010

111. Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, et al. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol Cell. 2008 Oct 24;32(2):232-46. doi: 10.1016/j.molcel.2008.08.022. PMID: 18951091

112. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. Long noncoding RNA as modular scaffold of histone modification complexes. Science. 2010

Aug 6;329(5992):689-93. doi: 10.1126/science.1192002. Epub 2010 Jul 8. PMID: 20616235

113. Bernard D, Prasanth KV, Tripathi V, Colasse S, Nakamura T, Xuan Z, et al. A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. EMBO J. 2010 Sep 15;29(18):3082-93. doi: 10.1038/emboj.2010.199. Epub 2010 Aug 20. PMID: 20729808
114. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell. 2010 Sep 24;39(6):925-38. doi: 10.1016/j.molcel.2010.08.011. PMID: 20797886
115. Chen LL, Carmichael GG. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. Mol Cell. 2009 Aug 28;35(4):467-78. doi: 10.1016/j.molcel.2009.06.027. PMID: 19716791
116. Bartonicek N, Maag JL, Dinger ME. Long noncoding RNAs in cancer: mechanisms of action and technological advancements. Mol Cancer. 2016 May 27;15(1):43. doi: 10.1186/s12943-016-0530-6. PMID: 27233618
117. Ifere GO1, Ananaba GA. Prostate cancer gene expression marker 1 (PCGEM1): a patented prostate- specific non-coding gene and regulator of prostate cancer progression. Recent Pat DNA Gene Seq. 2009;3(3):151-63. PMID: 19891595
118. de Kok JB, Verhaegh GW, Roelofs RW, Hessels D, Kiemeney LA, Aalders TW, et al. DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. Cancer Res. 2002 May 1;62(9):2695-8. PMID: 11980670
119. Chung S, Nakagawa H, Uemura M, Piao L, Ashikawa K, Hosono N, et al. Association of a novel long non-coding RNA in 8q24 with prostate cancer susceptibility. Cancer Sci. 2011 Jan;102(1):245-52. doi: 10.1111/j.1349-7006.2010.01737.x. Epub 2010 Sep 28. PMID: 20874843
120. Ji P1, Diederichs S, Wang W, Böing S, Metzger R, Schneider PM, et al. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene. 2003 Sep 11;22(39):8031-41. PMID: 12970751
121. Tano K1, Mizuno R, Okada T, Rakwal R, Shibato J, Masuo Y, Ijiri K, Akimitsu N. MALAT-1 enhances cell motility of lung adenocarcinoma cells by influencing the

expression of motility-related genes. FEBS Lett. 2010 Nov 19;584(22):4575-80. doi: 10.1016/j.febslet.2010.10.008. Epub 2010 Oct 13. PMID: 20937273

122. Zhang ZC, Tang C, Dong Y, Zhang J, Yuan T, Li XL Targeting LncRNA-MALAT1 suppresses the progression of osteosarcoma by altering the expression and localization of β-catenin. J Cancer. 2018 Jan 1;9(1):71-80. doi: 10.7150/jca.22113. eCollection 2018. PMID: 29290771

123. Jie Y, Zhao H2 LncRNA MALAT1 induces colon cancer development by regulating miR-129-5p/HMGB1 axis J Cell Physiol. 2017 Dec 11. doi: 10.1002/jcp.26383. [Epub ahead of print] PMID: 29226325

124. Yang Z1, Zhou L, Wu LM, Lai MC, Xie HY, Zhang F, Zheng SS Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. Ann Surg Oncol. 2011 May;18(5):1243-50. doi: 10.1245/s10434-011-1581-y. Epub 2011 Feb 15. PMID: 21327457

125. Niinuma T1, Suzuki H, Nojima M, Nosho K, Yamamoto H, Takamaru et al Upregulation of miR-196a and HOTAIR drive malignant character in gastrointestinal stromal tumors Cancer Res. 2012 Mar 1;72(5):1126-36. doi: 10.1158/0008-5472.CAN-11-1803. Epub 2012 Jan 18. PMID: 22258453

126. Liu M1, Jia J1, Wang X1, Liu Y2, Wang C2, Fan R3. Long non-coding RNA HOTAIR promotes cervical cancer progression through regulating BCL2 via targeting miR-143-3p. Cancer Biol Ther. 2018 Jan 16:1-9. doi: 10.1080/15384047.2018.1423921. [Epub ahead of print] PMID: 29336659

127. Gupta RA1, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ et al Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis Nature. 2010 Apr 15;464(7291):1071-6. doi: 10.1038/nature08975. PMID: 20393566

128. Panzitt K, Tschernatsch MM, Guelly C, Moustafa T, Stradner M, Strohmaier HM, et al. Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. Gastroenterology. 2007 Jan;132(1):330-42. Epub 2006 Aug 14. PMID: 17241883

129. Matouk IJ, Abbasi I, Hochberg A, Galun E, Dweik H, Akkawi M. Highly upregulated in liver cancer noncoding RNA is overexpressed in hepatic colorectal metastasis. Eur J Gastroenterol Hepatol. 2009 Jun;21(6):688-92. PMID: 19445043

130. Gutschner T, Diederichs S. The hallmarks of cancer: a long non-coding RNA point of view. RNA Biol. 2012 Jun;9(6):703-19. doi: 10.4161/rna.20481. Epub 2012 Jun 1. Review. PMID: 22664915
131. Kanno T, Kanno Y, LeRoy G, Campos E, Sun HW, Brooks SR, et al. BRD4 assists elongation of both coding and enhancer RNAs by interacting with acetylated histones. Nat Struct Mol Biol. 2014 Dec;21(12):1047-57. doi: 10.1038/nsmb.2912. Epub 2014 Nov 10. PMID: 25383670
132. Chu C, Quinn J, Chang HY. Chromatin isolation by RNA purification (ChIRP). J Vis Exp. 2012 Mar 25;(61). pii: 3912. doi: 10.3791/3912. PMID: 22472705
133. G Hendrickson D, Kelley DR, Tenen D, Bernstein B, Rinn JL. Widespread RNA binding by chromatin-associated proteins. Genome Biol. 2016 Feb 16;17:28. doi: 10.1186/s13059-016-0878-3. PMID: 26883116
134. Melé M, Rinn JL. "Cat's Cradling" the 3D Genome by the Act of LncRNA Transcription. Mol Cell. 2016 Jun 2;62(5):657-64. doi: 10.1016/j.molcel.2016.05.011. PMID: 27259198
135. He W, Zhong G, Jiang N, Wang B, Fan X, Chen C, et al. Long noncoding RNA BLACAT2 promotes bladder cancer-associated lymphangiogenesis and lymphatic metastasis. J Clin Invest. 2018 Feb 1;128(2):861-875. doi: 10.1172/JCI96218. Epub 2018 Jan 22.
136. Fu Z, Chen C, Zhou Q, Wang Y, Zhao Y, Zhao X, et al. LncRNA HOTTIP modulates cancer stem cell properties in human pancreatic cancer by regulating HOXA9. Cancer Lett. 2017 Dec 1;410:68-81. doi: 10.1016/j.canlet.2017.09.019. Epub 2017 Sep 22. PMID: 28947139
137. Hiller M, Pudimat R, Busch A, Backofen R. Using RNA secondary structures to guide sequence motif finding towards single-stranded regions. Nucleic Acids Res. 2006;34(17):e117. Epub 2006 Sep 20.
138. Maticzka D, Lange SJ, Costa F, Backofen R. GraphProt: modeling binding preferences of RNA-binding proteins. Genome Biol. 2014 Jan 22;15(1):R17. doi: 10.1186/gb-2014-15-1-r17.
139. Li JH, Liu S, Zheng LL, Wu J, Sun WJ, Wang ZL et al. Discovery of Protein-lncRNA Interactions by Integrating Large-Scale CLIP-Seq and RNA-Seq Datasets. Front Bioeng Biotechnol. 2015 Jan 14;2:88. doi: 10.3389/fbioe.2014.00088. eCollection 2014.

Publications during PhD

1. *RUNX2 expression in thyroid and breast cancer requires the cooperation of three non-redundant enhancers under the control of BRD4 and c-JUN.* Sancisi V, Manzotti G, Gugnoni M, **Rossi T**, Gandolfi G, Gobbi G, Torricelli F, Catellani F, Faria do Valle I, Remondini D, Castellani G, Ragazzi M, Piana S, Ciarrocchi A. Nucleic Acids Res. 2017 Nov 2;45(19):11249-11267. doi: 10.1093/nar/gkx802.
2. *Efficacy of NEDD8 Pathway Inhibition in Preclinical Models of Poorly Differentiated, Clinically Aggressive Colorectal Cancer*. Picco G, Petti C, Sassi F, Grillone K, Migliardi G, **Rossi T**, Isella C, Di Nicolantonio F, Sarotto I, Sapino A, Bardelli A, Trusolino L, Bertotti A, Medico E. J Natl Cancer Inst. 2016 Oct 22;109(2). pii: djw209. Print 2017 Feb.
3. *Nutrigenomics of extra-virgin olive oil: A review*. Piroddi M, Albini A, Fabiani R, Giovannelli L, Luceri C, Natella F, Rosignoli P, **Rossi T**, Taticchi A, Servili M, Galli F. Biofactors. 2017 Jan 2;43(1):17-41. doi: 10.1002/biof.1318. Epub 2016 Sep 1. Review.
4. *Potential chemopreventive activities of a polyphenol rich purified extract from olive mill wastewater on colon cancer cells.* Bassani B*, **Rossi T***, De Stefano D, Pizzichini D, Corradino P, Macrì N, Noonan DM, Albini A, Bruno, A... J Funct Foods. 2016;27: 236-248
5. *Hop derived flavonoid xanthohumol inhibits endothelial cell functions via AMPK activation*. Gallo C, Dallaglio K, Bassani B, **Rossi T**, Rossello A, Noonan DM, D'Uva G, Bruno A, Albini A. Oncotarget. 2016 Sep 13;7(37):59917-59931. doi: 10.18632/oncotarget.10990.
6. *Aspirin and atenolol enhance metformin activity against breast cancer by targeting both neoplastic and microenvironment cells*. Talarico G, Orecchioni S, Dallaglio K, Reggiani F, Mancuso P, Calleri A, Gregato

G, Labanca V, **Rossi T**, Noonan DM, Albini A, Bertolini F. Sci Rep. 2016 Jan 5;6:18673. doi: 10.1038/srep18673.

7. *Effect of a Purified Extract of Olive Mill Waste water on Endothelial Cell Proliferation, Apoptosis, Migration and Capillary-Like Structure in vitro and in vivo.* **Rossi T**, Bassani B, Gallo C, Maramotti S, Noonan DM, Albini A, Bruno A. J Bioanal Biomed. 2015:S12: 006.
8. *Preliminary Evidence on the Diagnostic and Molecular Role of Circulating Soluble EGFR in Non-Small Cell Lung Cancer.* Lococo F, Paci M, Rapicetta C, **Rossi T**, Sancisi V, Braglia L, Cavuto S, Bisagni A, Bongarzone I, Noonan DM, Albini A, Maramotti S. Int J Mol Sci. 2015 Aug 19;16(8):19612-30. doi: 10.3390/ijms160819612.