UNIVERSITÀ DEGLI STUDI DI CATANIA

DOTTORATO DI RICERCA IN BIOMEDICINA TRASLAZIONALE

CURRICULUM IN BIOMEDICINA MOLECOLARE GENOMICA E DEI SISTEMI COMPLESSI: BASI CELLULARI E MOLECOLARI DEL FENOTIPO - XIX CICLO

DIPARTIMENTO DI SCIENZE BIOMEDICHE E BIOTECNOLOGICHE

SEZIONE DI BIOLOGIA E GENETICA "GIOVANNI SICHEL"

Cristina Barbagallo

_____

LncRNA and circRNA expression profiles in tissues and serum exosomes of colorectal cancer patients and cell lines

_____

TESI DI DOTTORATO

_____

Coordinatore del Dottorato: Prof. Lorenzo Malatino

Coordinatore del Curriculum: Prof. Michele Purrello

Tutor: Prof. Marco Ragusa

Anno Accademico 2015-2016

Table of contents

1. Abstract

In the last few years several studies demonstrated the fundamental role of non-coding RNAs (ncRNAs) in tumor onset and progression. While the involvement and the mechanism of action of microRNAs (miRNAs) have been widely investigated, little is known about long non-coding RNAs (lncRNAs) and circular RNAs (circRNAs), which seem to act through a plethora of molecular mechanisms regulating essential biological processes, such as cell cycle, splicing, chromatin remodeling, apoptosis, adhesion and migration. LncRNAs are non-coding RNAs longer than 200 nucleotides, with or without 5'-cap and poly(A) tail; circRNAs are circular molecules lacking free ends and thus resistant to exonucleasic degradation, transcribed from protein-coding genes. Several evidences showed that aberrant expression of lncRNAs and circRNAs is associated with various tumors, including CRC. In our study we analysed through Real-Time PCR the expression of a set of 17 lncRNAs and 31 circRNAs selected from literature in 20 CRC tissues compared to normal adjacent tissues (NATs), and in serum exosomes of 20 CRC patients compared to 20 healthy individuals. We identified 8 ncRNAs (CCAT1, CCAT2, CDR1AS, HOTAIR, MALAT1, TUG1, UCA1 and ZEB2AS1) differentially expressed (DE) in tissues and 3 ncRNAs (circ16, TUG1 and UCA1) DE in serum exosomes. Through gene ontology analysis we verified the involvement of DE ncRNAs in pathways associated with tumor progression; we also evaluated the diagnostic accuracy of ncRNAs deregulated in serum exosomes through ROC curves, suggesting their possible application in CRC non-invasive diagnosis. We observed that inhibition of MAPK pathway, associated with CRC, altered the expression of HOTAIR, MALAT1, TUG1 and UCA1 in HCT-116 colon cancer cells treated with U0126 (MEK1/2 specific inhibitor) compared to untreated cells, suggesting a possible connection between ncRNAs and MAPKs. Also, we performed a computational analysis to identify mRNAs and miRNAs involved in CRC and characterized by correlation of expression and sequence complementarity with ncRNAs deregulated in CRC tissues.

Finally, we performed *in vitro* functional assays by silencing UCA1 expression through ASOs in HCT-116 cells and analysed the expression of its mRNA and miRNA targets, aiming to investigate the eventual degradation triggered by the interaction between ncRNAs and their miRNA/mRNA targets. Our study provide new data about aberrant expression of ncRNAs in CRC and their involvement in carcinogenesis; further analyses will be necessary to fully understand ncRNA molecular mechanisms and to evaluate their possible application in CRC diagnosis and therapy.

2. Introduction

A few years after the discovery of DNA structure, Francis Crick stated the central dogma of molecular biology: "*Once information has got into a protein it can't get out again*" (1); with this simple sentence he summarized the flow of genetic information through DNA, RNA and proteins (Figure 2.1).
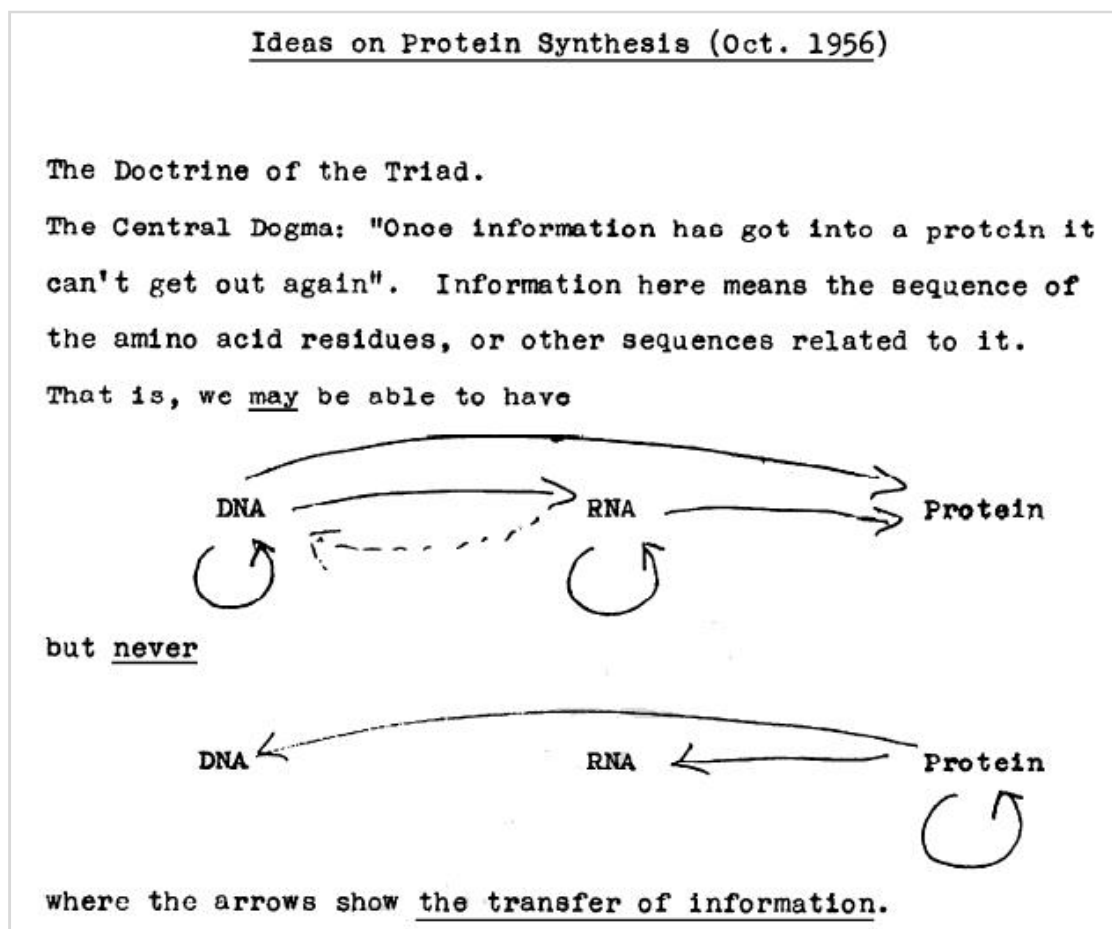
Figure 2.1: the original concept diagram published by Francis Crick in 1956.

Shortly after, François Jacob and Jacques Monod confirmed the central role of RNA in genetic information transfer from DNA to proteins (2). At that time, with few evidences, it was assumed that all biological processes in cells were regulated by proteins, characterized by more complex structures, while nucleic acids were considered genetic information reservoirs. To date, a

variety of new technologies and skills have been improved and a huge amount of sequencing data has been produced, leading researchers all over the world to go beyond these models and reconsider RNA role in cellular biology.

In 1990 an international consortium (including USA, United Kingdom, France, Australia, and China) launched the Human Genome Project, aiming to determine the entire sequence and to detect all genes of human genome (3). In 2001 a first draft of the human genome was published (4), followed by the final version in 2004 (5). Until then, scientists tried to indirectly estimate the total number of human protein-coding genes using many approaches, such as the typical vertebrate tissue mRNA complexity, the approximate ratio of typical gene size to the genome size, the number of CpG islands and the frequency of their association with known genes; finally it was supposed a total of 30,000-35,000 protein-coding genes in human genome (3). The final results of Human Genome Project revealed that our genome includes approximately 20,000-25,000 protein-coding genes (5), but this number has been recently corrected in about 19,000 genes (6). Surprisingly, this number is lower than expected, and very close to total protein-coding gene number of less complex organisms (*Caenorhabditis elegans*, 19,735 genes, *Gallus gallus*, 16,736 genes, *Drosophila melanogaster*, 14,889 genes, *Mus musculus* (21,839 genes) (7) (Figure 2.2).
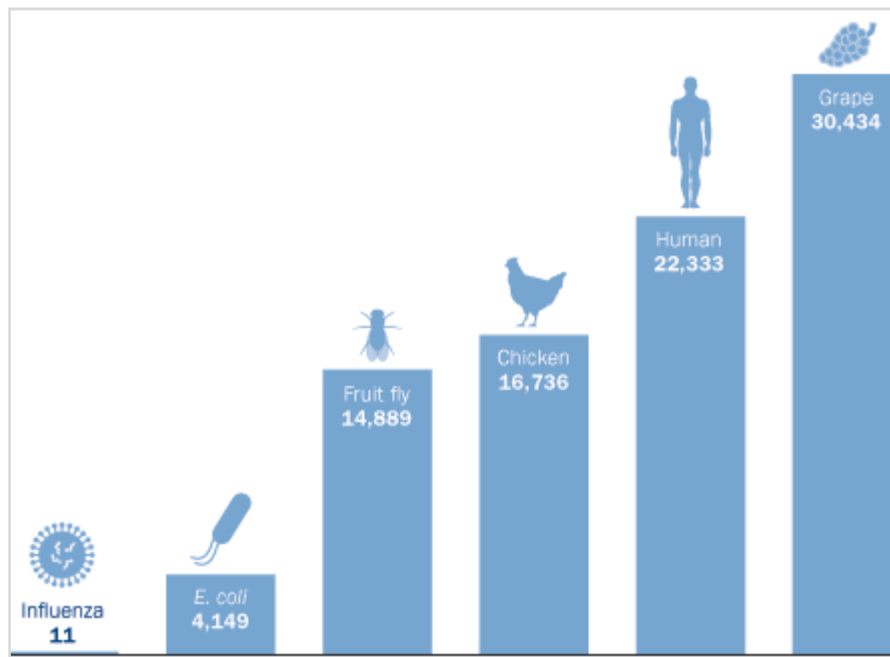
Figure 2.2: total number of protein-coding genes comparison between *Homo sapiens* and inferior organisms (from *https://www.sciencenews.org/article/more-chicken-fewer-grape*).

This discovery surprised scientists, which believed that *Homo sapiens* higher complexity was due to higher gene number compared to inferior organisms. Also, it was observed that all genes identified during the Human Genome Project represent less than 2% of the entire genome, while in other organisms, eukaryotic or prokaryotic, the coding fraction of genome is considerably higher (8) (Figure 2.3).
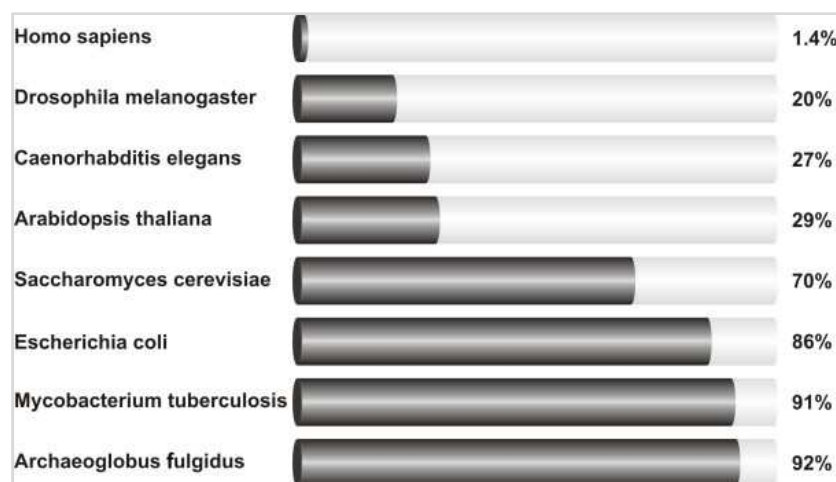
Figure 2.3: coding genome percentage comparison between Homo sapiens and other organisms, both prokaryotic and eukaryotic (from *Sana et al., J Transl Med. 2012*).

This illusory paradox was partially explained by alternative splicing events and post-translational modifications, which allow cells to produce more proteins with different functions from a single gene. However, scientists concluded that a large fraction of the human genome, even 98%, was represented by "*junk DNA*", that is DNA with no function.

In order to understand the molecular bases of *Homo sapiens* higher complexity and the "junk DNA" function, the ENCODE (*Encyclopedia of DNA Elements*) Project was launched with the purpose to identify all functional elements in the human genome (9). Final results, confirmed by following studies, showed that the 85-90% of human genome is pervasively transcribed: numerous new transcription start sites were identified, as well as many new transcripts, originated from genes located at non-coding regions or overlapping protein-coding *loci* (10, 11, 12). Since these new transcripts were not translated into proteins, they were defined non-coding RNAs (ncRNAs). It was recently estimated that non-coding gene number tends to exponentially increase ascending the evolutionary scale, while protein-coding gene number reaches a *plateau*: therefore, *Homo sapiens* higher complexity in comparison to inferior organisms with similar protein-coding gene number may be due to this new class of RNAs (13) (Figure 2.4).
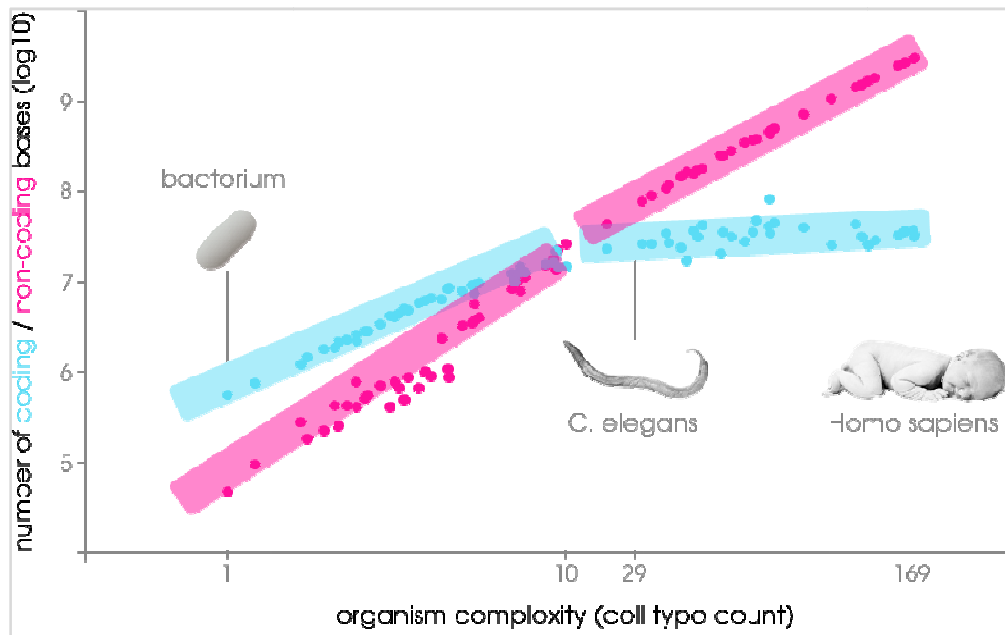
Figure 2.4: correlation between organism complexity and coding/non-coding genome fraction (from *Kristiaan Vanhercke, The power of RNA: discover next generation biomarkers and therapeutic targets*).

Next Generation Sequencing (NGS) new technologies led to the identification of ncRNA sequences, stressing their great structural heterogeneity. The same heterogeneity can also be observed in the structure of ncRNA genes: i) their length is extremely variable, from few tens of nucleotides to several hundreds of kilobases; ii) they lack ORFs (Open Reading Frames); iii) evolutionary conservation is often low or absent; iv) there is no preferential localization within the genome. Also, a relative tolerance to point mutations has been observed (14, 15). These features make ncRNA genes difficult to identify, which is why their exact number is to date unknown; however, it has been estimated that the human genome includes from 10,000 to 20,000 non-protein coding genes (16). Classifying these molecules is also difficult: two different classification are currently used, based on ncRNA length or function. According to length, ncRNAs are divided into i) long non-coding RNAs (lncRNAs), which are longer than 200 nucleotides, and ii) small non-coding RNAs, whose length is equal or inferior to 200 nucleotides; the latter includes ncRNA molecules deeply analysed, such as microRNAs (miRNAs), small interfering RNAs (siRNAs), and PIWI-interacting RNAs (piRNAs). The second classification, based on function, consists of i)

housekeeping ncRNAs, constitutively expressed and responsible of fundamental physiological processes; ii) regulatory ncRNAs, whose expression can be induced in different cell histotypes, in defined development and differentiation stages or following external *stimuli*. Housekeeping ncRNAs include transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), and telomerase RNAs, while regulatory ncRNAs comprise miRNAs, siRNAs, piRNAs and lncRNAs (14, 15).

2.1 Long non-coding RNAs

LncRNAs represent a highly heterogeneous class of ncRNAs, characterized by length greater than 200 nucleotides (cut-off value derived from biochemical techniques that allow to exclude all known classes of small RNAs) frequently reaching 100 kb (14). These transcripts are considered non-protein coding because they lack long ORF (more than 100 codons) and/or codon conservation; however, recent studies showed that some lncRNAs can code for small proteins (17). LncRNA genes are difficult to identify because of their variable features; indeed, they can:

 i. Be located within intergenic regions, generating long intergenic non-coding RNAs (lincRNAs) (18);

 ii. Be located inside introns of protein-coding genes (19);

 iii. Partially overlap untranslated regions (UTRs) or protein-coding gene promoters (20, 21);

 iv. Be pseudogenes whose transcripts (sense or antisense) control the expression of their protein-coding paralog genes (22, 23);

 v. Lie within transcribed ultra-conserved regions (tUCRs), characterized by high evolutionary conservation and located at intra- or intergenic regions (24);

 vi. Be located within the mitochondrial genome and regulated by proteins coded by nuclear genes (25).

The discovery of ncRNAs amazed researchers and started off a debate about the effective functions of these molecules, initially considered "transcriptional noise" (26, 27). Nevertheless, numerous studies have highlighted that ncRNA expression is tissue-specific, and it varies in different development or differentiation stages or in response to external *stimuli* (28, 29, 30, 31). These data allowed to go beyond the "transcriptional noise" hypothesis and sparked scientist interest in lncRNAs: recent studies showed that these molecules play an important role in biological processes, both physiological, such as cell cycle (32, 33), apoptosis (34, 35), differentiation (36), stemness (37), aging (38), and pathological (39, 40, 41). In particular, many researchers focused their attention on the possible involvement of lncRNAs in neoplastic processes, which by now seems to be demonstrated (42, 43, 44, 45, 46).

2.1.1 LncRNA biogenesis

Most of lncRNAs share with mRNAs the transcription by RNA polymerase II, the epigenetic regulation (histone modifications) and splicing signals (47); only a small part of them is transcribed by RNA polymerase I and III (48, 49). About half of lncRNAs are post-transcriptionally processed undergoing the same steps of mRNAs (5'-cap addition, polyadenylation, splicing), although some of them are monoexonic and non-polyadenylated (47, 50). Poly(A) tail protects 3'-ends from enzymatic degradation; in its absence, the presence of alternative structures with similar function, such as triple helix, has been shown. This kind of structure has been identified at 3'-end of MALAT1 (metastasis-associated lung adenocarcinoma transcript 1), a 8.7-kb lncRNA transcribed from a gene located at chromosome 11. During MALAT1 processing, the primary transcript loses from its 3'-end a 61-nucleotide fragment, called mascRNA (MALAT1-associated small cytoplasmic RNA); hence, the new 3'-end of mature transcript presents a stem-loop structure and three conserved motifs, composed of two U-rich and one A-rich tracts; this A-rich tract is similar to a poly(A) tail, but it's too short to be recognized by poly(A)-binding proteins. These three motifs

interact with each other and form a triple helix structure that stabilizes MALAT1 3'-end and protects it from enzymatic degradation (51) (Figure 2.5).
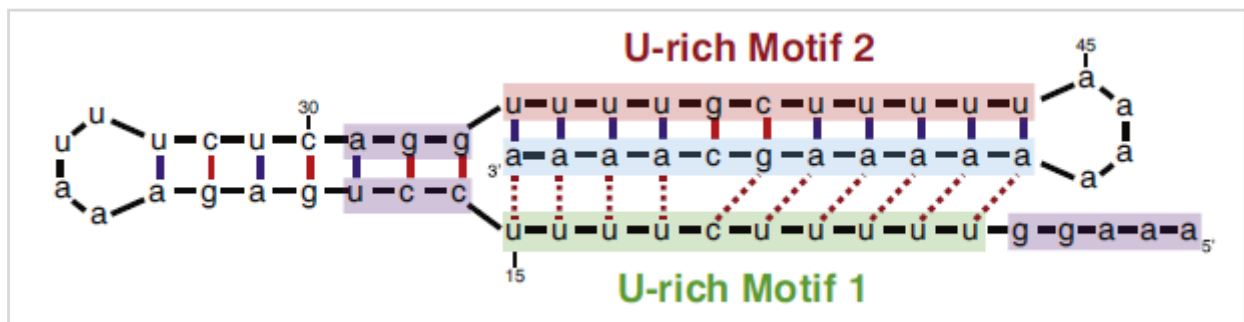
Figure 2.5: triple helix structure stabilizing MALAT1 3'-end in absence of poly(A) tail. Dashed lines show base triplets (from *Wilusz, Biochim Biophys Acta. 2016*).

2.1.2 LncRNA functions

LncRNAs play a crucial role in cellular physiology through a plethora of functions and molecular mechanisms. It is now clear that they can regulate gene expression at different levels (transcriptional, post-transcriptional, translational, post-translational, epigenetic) (52) (Figure 2.6); however, the exact molecular mechanism has been elucidated only for few lncRNAs.
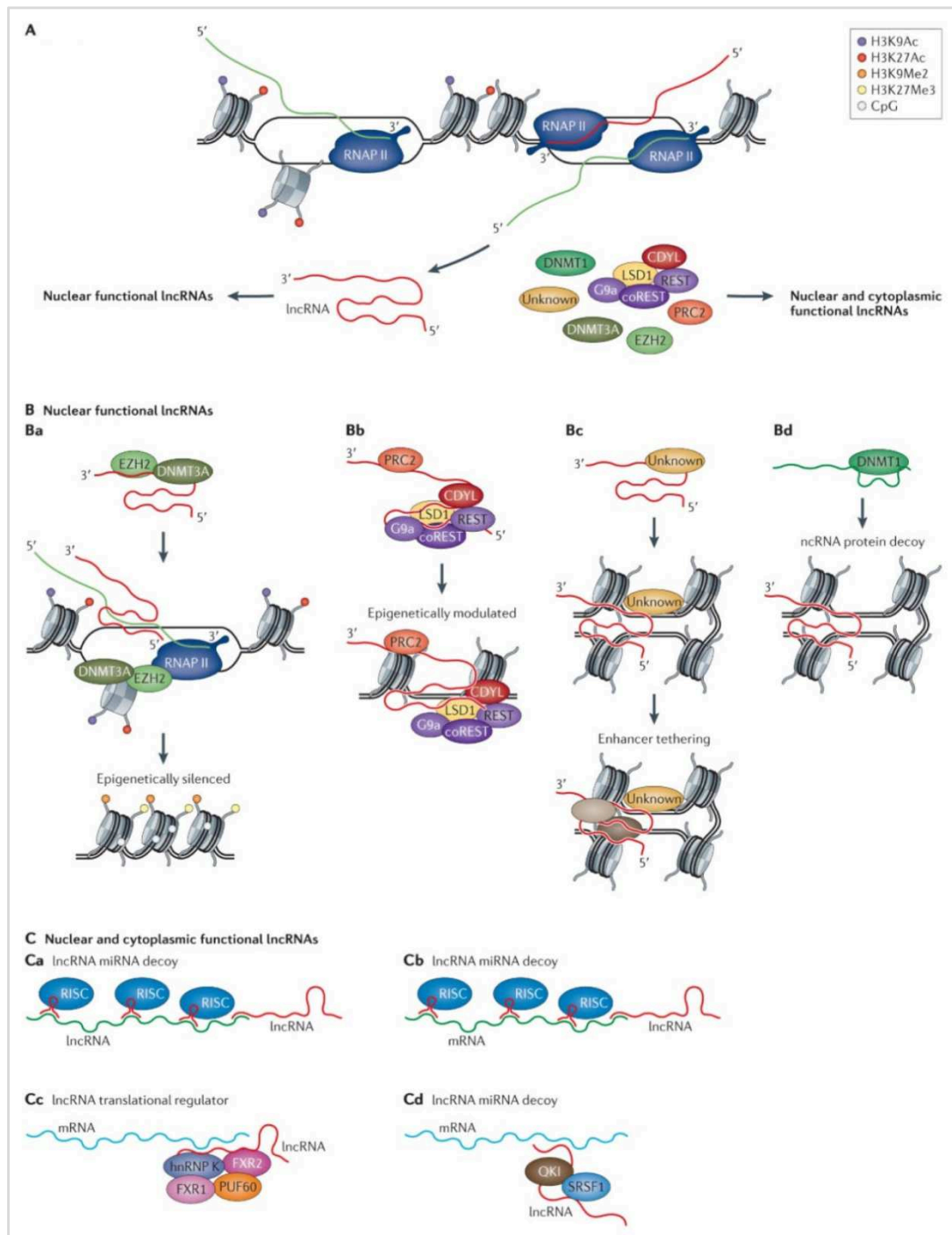
Figure 2.6: lncRNAs regulate cell physiology performing several functions (from *Morris and Mattick, Nat. Rev. Genet. 2014*).

Xist (X inactive specific transcript) is one of the first lncRNAs identified (53, 54). Xist is a 19-kb lncRNA transcribed by the X chromosome; it is involved in the X-chromosome inactivation process, the epigenetic mechanism leading to the random inactivation of one of the two X

chromosomes in female individual cells. X inactivation depends on histone modifications (in particular lysine methylation) performed by PRC1 and PRC2 (Polycomb-repressing complex 1 and 2), which in turn are recruited on the inactive X chromosome by Xist (55). Similarly, other lncRNAs can modulate gene expression by recruiting chromatin remodeling complexes: one of the best known is HOTAIR (Hox transcript antisense intergenic RNA), a 2.2-kb lncRNA transcribed from the antisense strand of HOXC (Homeobox C) cluster on chromosome 12 (56, 57). HOTAIR also interacts with PRC2, inducing the trimethylation of histone H3 lysine 27, and the consequent transcription inhibition, of HOXD (Homeobox D) cluster on chromosome 2 (58). Moreover, HOTAIR interacts not only with PRC2 through a 5'-domain, but also with LSD1 (lysine demethylase 1A) through a 3'-domain, thus modulating demethylation and transcriptional activation of the same cluster (57, 59). TUG1 (taurine up-regulated 1) is a 7.5-kb lncRNA located at chromosome 22; its expression is p53-mediated thanks to many highly conserved p53-binding sites included in its promoter. After DNA damage, p53 binds TUG1 promoter and activates TUG1 expression; the lncRNA in turn recruits PRC2 and inhibits expression of cell cycle regulator genes (60).

LncRNAs can also bind transcriptional factors, recruiting them on DNA or sequestering them, therefore regulating the expression of specific genes. PANDAR (promoter of CDKN1A antisense DNA damage activated RNA) is a 1.5-kb polyadenylated and capped lncRNA, antisense transcript of CDKN1A (cyclin-dependent kinase inhibitor 1A) on chromosome 6. Its expression is regulated by p53 after DNA damage; PANDAR binds the transcription factor NFYA (nuclear transcription factor Y subunit alpha), impeding it to bind the promoters of pro-apoptotic genes (61).

Some lncRNAs can interact with DNA inducing loop formation, thus moving closer two otherwise very distant regions; these lncRNAs act as "bridges" between promoters and enhancers or silencers. Their mechanism of action is not known yet, although it has been proposed the formation of a RNA-DNA-DNA triplex or a RNA-protein-DNA complex (62). The DHFR (dihydrofolate

reductase) *locus* on chromosome 5 contains two promoters: the major promoter codes for the DHFR mRNA, while the upstream minor promoter codes for a ncRNA which negatively regulates mRNA levels in *cis* and in *trans*: the DHFR ncRNA interacts with the major promoter, forming a stable complex, and with TFIIB (transcription factor IIB), causing PIC (pre-initiation complex) dissociation and transcription inhibition (63) (Figure 2.7).
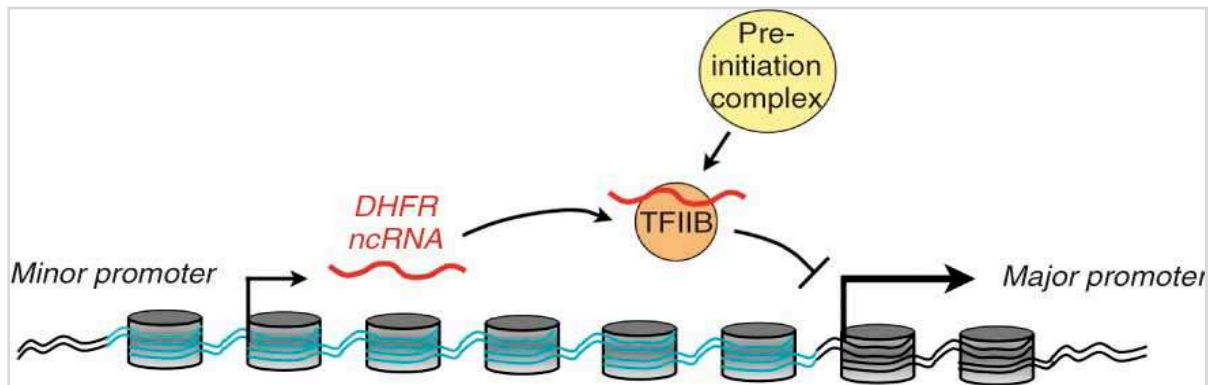
Figure 2.7: DHFR ncRNA inhibits the transcription of DHFR mRNA by interacting with the major promoter and TFIIB (from *Martianov et al., Nature 2007*).

LncRNAs can act at post-transcriptional level modulating mRNA stability, translation and splicing. BACE1AS (BACE1 antisense RNA) is a 825-nucleotide lncRNA transcribed from a gene located at chromosome 11, apparently involved in Alzheimer's disease pathogenesis. BACE1AS is the natural antisense of BACE1 (beta-secretase 1), a transmembrane protease that catalyzes the first step in amyloid beta peptide synthesis; amyloid beta peptide accumulates in Alzheimer's patient cerebral neurons, generating the typical amyloid plaques. Thanks to sequence complementarity, BACE1AS binds BACE1 mRNA, thus stabilizing it and inducing increased amyloid beta production (64). ZEB2AS1 (ZEB2 antisense RNA 1), antisense transcript of ZEB2 (zinc finger E-box binding homeobox 2), is a 430-nucleotide lncRNA transcribed from chromosome 2. ZEB2 plays a key role in neoplastic processes by repressing E-cadherin expression, thus promoting epithelial-mesenchymal transition (EMT). ZEB2AS1 is abundantly expressed in mesenchymal cells through the activation of an alternative promoter; the lncRNA binds the 5'-UTR of ZEB2 mRNA,

masking splicing sites and preventing the removal of a long intron (about 2.5 kb) including an internal ribosome entry site (IRES). So, in presence of its antisense, ZEB2 mRNA is recognised by ribosomes and translated, producing a transcription factor that induces E-cadherin and EMT repression (20) (Figure 2.8).
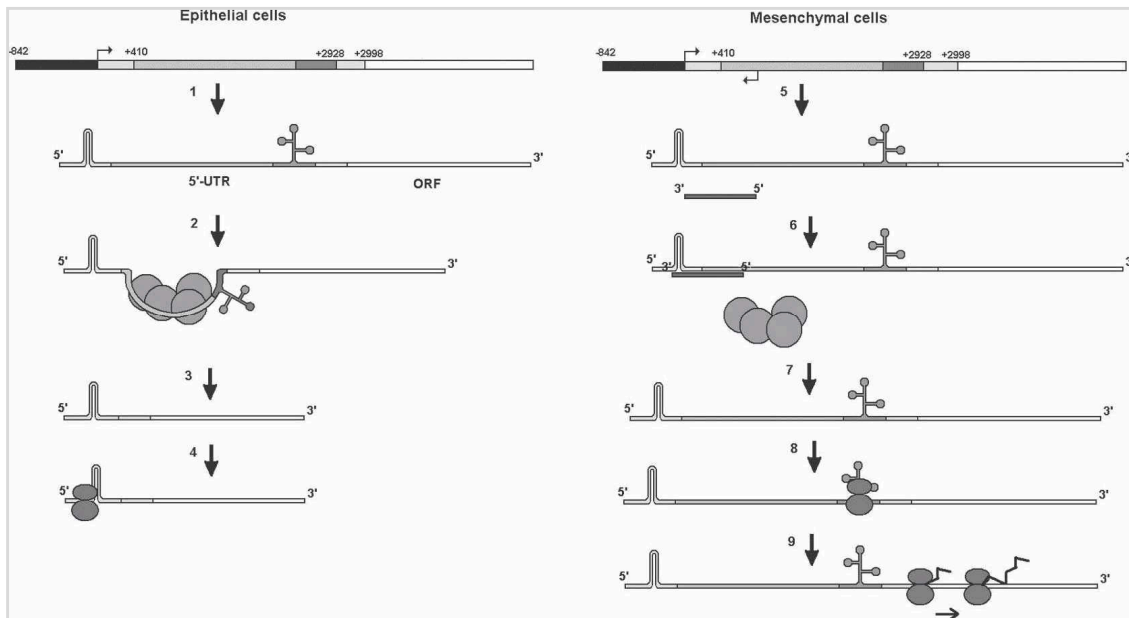
Figure 2.8: ZEB2 mRNA alternative splicing mediate by its antisense ZEB2AS1 in mesenchymal cells; ZEB2 main promoter is shown in black, IRES-containing 3-kb 5'-UTR in gray, ZEB2 ORF in white (from *Beltran et al., Genes Dev. 2008*).

Another function of lncRNAs is to regulate cellular localization of proteins, thus modulating their activity. NRON (non-protein coding RNA, repressor of NFAT), a 2.7-kb transcript of chromosome 9, controls the trafficking of NFAT (nuclear factor of activated T cells) transcription factor: NFAT is normally contained in the cytoplasm, but it is imported to the nucleus in response to calcium-dependent signals. NRON interacts with several proteins, including members of the importin-beta superfamily, thus inhibiting NFAT nuclear accumulation (52, 65).

Some lncRNAs play structural roles, allowing correct formation of specific cellular structures. NEAT1 (nuclear paraspeckle assembly transcript 1) is a 4-kb monoexonic polyadenilated lncRNA transcribed from chromosome 11; it is contained in the nucleus overlapping SC35 splicing

factor enriched regions; NEAT1 presence in the nucleus is essential for the formation of paraspeckles, nuclear domains involved in nuclear retention of specific mRNAs (66).

Recent studies revealed a connection between lncRNAs and small non-coding RNAs, such as miRNAs: some lncRNAs can be miRNA precursors or can interact with them, sequestering them and inhibiting their transcriptional repression activity; this function has been defined "miRNA sponge". H19, the first lncRNA identified (67), is transcribed from a gene localized on chromosome 11, near IGF2 (insulin-like growth factor 2) *locus*; both genes are epigenetically silenced: only the maternal allele of H19 and the paternal allele of IGF2 are expressed, while the other ones are silenced by methylation (68). H19 gene produces a 2.3-kb capped, polyadenilated and spliced lncRNA which is expressed from the earliest stages of embryonic development in several organs and quite completely silenced after birth. H19 is the precursor of miR-675, which targets, among others, RB1 (retinoblastoma 1), a key cell cycle regulator (69). Plus, H19 contains many binding sites for let-7 miRNA family, well known tumor suppressor miRNAs; H19 can bind and sequester let-7 miRNAs, acting as "miRNA sponge" (70). TUSC7 (tumor suppressor candidate 7), a 2.1-kb lncRNA transcribed from chromosome 3, also acts as "sponge" for miR-211 thanks to two binding sites lying within its fourth exon (71). MALAT1 is highly expressed and retained inside the nucleus, where it acts as "riboregulator" modulating the expression of its target genes, involved in cell cycle regulation (51, 72). MALAT1 lacks poly(A) tail, and its 3'-end folds to form a tRNA-like secondary structure, recognised and cleaved by RNAse P; this cleavage creates a short fragment, subsequently processed by enzymes involved in tRNA biogenesis; at the end of the process, a 61-nucleotide mature transcript called mascRNA is produced (Figure 2.9). MascRNA is exported to the cytoplasm but, despites its structure, it doesn't contribute to protein synthesis, and its function is still unknown (51). A recent study suggested a mascRNA involvement in immune system regulation: high levels of mascRNA (but not of MALAT1) have been observed in human immune system circulating cells; furthermore, mascRNA downregulation in monocytes-macrophages

strongly affects the expression of fundamental genes, including FASLG (Fas ligand), FAS (Fas cell surface death receptor), TNF-α (tumor necrosis factor alpha), and IL6 (interleukin 6) (73).
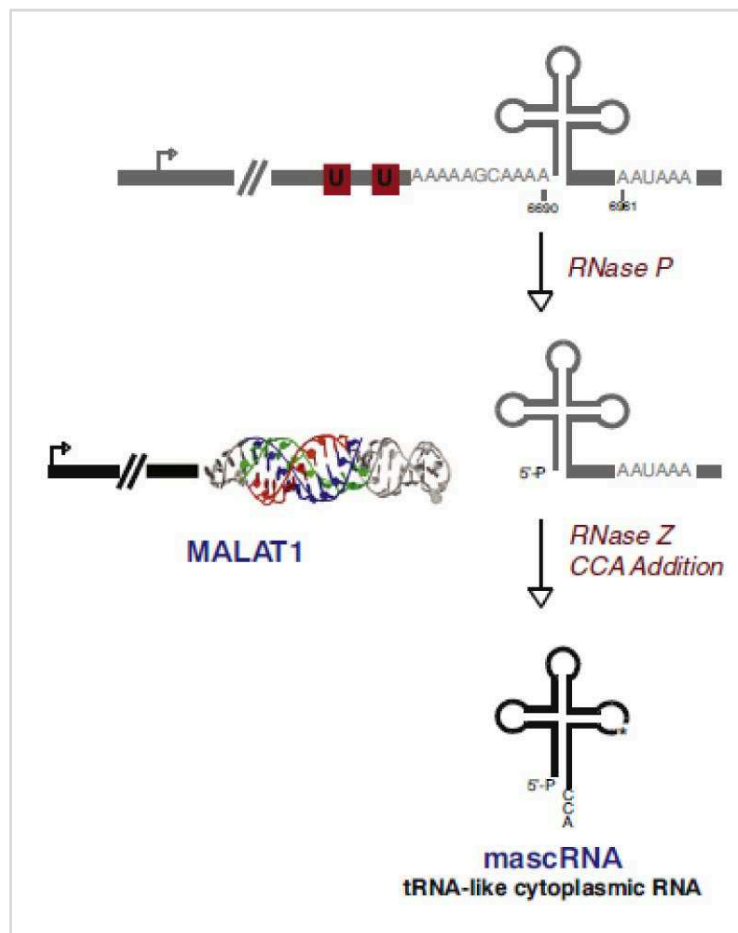
Figure 2.9: mascRNA biogenesis mediated by RNAse P cleavage of 3'-end of MALAT1 (from *Wilusz, Genes Dev. 2012*).

Therefore, it is clear that lncRNAs perform multiple functions through very heterogeneous molecular mechanisms, escaping any attempt at classification.

2.2 Circular RNAs

Next Generation Sequencing techniques led to the identification of a unique class of lncRNAs represented by circular non-coding RNA molecules, called circular RNAs (circRNAs). The first single-stranded circular RNA molecule was identified in 1976 within viroids, infectious agents similar to viruses but lacking capsids (74); since then other circular RNAs have been

observed in other organisms, including animals, but they have been considered by-products of mRNA processing derived from splicing errors (75). Recently, several new circRNAs previously unknown have been identified, characterized by high expression levels in mammalian cells and, sometimes, by high evolutionary and expression conservation, suggesting a specific function in cellular physiology (76, 77).

2.2.1 CircRNA biogenesis

As it is known, one of the characteristics distinguishing prokaryotes and eukaryotes is the alternation of introns (non-coding segments) and exons (coding segments) in eukaryotic gene structure; after transcription, pre-mRNA undergoes a processing which includes splicing, that is intron removal by a ribonucleoprotein complex called spliceosome. Splicing has been widely studied; it consists of two phases, both regulated by the spliceosome: i) the 2'-OH of an intronic base, called branchpoint (usually an adenosine), performs a nucleophilic attack on the 5'-splicing site, localized at exon-intron boundary and identified by consensus dinucleotide GU, thus forming a phosphodiester 2',5'-bond and releasing the 3'-OH end of the exon; ii) this 3'-OH in turn attacks on the 3'-splicing site, localized at intron-exon boundary and identified by consensus dinucleotide AG: hence, the two exons are bound through a phosphodiester 3',5'-bond, while the intron is removed as a lariat, because of a 2',5'-bond (Figure 2.10). This is an essential event finely regulated but also flexible, that allows cells to produce several protein variants modulating splicing sites choice (alternative splicing) (78).
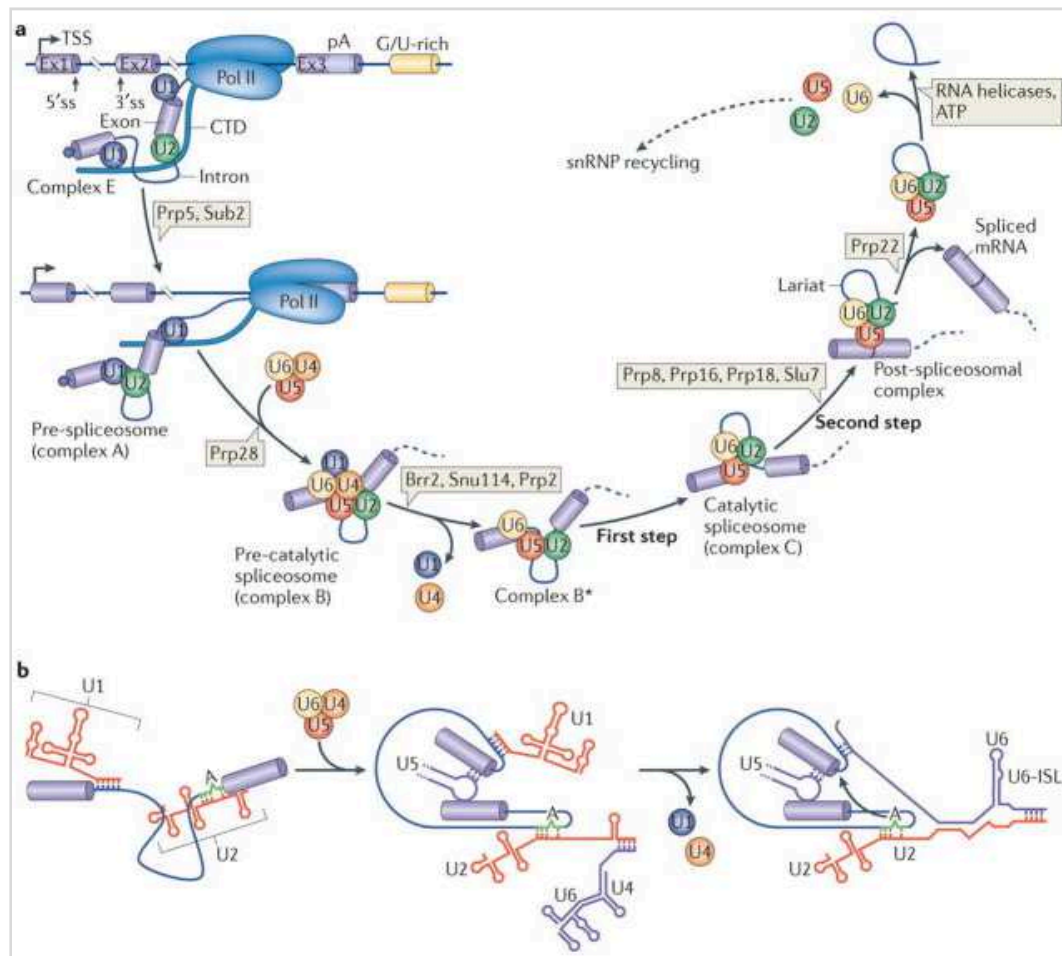
Figure 2.10: splicing mechanism regulated by the spliceosome (from Matera *et al,. Nat Rev Mol Cell Biol. 2014*).

Many circRNAs are transcribed from protein-coding gene exons, so they are called exonic circRNAs (ecircRNAs); it has been proposed that ecircRNA biogenesis ecircRNA biogenesis is catalyzed by the spliceosome following the same splicing mechanism. This hypothesis is supported by several evidences: both splicing inhibitors and mutagenesis of splicing sites (GU and AG consensus sequences) significantly reduce circular RNA levels (79). The exact mechanism of ecircRNA biogenesis is still unknown: however, it seems that spliceosome regulates backsplicing (or inverse splicing) events, where a downstream 5' splicing site (donor) is bound to an upstream 3' splicing site (acceptor); this mechanism produces a head-to-tail junction, typical of backsplicing events (80) (Figure 2.11).
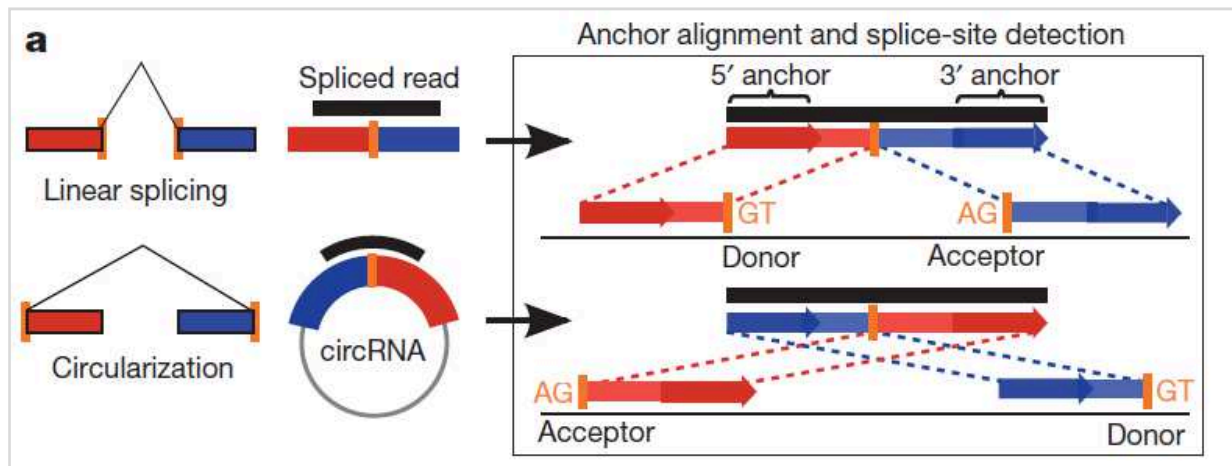
Figure 2.11: head-to-tail junction of two exons typical of backsplicing events (from *Memczak et al., Nature 2013*).

Three models for ecircRNA biogenesis have been proposed:

i. Lariat-driven circularization (Figure 2.12a): during splicing events, some exons can be skipped and removed together with introns (exon skipping); ecircRNAs can be produced by a second splicing on exon-containing lariats, before they are unraveled by debranching enzymes (77);

ii. Intron pairing-driven circularization (Figure 2.12b): intronic complementary motifs flanking circularized exons can anneal and promote circularization (77);

iii. Resplicing-driven circularization (Figure 2.12c): sometimes, a second splicing event can happen on a mature mRNA in presence of specific splicing signals. The first event is regulated by canonical splicing sites and removes introns from pre-mRNA; the second one requires cryptic splicing sites, which are recognised only in specific conditions, and induces exon circularization. Re-splicing is likely an aberrant phenomenon, often observed in carcinogenesis (81).
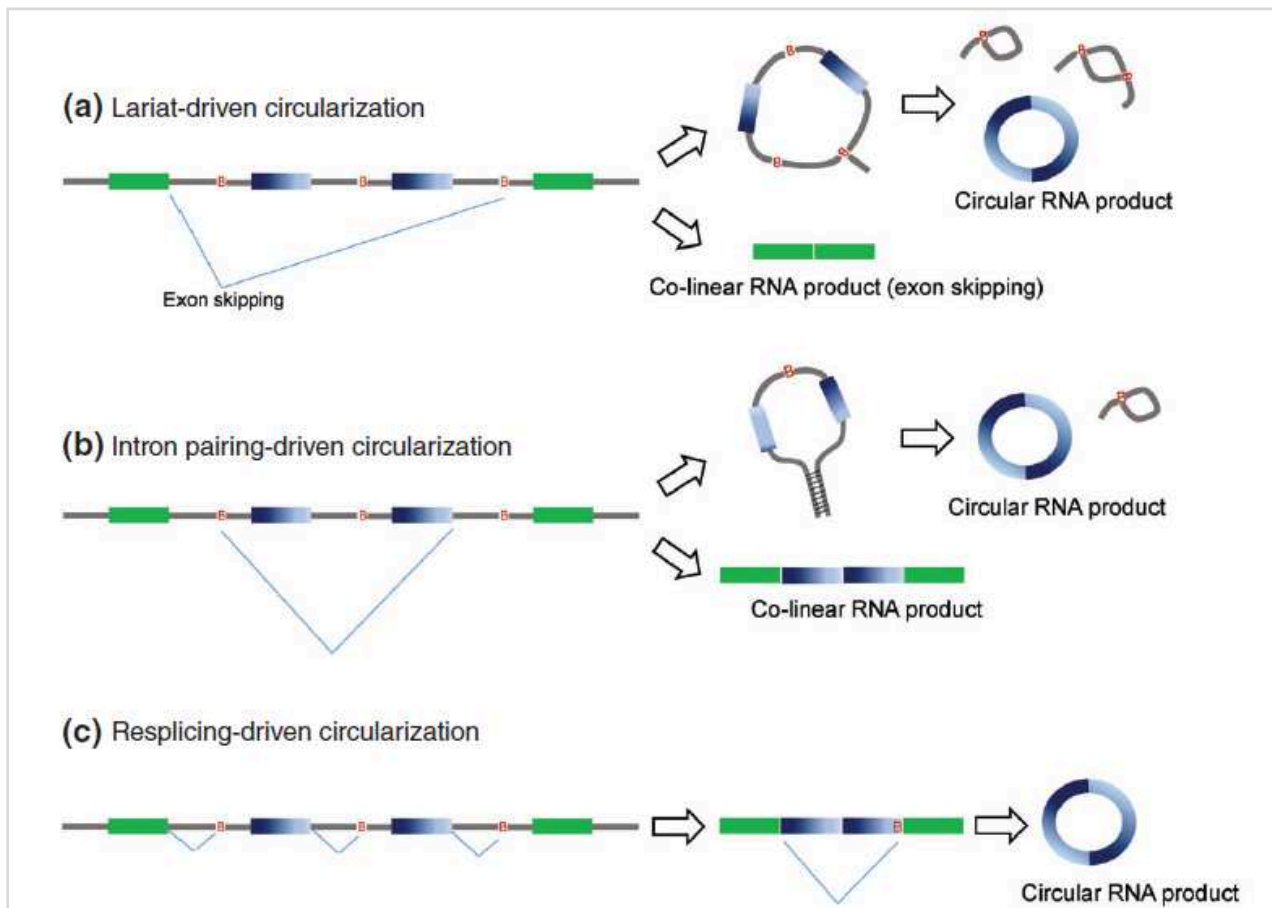
Figure 2.12: three models proposed to explain ecircRNA biogenesis a) Lariat-driven circularization: an exon-containing lariat undergoes a second splicing; b) intron pairing-driven circularization: the annealing of intronic complementary motifs flanking circularized exons promotes circularization; c) resplicing-driven circularization: a second splicing event on a mature mRNA induces exon circularization (from *Matera et al,. Nat Rev Mol Cell Biol. 2014*).

As suggested by the intron pairing-driven circularization model, the presence of reverse complementary sequences (RCSs) in flanking introns can move exons closer and promote their circularization; indeed, several studies reported a positive correlation between flanking intron RCSs (often represented by Alu repeats in primates) and exon circularization (77, 82, 83). However, RCSs are not essential, and sometimes longer RCSs can even inhibit circularization through increased base-pairing stability; moreover, a gene can include many RCS copies competing with each other, so alternative circularization events can occur, thus raising ecircRNA number produced by a single gene (79, 84). Another factor promoting ecircRNA biogenesis is flanking intron length, that is

three- to fivefold longer than in randomly selected introns both in human genome, (77, 82), and in other species (79). Again, long flanking introns are not essential for ecircRNA biogenesis, but longer introns are likely to include more cis-acting regulatory element (77, 85).

In addition to ecircRNAs, circular transcripts originated from spliced introns, called circular intronic RNAs (ciRNAs), have been identified. They are originated from the lariat, which is not debranched but circularized thanks to opportune signals, that are: i) a 7-nucleotide GU-rich sequence sited near to 5' splicing site; ii) an 11-nucleotide C-rich sequence near to branchpoint; molecules participating in this process are still unknown (86) (Figure 2.13).
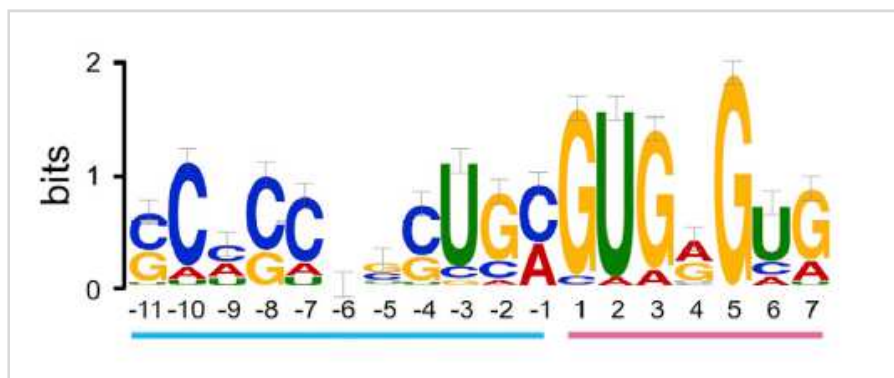
Figure 2.13: consensus sequences needed for ciRNA biogenesis: a 7-nucleotide GU-rich sequence sited near to 5' splicing site (pink) and an 11-nucleotide C-rich sequence near to branchpoint (blue) (from *Zhang et al., Mol Cell. 2013*).

It has been hypothesized that the same factors regulating canonical (or linear) splicing also modulate backsplicing events; therefore, the two phenomena compete with each other to bind these molecules. A recent study showed that Quaking (QKI), an RNA-binding protein regulating alternative splicing, controls the expression of several circRNAs during EMT. QKI binds RCSs, moving circularized exons closer; the addition of QKI-binding sites in flanking introns highly increases circRNA production (87). ADAR (adenosine deaminase acting on RNA) protein family includes enzymes involved in RNA editing, a process modifying some nucleotides during pre-mRNA maturation. In particular, ADAR enzymes bind the adenosines of double-stranded RNA molecules, converting them into inosines through deamination. Recently, it has been observed that

the downregulation of ADAR enzymes induces increased circRNA expression, suggesting that these enzymes could negatively regulate circRNA biogenesis (83).

2.2.2 CircRNA functions

CircRNAs have been initially considered as splicing errors. Successively, some circRNAs showed evolutionary conservation between human and mouse and among species, probably because of the conservation of splicing sites; this observation led to the hypothesis of a molecular function performed by this class of ncRNAs (79). Another evidence supporting this hypothesis is the high number of different circRNAs produced by cells: according to the database circBase (http://www.circbase.org/), approximately 100.000 circRNAs have been identified in human cells (89). It has been estimated that circRNAs represent 1-10% of the entire cell transcriptome (76, 85, 90); most of them have very low expression levels (0.1-1%) compared to co-linear mRNAs, although sometimes the circular transcript exceeds the linear one (77); moreover, circRNAs are often tissue-specific, so they are expressed only in some cell histotypes, even if they tend to be more abundant in brain tissues (79).

Several studies reported the presence of ecircRNAs in the cytoplasm, while their biogenesis occurs within the nucleus: nucleus export systems likely regulate the transport of these molecules (76, 77, 90). One of ecircRNA functions, also performed by lncRNAs, is to bind miRNAs and modulate their action. Some ecircRNAs are strongly associated with Argonaute (Ago) proteins, key factors of miRNA-induced gene silencing (RNA interference, RNAi), and include many miRNA-binding sites. The most known example is represented by a circRNA called CDR1AS/ciRs-7 (CDR1 antisense) and its 80 binding sites for miR-7. CircRNA-miRNA interaction causes a reduced miRNA availability within cell cytoplasm, and consequently a reduced translational repression of specific mRNA targets: increased CDR1AS expression induces an incremented expression of proteins coded by mRNA targets of miR-7. On the contrary, CDR1AS

downregulation provokes reduced protein levels of miR-7 targets, because miR-7 is not sequestered by the circRNA (80, 91) (Figure 2.14). A bioinformatic study showed that circRNAs present miRNA-binding sites lacking polymorphisms, strengthening their miRNA sponge or miRNA reservoir function (92); nevertheless, most of circRNAs include a number of miRNA-binding sites comparable to co-linear mRNAs, suggesting that only some circRNAs act as miRNA sponge (90, 93).
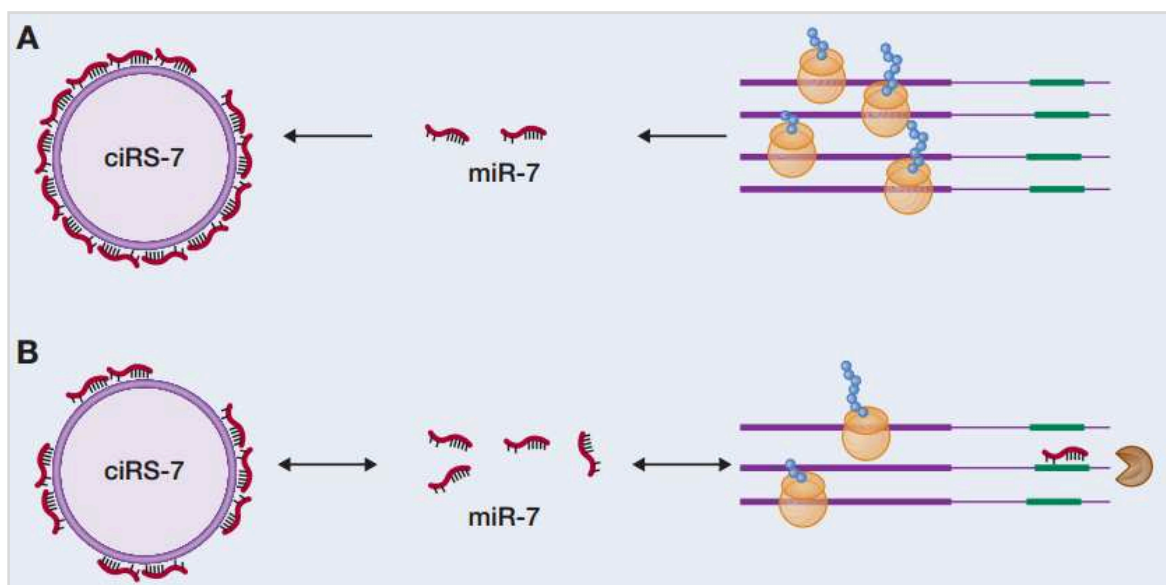
Figure 2.14: CDR1AS/ciRS-7 acts as miR-7 "sponge" (from *Hansen et al., Cancer Res. 2013*).

Moreover, ecircRNAs can modulate parental gene expression through transcription regulation. Recent studies showed that some ecircRNAs are retained within the nucleus and interact with RNA polymerase II, suggesting an involvement in gene expression regulation; these ecircRNAs contain, in addition to exons, non-spliced introns, and therefore they have been called exon-intron circRNAs (EIcircRNAs) (94, 95, 96). EIcircRNAs are prevalently expressed in the nucleus; they promote parental gene transcription by binding small nuclear ribonucleoprotein U1 (snRNPU1), a splicing regulator also involved in the initial stages of transcription; the EIcircRNA-snRNPU1 complex interacts with RNA polymerase II on the promoter of the parental gene, promoting its expression (94) (Figure 2.15).
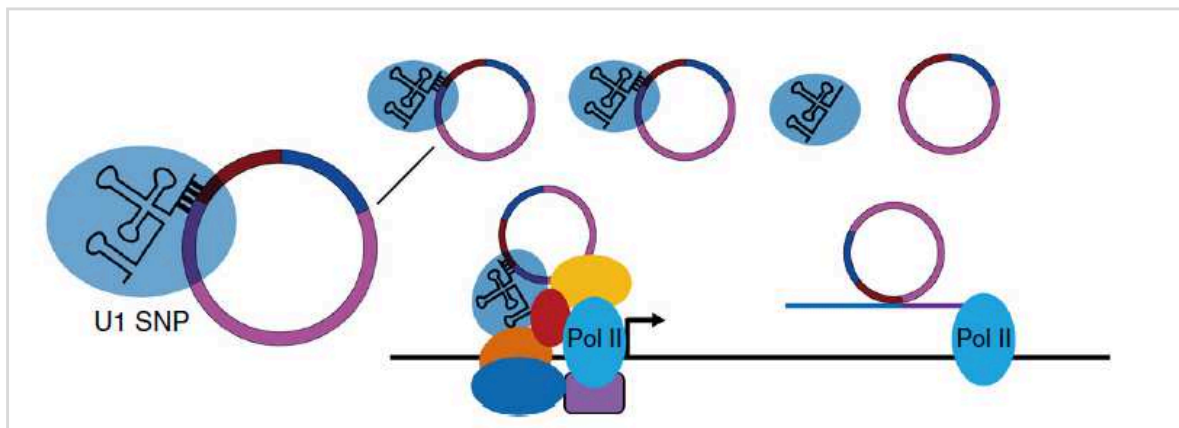
Figure 2.15: positive regulation of parental gene transcription mediated by EIcircRNAs: they bind snRNPU1 forming a complex which interacts with RNA polymerase II, promoting transcription (from *Li et al., Nat Struct Mol Biol. 2015*).

Another study suggested that circRNA biogenesis itself reduces co-linear mRNA expression. MBNL1 (muscleblind like splicing regulator 1) gene codes for a splicing regulating protein, but also for a circRNA called circMbl, produced co-transcriptionally. circMbl flanking introns include several binding sites for MBNL protein: when highly expressed, MBNL protein binds circMbl, promoting its circularization to the detriment of splicing. Thus, the cell reacts to MBNL accumulation in two ways: i) splicing is inhibited, reducing mature mRNA production; ii) protein availability is decreased through its circMbl-mediated sequestering (97).

Despite many ecircRNAs include ORFs, to date contrasting evidences suggesting peptide production has been reported; therefore, the possibility that ecircRNAs act as mRNAs needs further investigations (98).

CiRNA functions are not fully understood. A recent study showed that ciRNAs present reduced evolutionary conservation between human and mouse, accumulate in the nucleus and contain few miRNA-binding sites; these observations suggest different functions and mechanisms compared to ecircRNAs. CiRNAs interact with phosphorylated RNA polymerase II, suggesting that they act as positive regulator of gene expression, consistently with their nuclear localization. Interestingly, ciRNA expression is positively correlated with parental gene levels: it has been hypothesized that ciRNAs promote co-linear mRNA expression, as suggested by their accumulation

in the sites of synthesis. Low evolutionary conservation can be due to ciRNA origin: since introns don't code for functional proteins, they undergo no or very low evolutive pressure and accumulate more mutations compared to exons (86).

2.3 Small non-coding RNAs

Small non-coding RNAs, characterized by length equal or inferior to 200 nucleotides, have been widely studied for many years. By now, it is well known that they regulate many biological processes: i) RNA synthesis, processing and translation and transcriptional initiation [piRNAs, PASRs (promoter-associated small RNAs)]; ii) RNA maturation (snoRNAs); iii) RNA degradation or translation inhibition (miRNA, siRNA) (15). Among all small non-coding RNAs, the most known and studied are definitely miRNAs, whose biogenesis and function as post-transcriptional repressors have been well characterized.

2.3.1 MiRNA biogenesis

MicroRNAs are a class of endogenous single-stranded RNAs consisting of approximately 18-25 nucleotides and expressed both in animals and in plants. MiRNA-coding genes are located at all chromosomes in humans, except for Y chromosome; frequently, genes coding for different miRNAs are located at adjacent *loci* on the same chromosome, forming clusters (e.g. let-7 cluster on chromosome 9, including hsa-let-7a-1, hsa-let-7f-1 e hsa-let-7d). Transcription of a cluster is simultaneous and generates a polycistronic primary transcript, successively processed into single miRNAs. MiRNAs belonging to the same cluster are often related to each other, suggesting that clusters derive from gene duplication; consequently, miRNAs from the same cluster are often functionally correlated, since function depends on sequence. Genomic localization of miRNA-coding genes is highly heterogeneous; indeed, they can be located at: i) intergenic regions, ii) protein-coding gene introns, iii) non-coding gene introns, iv) non-coding gene exons; moreover,

some miRNA-coding genes have an independent promoter (99). According to miRBase database (http://www.mirbase.org/), to date 1881 precursors and 2588 mature miRNAs have been identified, even if a recent study hypothesized a higher number of miRNA-coding genes in human genome (100).

MiRNA biogenesis starts in the nucleus; transcription is prevalently performed by RNA polymerase II, rarely by RNA polymerase III; transcription produces a long (many kb) primary transcript called pri-miRNA, which folds into a double-stranded hairpin structure; the hairpin includes a loop and a double-stranded stem, made up of 33 complementary base pairs (bp), and ends with two long single-stranded traits (101, 102, 103). Successively, pri-miRNA undergoes two sequential cleavages: the first cleavage occurs in the nucleus by Drosha, the second one in the cytoplasm by Dicer, two endonucleases including conserved RNAse III catalytic domains (RIIIDa, RRIIDb); these enzymes act in association with proteins containing double-stranded RNA-binding domains (dsRBDs) (99). The first cleavage performed by a complex including Drosha and Pasha (DGCR8) converts the pri-miRNA into a 700-nucleotide long molecule with a hairpin structure called pre-miRNA, which is exported to the cytoplasm by exportin-5 (Exp5), a GTP-dependent nuclear/cytoplasmic transporter (104). In the cytoplasm, pre-miRNA interacts with a big multiprotein complex called RISC loading complex (RLC), consisting of the endonuclease Dicer, TRBP (Tar RNA Binding Protein), which contains three dsRBDs, PACT (protein activator of PKR), and Argonaute protein Ago-2, an RNAse with catalytic function (105, 106, 107, 108). Argonaute (Ago) proteins are expressed in all eukaryotes and contain specific domains: PAZ and MID domains to bind RNA targets at 3'-end, and PIWI domain al 5'-end to catalyze RNA cleavage (109). Dicer catalyzes the cleavage of the hairpin loop, producing a 22-nucleotide long double-stranded miRNA characterized by mismatches and two protrusive nucleotides at both 3'-ends. This RNA duplex is released from RLC and splits into two strands, defined: i) guide strand, complementary to the targets, and therefore functional; ii) passenger strand (miRNA*), initially

considered non-functional and degraded (110) (Figure 2.16). Recent studies showed that passenger strand also is functional and accumulates in the cytoplasm (111, 112, 113): by now, it is known that every miRNA-coding gene originates two mature molecules, defined -5p and -3p depending on the duplex strand from which they derive.
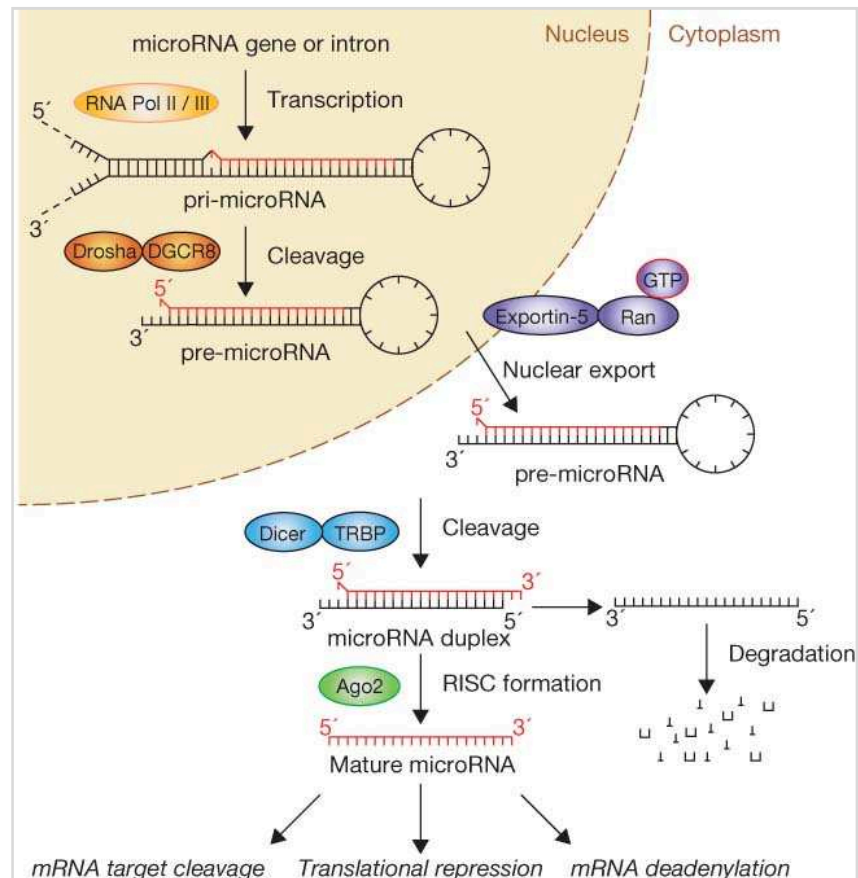
Figure 2.16: miRNA biogenesis starts in the nucleus with the transcription of a long pri-miRNA performed by RNA polymerase II or III; the pri-miRNA folds into a hairpin structure and undergoes an enzymatic cleavage performed by Drosha, generating the pre-miRNA, characterized by a stem-loop structure. The pre-miRNA is exported to the cytoplasm; the hairpin structure undergoes a cleavage performed by Dicer, generating a 22-nucleotide long double-stranded miRNA (from *Winter et al., Nat Cell Biol. 2009*).

A particular miRNA subfamily includes genes located within protein-coding gene introns, which code for miRNA called mirtrons. Mirtron biogenesis is Drosha-independent: after splicing, they are processed into a hairpin RNA, which is exported to the cytoplasm by exportin-5 and directly cleaved by Dicer (114) (Figure 2.17).
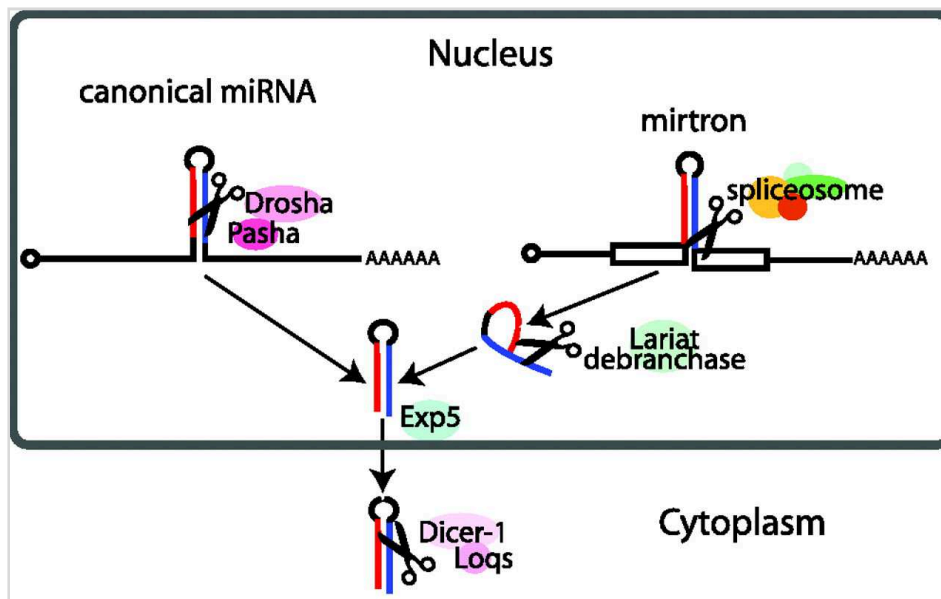
Figure 2.17: mirtron biogenesis is Drosha-independent: after removal, the intron lariat is exported to the cytoplasm by exportin-5 and cleaved by Dicer (from *Ruby et al., Nature. 2007*).

2.3.2 MiRNA functions

All miRNAs perform the same function of post-transcriptional repressors of gene expression through the same molecular mechanism.

Mature miRNAs bind Ago-2 protein and are led to the RNA-induced silencing complex (RISC), the true effector of miRNA-mediated silencing; RISC consists of several proteins, including Ago. Once associated to RISC, mature miRNA becomes active and recognizes its mRNA target: in animals, interaction occurs between the 3'-UTR of mRNAs, where many miRNA-binding sites can be located, and the seed region of miRNAs, included between the second and the eighth nucleotide of the 5'-end of mature miRNAs. MiRNA-mRNA complementarity determines mRNA destiny: i) a perfect match causes the degradation of the mRNA, which is first deprived of its poly(A) tail; ii) a partial complementarity induces the inhibition of mRNA translation (115) (Figure 2.18).
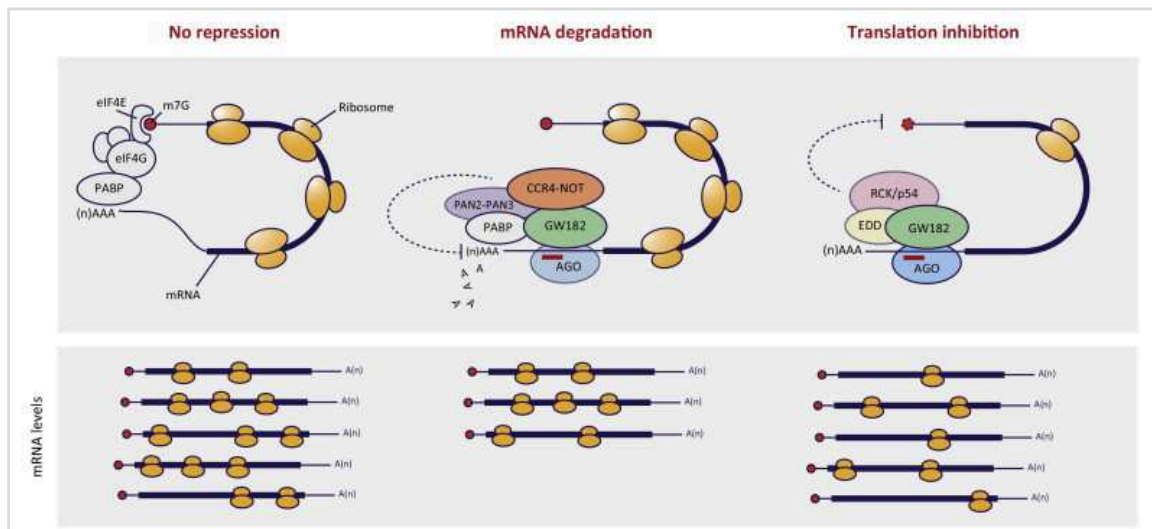
Figure 2.18: miRNA-mediated post-transcriptional silencing occurs through mRNA degradation or through inhibition of mRNA translation (from *Vidigal and Ventura, Trends Cell Biol. 2015*).

MiRNA-mediated mRNA degradation mechanism is still not clear; it seems that RISC recruits on mRNA target the effectors of deadenylation and degradation (exonucleases). This process starts in the cytoplasm and continues in specialized organelles called P-bodies (processing bodies), subcellular structures representing transcript storage and decay areas (109, 116). Translation inhibition seems to be due to RISC-included Ago proteins, which compete with eukaryotic translation initiation factors (eIFs) to bind mRNA 5'-cap: during the first stages of protein synthesis, eIF factors bind the 5'-cap, while PABPC1 [poly(A) binding protein cytoplasmic 1] binds poly(A) tail; eIFs-PABPC1 interaction induces the approach of the two ends of mRNA, promoting its recognition by ribosomes. When mRNA is recognized by miRNA loaded on RISC, Ago proteins bind its 5'-cap and prevent translation start (109).

It has been demonstrated that every miRNA can bind several mRNA targets, and, similarly, a single mRNA can be targeted by many miRNAs (99); therefore, miRNAs can be considered small effectors of wider regulatory pathways, controlling fundamental processes, such as cell cycle, differentiation, apoptosis, and also more complex processes regarding the entire organism, such as embryonic development, immune response, and many others.

2.4 Exosomes

In the last decade, the scientific community welcomed with great interest the discovery of a new mechanism of intercellular communication mediated by membranous vesicles, structures delimited by a phospholipid bilayer with transmembrane proteins; these vesicles carry the same molecules included within cell cytoplasm, such as hydrophilic solutes and biological macromolecules (proteins, lipids, RNAs), so they have been initially considered miniature cells. Further studies showed that cells secrete several types of vesicles, identified with different names (exosomes, exosome-like vesicles, microparticles, microvesicles, membrane-bound particles, apoptotic bodies, apoptotic microparticles), conferred in relation to vesicle dimension, density, origin and biogenesis. It is possible to classify vesicles in: i) exosomes, vesicles of endocytic origin secreted after multivesicular bodies and plasma membrane fusion; ii) shedding microvesicles, which directly origin from plasma membrane; iii) apoptotic bodies, derived from membrane blebbing during apoptosis (117, 118) (Figure 2.19).
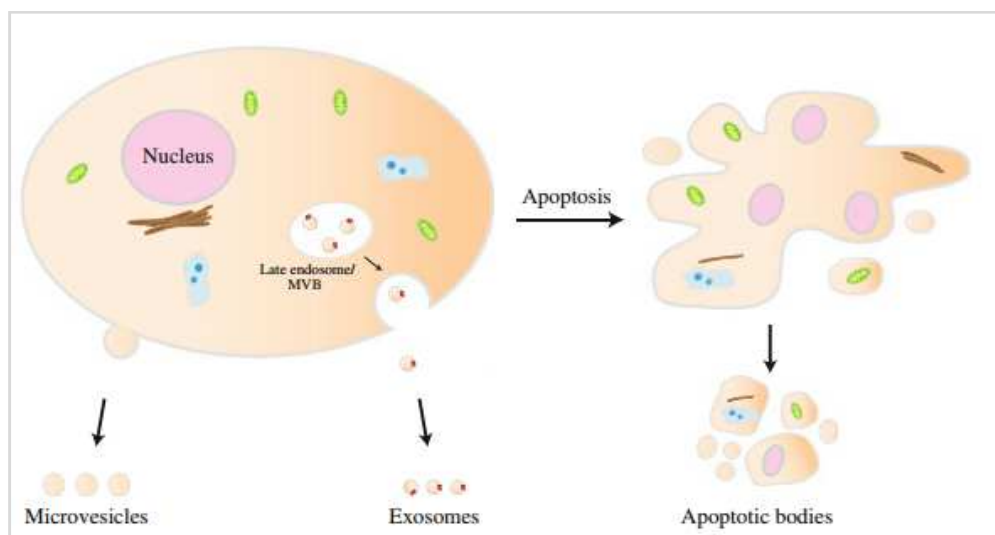
Figure 2.19: different biogenesis of shedding microvesicles, exosomes and apoptotic bodies (from *Akers et al., J Neurooncol. 2013*).

Exosomes are small membrane vesicles characteristically cup-shaped, with a diameter of 30-100 nm (119, 120) (Figure 2.20). Similarly to cells, exosomes present lipid rafts on their membrane and contain proteins and RNAs; typically, it is believed that exosomes lack DNA, even if some studies recently showed the contrary (121, 122).
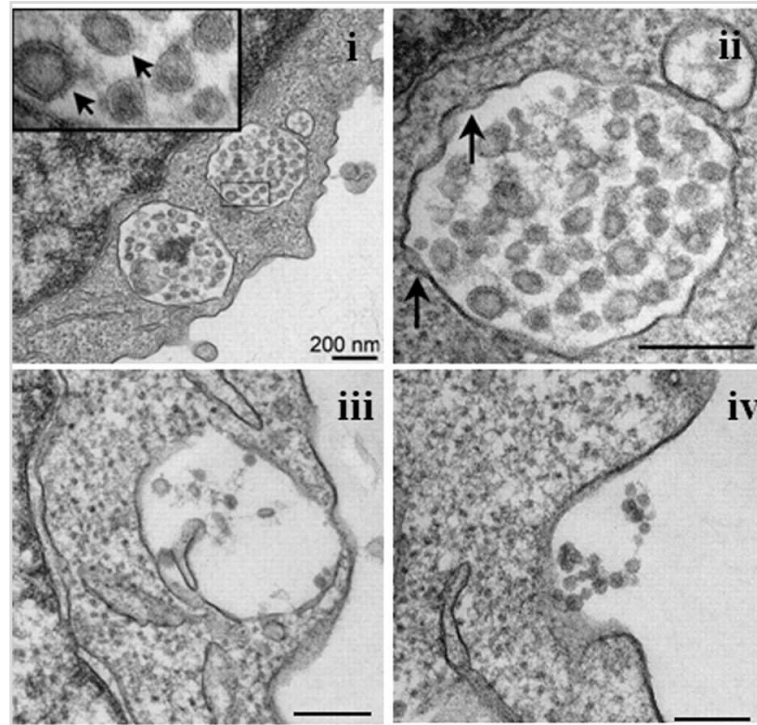
Figure 2.20: exosomes observed by transmission electron microscopy. Several stages of biogenesis are shown: i) intraluminal vesicles within multivesicular bodies; ii) intraluminal vesicles formation through inward invagination of multivesicular body membrane; iii) multivesicular bodies and plasma membrane fusion; iv) exosome secretion (from *Sahoo et al., Circ Res. 2011*).

Exosomes have been identified for the first time in 1987 as vesicles secreted during maturation of reticulocytes, red blood cells precursors (123); successively, it has been observed that similar structures are secreted by several cell histotypes, such as immune system cells (124, 125, 126), platelets (127), intestinal epithelial cells (128), cancer cells (129, 130). By now it is clear that many cell types, if not all, can secrete exosomes (131, 132, 133), which are released in extracellular space and biological fluids: indeed, exosomes have been identified in all body fluids analysed to

date, such as urine (134), seminal fluid (135), amniotic fluid (136), saliva (137), breast milk and plasma (138).

Exosome biogenesis begins on plasma membrane with an endocytosis event that generates a vesicle called early endosome, typically characterized by tube-like shape and cortical localization within the cytoplasm; early endosome undergoes a maturation process during which internal pH decreases because of the activation of membrane proton pumps, converting into a late endosome, spherical and closer to the nucleus. Late endosome evolves into a multivesicular body (MVB) when its membrane, through inward invaginations, originates intraluminal vesicles (ILVs). MVB fate is variable: indeed, MVBs can: i) fuse together with lysosomes, causing the degradation of their cargo by lysosome hydrolytic enzymes; ii) move toward plasma membrane, fuse with it and release ILVs in the extracellular space; secreted ILVs are called exosomes (139) (Figure 2.21)
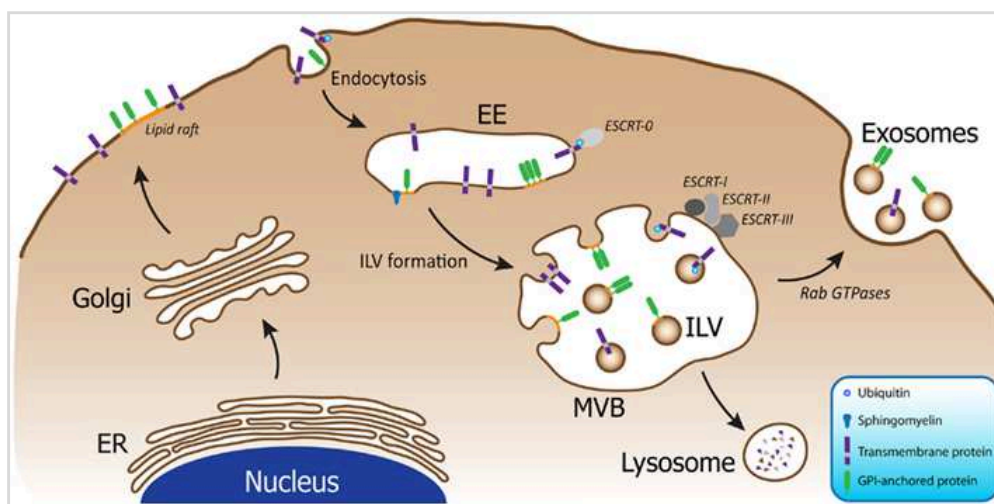
Figure 2.21: exosome biogenesis starts with an endocytosis event and the formation of an early endosome; after a maturation process, it becomes a late endosome, which converts into a multivesicular body after inward invaginations of its membrane, leading to the formation of intraluminal vesicles. Once secreted through MVBs and plasma membrane fusion, ILVs are called exosomes (from *Bellingham et al., Front. Physiol. 2012*).

Exosome biogenesis regulation is still under investigation. A calcium-mediated mechanism modulating vesicles secretion has been proposed; calcium ions control similar events, such as neurotransmitter release through synaptic vesicle secretion, or lysosome fusion with the plasma

membrane; in particular, exosome secretion is induced by increased cytoplasmic levels of calcium ions in haematopoietic cells and oligodendrocytes (140, 141, 142). Similarly, plasma membrane depolarization due to the increase of potassium ions promotes exosome release in cortical neurons (143). Moreover, exosome biogenesis is impaired by inhibition of sphingomyelinases, ubiquitous enzymes controlling ceramide production through hydrolysis of sphingomyelin molecules included in plasma membranes: this evidence suggests that ceramide, of which exosome membranes are enriched, can act as exosome biogenesis regulator (144). Rab protein family includes small GTPases involved in trafficking and fusion of intracellular membranes; many members of this family (in particular Rab27a and Rab27b) regulate exosome biogenesis, even if it's not clear if different members act in different cell types or at different stages of the process (145, 146, 147, 148, 149). Other proteins involved in exosome production are SNAREs (soluble N-ethylmaleimide-sensitive factor attachment protein receptors), a family of membrane fusion regulators; SNAREs modulate MVBs fusion with the plasma membrane. Vesicular SNAREs (v-SNAREs), included within MVB membrane, interact with target SNAREs (t-SNARE), located on the cytoplasmic side of the plasma membrane; this interaction causes the formation of a membrane-bridging SNARE complex, which controls the fusion of the membranes (150, 151) (Figure 2.22).
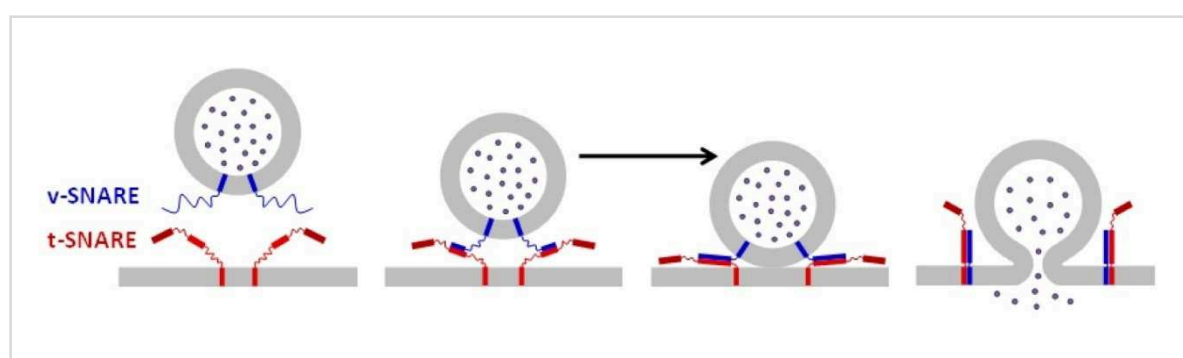
Figure 2.22: membrane fusion mediated by SNARE proteins: vesicular SNAREs interact with target SNAREs, moving the two membranes closer and triggering their fusion (from *David Tareste, Biophysics of membrane fusion*).

Also the exosome cargo is still under investigation; ExoCarta database (http://www.exocarta.org/) lists all molecules which have been indentified to date within exosomes

secreted by several cell histotypes. However, several studies underline a specific cargo including lipids, proteins and RNAs:

i. Exosomal membranes are particularly enriched in cholesterol, ceramide, sphingomyelin, gangliosides and polyunsaturated fatty acids, while phosphatidylcholine and diacylglycerol levels are lower compared to donor cell membranes (152, 153); moreover, compared to cell membranes, exosomal membranes contain more phosphatidylserine on outer side, which could promote their internalization by recipient cells (154). In its entirety, exosomal membrane is more rigid than cell membrane, and this rigidity could be pH-dependent. This hypothesis isn't surprising considering exosome biogenesis, since MVB pH is lower than cytosolic one; plus, it has been observed that exosome uptake increases in acid microenvironment, such as the cancer niche, where decreased pH makes exosome membrane more fluid, promoting its fusion with recipient cells (153, 155). In addition, exosome membrane rigidity confers on them greater resistance to degradation, making exosomes excellent vehicles for several different molecules (156, 157). Exosomes also contain lipids acting as signal mediators, such as prostaglandins, and enzymes essential for their production, among which phospholipases A and D (158). It has been demonstrated that, overall, exosomes present an 8.4-fold enrichment of lipids per mg of protein (152).

ii. Proteins identified within exosomes always come from donor cell cytoplasm, while organelle characteristic proteins have never been found within exosomes (159, 160); in particular, exosomes transport heat shock proteins, such as Hsc73 and Hsc90 chaperons, proteins constituting ESCRT (endosomal sorting complex required for transport) complex, such as Tsg101 (tumor susceptibility gene 101) and Alix; tetraspanins, among which CD63 and CD81; cell cycle regulators, such as

phosphatidylinositol-3-kinase (PI3K), mitogen-activated protein kinases (MAPKs) and ErbB family members; integrins; cytoskeleton proteins, such as actin, myosin and tubulins; metabolism enzymes, among which glyceraldehyde 3-phosphate dehydrogenase (GAPDH); proteins involved in signal transduction, such as beta-catenin (120, 161, 162, 163). Among all these proteins, tetraspanins are particularly abundant in exosomes, so they are considered exosome specific markers.

iii.   Exosomes are enriched in RNA molecules long less than 200 nucleotides, typically mRNAs and miRNAs, while they lack ribosomal RNAs (164); recently, lncRNAs and circRNAs have been identified within exosomes (165, 166, 167, 168).

It has been hypothesized that exosome cargo is casually determined during ILV formation, because late endosome membrane invaginations could lead to random incorporation of surrounding portions of the cytoplasm (139); recently, a second theory is emerging, according to which cells strictly control molecular trafficking toward exosomes. This theory is supported by several studies demonstrating that exosome cargo depends on donor cell histotype and physiological state (169, 170, 171, 172, 173). Nevertheless, it is not clear yet how cells recognize those molecules that have to be directed toward exosomes: protein sequence or structure motifs (e.g., acylation sites or phospholipid-binding domains) and membrane myristoylation and palmitoylation sites which may mediate protein recognition have been identified (174); also RNA molecules seem to be identified through specific sequence motifs discriminating cellular and exosomal miRNAs; these consensus sequences seem to be recognized by ribonucleoproteins, such as hnRNPA2B1 (heterogeneous nuclear ribonucleoprotein A2B1) (173). ESCRT machinery involvement in exosomal cargo sorting seems to be certain. ESCRT consists of four multiprotein complexes, called ESCRT-0, ESCRT-I, ESCRT-II and ESCRT-III, usually recruited on the cytoplasmic side of endosome membrane to sort specific proteins within ILVs. This process requires ubiquitination of target proteins, recognized by Tsg101, included in ESCRT-I; the interaction between Tsg101 and target proteins activates

ESCRT-II: ESCRT-I and ESCRT-II together regulate the inward invaginations of MVB membrane, allowing cytoplasmic mRNAs, miRNAs and proteins to enter into the nascent ILVs. ESCRT-III is recruited by ESCRT-II to cleave the nascent ILVs into free vesicles within the MVB. All ESCRT subunits and free ubiquitin molecules are released within the cytoplasm to be recycled (Figure 2.23). However, ESCRT-independent mechanisms regulating exosomal cargo sorting have been proposed (175).
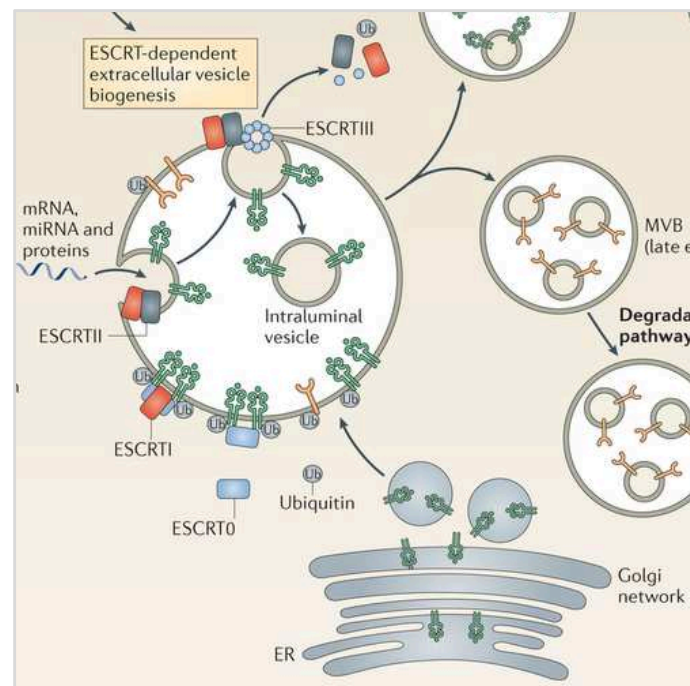
Figure 2.23: ESCRT complex involvement in exosomal cargo sorting: the ESCRT-0 complex recognizes ubiquitinated target proteins in the cytoplasm and binds the ESCRT-I complex, which in turn recruits ESCRT-II subunits. ESCRT-I and ESCRT-II regulates the inward invaginations of MVB membrane, allowing cytoplasmic mRNAs, miRNAs and proteins to enter into the nascent ILVs. ESCRT-II recruits ESCRT-III subunits, which cleave the nascent ILVs into free vesicles. All ESCRT subunits and ubiquitin molecules are released in the cytoplasm to be recycled (from *Robbins and Morelli, Nat Rev Immunol. 2014*).

2.4.1 Exosomes as mediators of intercellular communication

Despite exosomes were initially considered cell "waste bins" secreted to remove dangerous or useless compounds, it is by now clear that their main function is to mediate intercellular communication. Once secreted, exosomes interact with recipient cells influencing their physiology.

Exosomes can act on the same donor cell in autocrine manner, on recipient cells near to donor cell in paracrine manner, or they can reach distant target cells through bloodstream and act in endocrine manner (176, 177, 178). The exact molecular mechanism of this interaction is still under investigation, even if four models have been proposed: i) a protein of exosome membrane interact with a surface receptor of the recipient cell, acting as a ligand (Figure 2.24b); ii) exosome membrane fuses with target cell membrane, transferring in its cytoplasm their entire cargo (Figure 2.24c); iii) a protein of exosome membrane undergoes a proteolytic cleavage in the extracellular space, generating a soluble ligand that interacts with its receptor on the surface of the target cell (Figure 2.24d); iv) a recipient cell with phagocytic activity absorbs the entire vesicle through phagocytosis (Figure 2.24e) (155, 179, 180, 181).
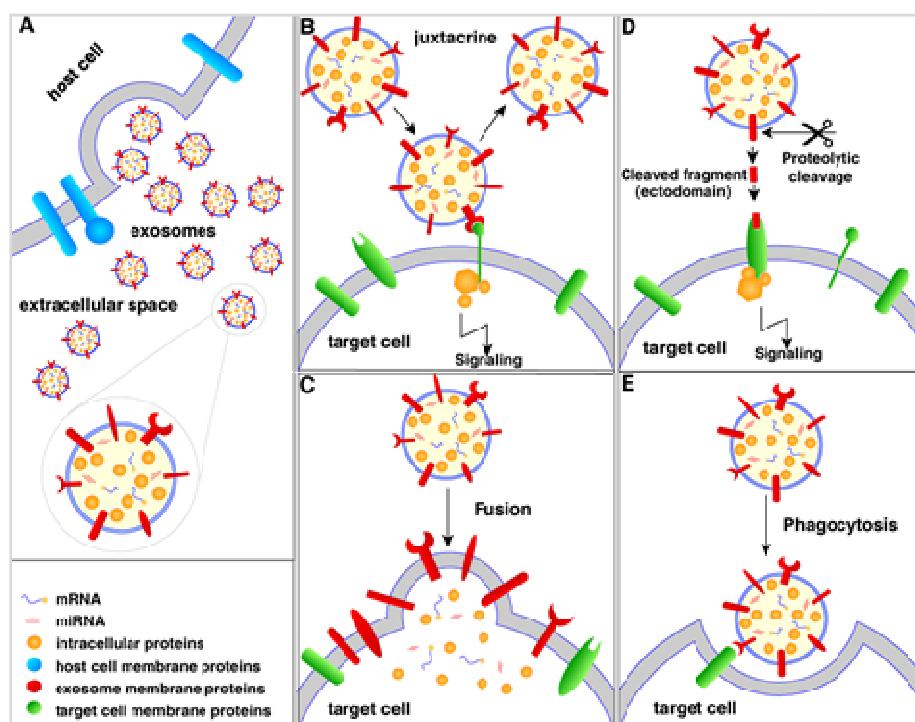
Figure 2.24: proposed models explaining the molecular mechanism of interactions between exosomes and recipient cells. A) exosomes are released by donor cells; B) a protein of exosome membrane interacts with a surface receptor on recipient cell membrane; C) exosome membrane fuses with target cell membrane; D) through a proteolytic cleavage in the extracellular space, a protein of exosome membrane generates a soluble ligand, which binds its receptor on recipient cell surface; E) a phagocytic cell absorbs the entire exosome through phagocytosis (from *Mathivanan, J Biotechnol Biomater. 2012*).

Whatever the mechanism, the interaction between exosomes and recipient cells induces a response modulating cell phenotype, because of a signal transduction event or the entrance of specific molecules carried by vesicles. Once entered into target cells, exosome mRNAs are actively translated into functional proteins (164); in particular, exosomes secreted by cells cultured in oxidative stress conditions, induced by hydrogen peroxide administration, confer on recipient cells higher resistance to the same treatment through RNA molecules transfer (169).

Among the first exosome-secreting cells identified there are several immune system cells; indeed, exosomes contribute to immune response modulation both mediating intercellular communication and acting as antigen-presenting vesicles thanks to their enrichment in class II MHC (major histocompatibility complex) molecules (124). Particularly, MHC-antigen complexes on exosome surface can directly interact with T-cell receptor; otherwise, the whole microvesicle can be captured by antigen-presenting cells (APCs) and indirectly activate T lymphocytes (182, 183, 184, 185); in addition, exosomes secreted by mature dendritic cells are more efficient in the activation of T lymphocytes than exosomes secreted by immature dendritic cells, suggesting that donor cell type influences exosome function (184, 186, 187).

As suggested by the presence of several molecules involved in signal transduction pathways within exosomes, these microvesicles mediate the transfer of pro-proliferative *stimuli*. First proteomic studies showed that beta-catenin, the main cytosolic regulator of Wnt pathway, is contained into exosomes; in addition, beta-catenin packaging into exosomes increases in presence of CD9 and CD82, two tetraspanins inhibiting Wnt pathway, and cytoplasmic protein levels decrease through its expulsion via exosome secretion (188).

Recently, the role played by exosomes in nervous system physiology has attracted increasing attention, since exosome secretion has been observed in neurons, Schwann cells, oligodendrocytes, astrocytes and microglia (133, 142, 143, 189). Exosomes produced by oligodendrocytes transfer enzymes, such as catalase and superoxide dismutase I, to neurons, helping

them fighting oxidative stress (190). Moreover, exosomes regulate synaptic plasticity by transporting AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) glutamatergic receptor, thus modulating their number on post-synaptic membrane (189).

2.5 Colorectal cancer

Colorectal cancer (CRC) is one of the most common diseases in industrial society. About 875,000 new cases are reported every year, among which about 500,000 are fatal. In Western society, CRC represents the second leading cause of cancer death and the third most frequent cancer type (191).

CRC has the highest incidence rate among tumors in the Italian population: among men it represents 14% of all new tumor cases, being preceded only by prostate and lung cancer, while among women it represents 13% of new cases, being preceded only by breast cancer. The causes of CRC increased incidence observed in the last decades are population aging and higher spread of risk factors, such as the consumption of red meat, sausages, refined flours and sugars, obesity and reduced physical exercise, smoking and alcohol excess; however, prevention campaigns led in recent years to reduced incidence and mortality (192). Average onset age is 62 years, with an increased risk after 40 for both men and women (191). Prognosis is generally favorable and tends to get better, being 5-years survival rate increased from 50% in early 90s to 64% in men in 2006-2007, and from 51% to 63% in women (192).

The disease begins as a benign adenomatous polyp, which successively evolves into advanced adenoma with high-grade dysplasia and finally in invasive cancer. TNM (tumor-node-metastasis) classification allows to classify tumor stage according on three parameters: i) tumor extension (T), ii) lymph node involvement (N), iii) distant metastasis (M). Invasive cancers confined within colon wall (TNM stage I and II) are curable, but if not treated they spread to regional lymph nodes (TNM stage III) and finally cause distant metastasis formation (TNM stage

IV). Stage I and II cancers are curable through surgical excision, and up to 73% of stage III tumors are curable by surgical excision associated with chemotherapy. Even if recent progresses in chemotherapy led to increased survival, stage IV cancers are usually incurable (191, 193).

Clearly, CRC prognosis depends largely on pathological stage at diagnosis and is influenced by several factors, among which the most important is tumor stage; bowel wall invasion degree, lymph node and distant metastases negatively influence prognosis (191).

2.5.1 CRC genetic alterations

CRC can be classified into two different forms: i) sporadic CRC is more common, representing about 75-80% of all cases; ii) familial (hereditary) CRC is rarer, even if it has been observed that a high percentage of CRC patients has relatives affected, with a familial degree comprised between the first and the third; this finding suggests that even in sporadic cancers familiarity leads to a higher risk of disease onset. According to genetic alterations involved in carcinogenesis, sporadic CRC can be further classified into two subtypes, associated with two main pathways involved in CRC: i) the "suppressor" or "canonical" pathway, which causes chromosomal instability (CIN); ii) the "mutator" pathway, associated with microsatellite instability (MSI) (194).

It is believed that the canonical pathway (about 80-85% of sporadic CRC cases) follows the sequence of molecular events described in Fearon and Vogelstein model (195), according to which specific genetic events correlate with tissue morphology during CRC carcinogenesis (Figure 2.25): the first genetic alteration is represented by tumor suppressor APC gene mutation during the adenoma stage, followed by KRAS and p53 mutations and chromosome 18q deletions. This model, although considered linear and simplistic, shows that mutations of tumor suppressor genes, such as APC and p53, and of oncogenes, such as KRAS, are fundamental events in the canonical pathway. Cancers characterized by CIN also present allelic imbalance, with frequent allelic loss on chromosomes 5q, 8p, 17p and 18q, and chromosomal duplications and translocations (194).
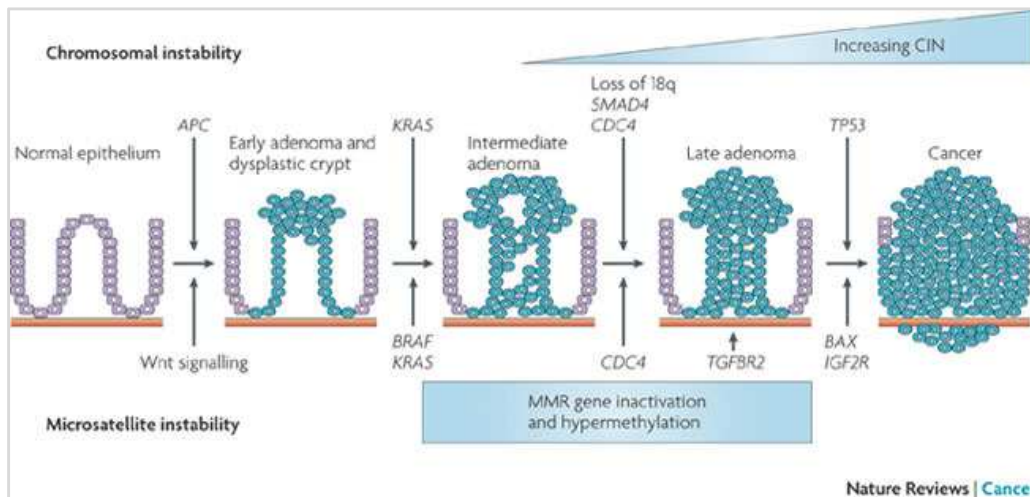
Figure 2.25: molecular alterations during CRC carcinogenesis; the upper part shows mutations associated with the "suppressor" pathway, the bottom those associated with the "mutator" pathway (from *Walther et al., Nat Rev Cancer. 2009*).

On the contrary, the "mutator" pathway is rarer, representing 15-20% of sporadic CRC cases. MSI cancers are characterized by high microsatellite instability (MSI-H), that is the accumulation of mutations in microsatellites, short non-coding DNA sequences repeated in tandem throughout the genome. This accumulation of mutations is due to other mutations occurring in genes coding for the mismatch repair (MMR) system, which normally corrects mistakes made by DNA polymerase during DNA replication; mutation-induced MMR system inactivation causes a further accumulation of mutations in microsatellite and in the entire genome. MSI-H sporadic cancers generally don't show cytogenetic abnormalities or aneuploidy; plus, they present rare or no mutations typical of "suppressor" pathway, while other mutations have been described in microsatellite sequences of genes involved in carcinogenesis, such as TGFβRII, BAX, MSH3, MSH6, caspase 5, APC, beta-catenin. In addition, most of MSI-H sporadic cancers presents the CpG island methylator phenotype (CIMP), characterized by DNA aberrant methylation which causes silencing of functional genes (194).

At molecular level, it is by now known that some signal transduction pathways play a crucial role in CRC neoplastic transformation: among them there are Wnt, MAPK, PI3K and p53 pathways.

The first event in CRC pathogenesis is represented by the activation of Wnt signaling pathway, which regulates proliferation, differentiation and motility of intestinal epithelial cells. In physiological conditions, the APC (adenomatous polyposis coli) protein interacts with Axin in the cytoplasm; the APC-Axin complex recruits the serine/threonine protein kinase GSK3β (glicogen synthase kinase 3β) and the kinase CK-1 (casein kinase 1), forming the beta-catenin "destruction complex"; this complex binds and phosphorylates cytoplasmic beta-catenin, earmarking it for the ubiquitin-proteasome-mediated degradation (Figure 2.26). Beta-catenin plays a dual role within cells: it acts as transcriptional factor and as cellular adhesion stabilizer. In the cytoplasm, beta-catenin binds E-cadherin, a transmembrane glycoprotein involved in formation and maintenance of adherens junctions between epithelial cells. In late carcinogenesis, E-cadherin loses its function; consequently, adherens junctions are disrupted and cell motility and migration increase. Normally, E-cadherin and APC compete for beta-catenin binding, controlling its cytoplasmic levels by inducing its degradation; thus, beta-catenin doesn't accumulate within the cytoplasm and is not imported to the nucleus. On the contrary, when APC is mutated and non-functional, E-cadherin binds beta-catenin until saturation, inducing a cytoplasmic accumulation of beta-catenin; therefore, it is imported to the nucleus to perform its function of transcriptional co-factor, inducing the expression of genes regulating proliferation, differentiation and cell migration, such as c-Myc, MMP-7 (matrix metalloproteinase-7), and cyclin D1. Since APC is a key regulator of Wnt pathway, it is believed that its mutation with loss of function is the initial event in CRC pathogenesis: mutated APC has been observed in about 80% of CRC patients, while beta-catenin mutations are rarer, being reported in about 5% of cases. APC mutations can be due to hypermethylation or

deregulation of APC gene; also, they can be inherited, as in FAP (familial adenomatous polyposis), an inherited syndrome causing predisposition to cancer (196).
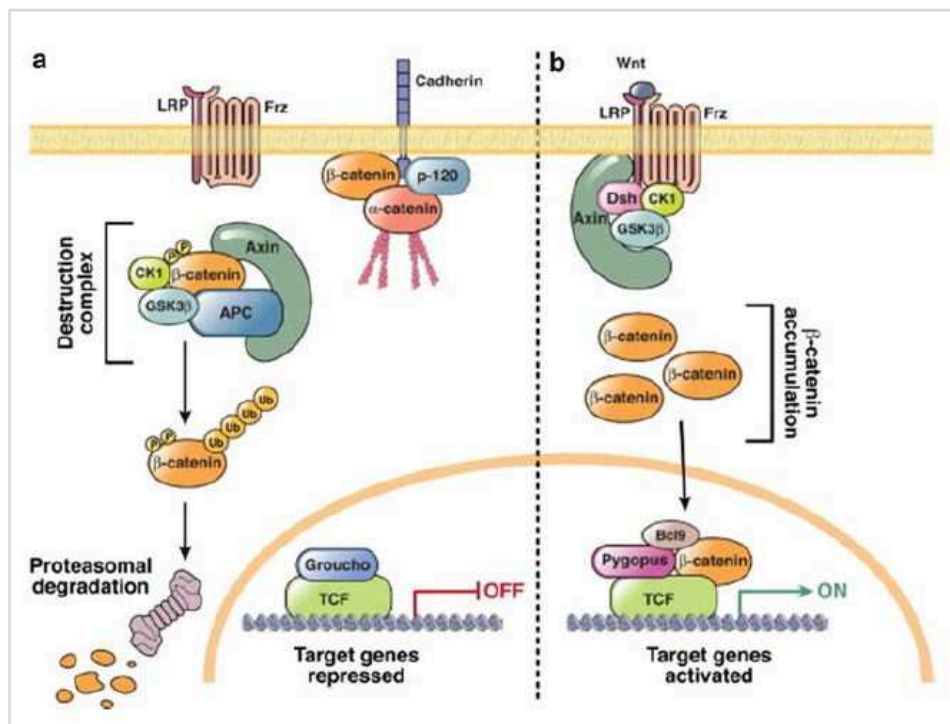
Figure 2.26: schematic representation of Wnt signaling pathway (from *Roper and Hung, Molecular Pathogenesis of Colorectal Cancer, Chapter 2. Springer New York, 2013*).

Another pathway classically associated with CRC is MAPK pathway, involved in both sporadic CRC subtypes: KRAS (KRAS proto-oncogene, GTPase) mutations have been observed in 30-35% of patients and are typically associated with chromosomal instability, while BRAF (B-Raf proto-oncogene, serine/threonine kinase) mutations are less frequent (9-11%) and are often associated with microsatellite instability (197). KRAS is a small monomeric G-protein involved in the transduction of mitogenic signals activated by growth factor receptors located on cell membranes. KRAS activation induces a sequence of downstream phosphorylations performed by several kinases, such as MAPK1/ERK (mitogen-activated protein kinase 1) and PI3K/AKT. In physiological conditions, the growth factor binds its receptor tyrosine kinase (RTK), causing its dimerization and autophosphorylation; the phosphorylated dimeric RTK, thus activated, recruits in the cortical cytoplasm GRB2/SOS (growth factor receptor/bound protein2/son of sevenless), which

in turn interacts with KRAS, activating it. Through GDP-GTP exchange, KRAS changes conformation and binds effector molecules, inducing the downstream phosphorylation sequence ending with ERK phosphorylation and translocation to the nucleus; here, p-ERK phosphorylates and activates transcription factors such as c-FOS and c-JUN, inducing the expression of genes regulating proliferation, differentiation, apoptosis (Figure 2.27). After GTP hydrolysis, KRAS is inactivated. All the different KRAS point mutations to date identified cause the loss of hydrolytic function, resulting in a constitutively active protein which induces an uncontrolled cell proliferation. KRAS mutations also alter Wnt pathway: KRAS activation induces an increased expression of beta-catenin, which accumulates in the cytoplasm; KRAS also phosphorylates and activates PI3K, causing GSK3β phosphorylation and inducing beta-catenin nuclear translocation. Hence, in physiological conditions, GSK3β acts as negative regulator of two fundamental pathways involved in CRC pathogenesis (196).
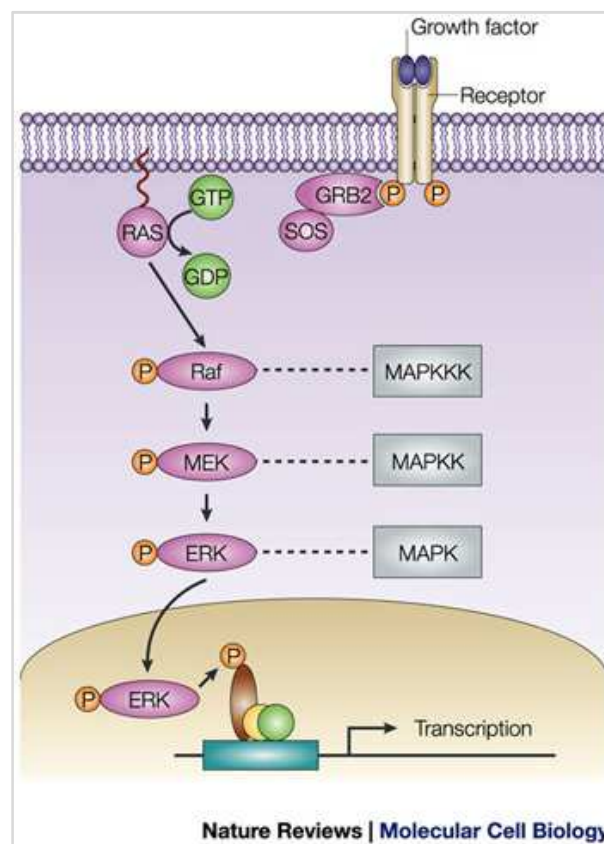
Figure 2.27: schematic representation of KRAS and downstream MAPK signaling pathway (from *Kim and Bar-Sagi, Nat Rev Mol Cell Biol. 2004*).

PI3K/AKT pathway also plays a crucial role in CRC carcinogenesis, inducing malignant transformation, cell proliferation, migration and survival, and promoting tumor progression. PI3K (phosphoinositide 3-kinase), belonging to a lipid kinase family, is made up of a regulatory subunit (p85) and a catalytic subunit (p110); there are three variants of p85 subunit (p85α, p85β and p85γ, coded respectively by PIK3R1, PIK3R2 and PIK3R3 genes) and three variants of p110 subunit (p110α, p110β and p110γ coded respectively by PIK3CA, PIK3CB and PIK3CD genes). Mutations inducing aberrant pathway activation have been observed in 30-35% of CRC patients, most frequently in PIK3CA and PIK3CB genes (196). Normally, PI3K is activated by external *stimuli* causing RTK activation, among which also KRAS-induced *stimuli*; p85 subunit recognizes phosphorylated tyrosine residues in the receptor intracellular domain, binding them and releasing p110 subunit, thus activated; p110 subunit phosphorylates phosphatidylinositol 4,5 biphosphate (PIP2) to phosphatidylinositol 3,4,5 triphosphate (PIP3), inducing PIP3 accumulation at cell membrane. PIP3 recruits AKT (AKT serine/threonine kinase 1), which is phosphorylated by PDK1 (pyruvate dehydrogenase kinase 1) and mTORC2 complex, including among others MTOR (mechanistic target of rapamycin). AKT promotes growth and survival in several ways: i) inhibition of pro-apoptotic proteins belonging to Bcl-2 family; ii) increased cytoplasmic availability of MDM2 (MDM2 proto-oncogene) and subsequent increase of p53 degradation; iii) induced expression of anti-apoptotic genes through NF-kB (nuclear factor kappa B) activation. PI3K pathway is inhibited by PTEN (phosphatase and tensin homologue), a tumor suppressor phosphatase which dephosphorylates PIP3 to PIP2, inhibiting AKT activation; therefore, PTEN is associated with inhibition of cell cycle progression and promotion of cell death and angiogenesis. CRC cases with PTEN inactivation, due to truncating mutations or promoter methylation-induced silencing, have been reported (198).

In late CRC carcinogenesis, p53 mutation has been reported in 50% of patients; p53 is a tumor suppressor protein which acts as transcription factor, thus stopping cell cycle and activating apoptosis in presence of DNA damage; for this reason, p53 has been defined the "genome guardian". In physiological conditions, low p53 protein levels are maintained through ubiquitin-proteasome degradation: MDM2, an ubiquitin-ligase enzyme, binds p53 and causes its degradation. Interestingly, MDM2 itself is a target of p53, which activates its expression; this mechanism represents a negative feedback regulation of p53 expression and function. After DNA damage, p53 phosphorylation within the MDM2-binding domain impedes p53-MDM2 interaction; thus, p53 is more stable, translocates to the nucleus and binds DNA, activating the transcription of several genes, including p21 (cyclin-dependent kinase inhibitor 1A, CDKN1A) and the pro-apoptotic factor BAX (BCL2-associated X protein); in this way, p53 arrests cell cycle in G1 phase, allowing DNA repair, or induces apoptosis if DNA damage cannot be repaired. About 80% of p53 mutations are missense and cause a stable but non-functional protein, which accumulates in the cytoplasm; this cytoplasmic accumulation of p53 is strongly correlated to metastatization and poor prognosis (196, 197).

Finally, epigenetic alterations also participate to tumor progression and have been associated with CRC; they include DNA methylation, histone modifications and chromatin remodeling, reversible alterations modulating gene expression without modifications in DNA sequence. DNA methylation is mediated by DNA methyltransferases; in eukaryotic genomes, it affects prevalently cytosines of CpG islands (CpG dinucleotides enriched regions), located within the promoter of the 60% of human genes; typically, methylation of promoter CpG islands provokes silencing, while methylation of CpG islands external to promoter fosters transcriptional activation. Carcinogenesis induces remarkable modification in the genome methylation pattern, with demethylation and hypermethylation of many promoters: among hypermethylated genes in CRC there are genes involved in Wnt pathway (APC and AXIN2), DNA repair (MLH1 and MLH2), cell cycle regulation (CDKN2A), and KRAS pathway (RASSF1A and RASSF1B). Histone modifications consist of

addition or removal of functional groups, such as acetyl groups (acetylation or deacetylation) or methyl groups (methylation or demethylation), on specific amino acid residues of N-terminal tails of histones, proteins around which DNA is wrapped within the nucleus. Specific histone modifications are associated with silencing or transcription activation: deacetylation, catalyzed by histone deacetylases (HDACs), induces transcription repression, while acetylation, catalyzed by histone acetyltransferases (HATs), promotes gene expression. The upregulation of many HDAC, including HDAC1-3, HDAC5 and HDAC7, has been reported in CRC, together with the downregulation of genes involved in Wnt pathway. Moreover, HDAC2 upregulation observed in early carcinogesis is associated with H4K12 and H3K18 histone ipoacetylation during the adenoma-carcinoma sequence: this evidence suggests that increased HDAC2 expression and subsequent ipoacetylation participate in CRC progression (199).

2.5.2 CRC diagnosis and therapy

It has been established that favorable prognosis of cancer patients depends on many factor, including early diagnosis; for this reason, several screening projects are implemented every year, aiming to diagnose early stage tumors, before symptom onset. Screening tests are usually minimally invasive and easy to perform; only patients positive to these tests undergo more deepened and often more invasive analyses.

Many epidemiological studies showed that environmental factors, such as diet and lifestyle, are associated with CRC pathogenesis: red meat high consumption, reduced physical exercise, obesity, alcohol and smoke increase the risk of tumor. It is believed that in most cases colorectal cancer begins as adenoma, a benign lesion that can evolve into a tumor; this process can require a long time, even ten years, making CRC a good candidate for screening and prevention. In addition, adenoma resection can prevent tumor development, reducing incidence and mortality. To date, two

principal screening strategies are used: faecal occult blood testing (FOBT) and endoscopy (flexible sigmoidoscopy and colonoscopy).

Tumors bleed, pouring into the intestinal lumen small amounts of blood that can be detected in stool. Therefore, FOBT represents a non-invasive and cheap test, easy to perform at home by the patient himself. This test has to be repeated every one or two years to increase its sensitivity for cancer. However, this test is characterized by low sensitivity and specificity, giving false positive results when bleeding is caused by other pathologies, such as ulcers or inflammatory diseases of the intestinal tract.

Flexible sigmoidoscopy (FS) allows to inspect the mucosa, perform tissue biopsies and remove polyps in the distal part of the colon, but not in the region proximal to sigmoid colon because of the endoscope limited length. FS is a high sensitivity and specificity test both for adenomas and carcinomas within the inspected area.

On the contrary, colonoscopy allows to inspect the entire colon mucosa, perform tissue biopsies and remove polyps in the whole colon-rectum in one single session; therefore, it is the ideal test for CRC prevention and diagnosis. Colonoscopy sensitivity and specificity to identify advanced adenomas and tumors are about 100%, and it is indeed the last test performed when all other screening tests gave positive results. For these reasons, colonoscopy is considered the gold standard for colonic disease diagnosis, but a great variability depending on endoscopists has been observed. Plus, it is an invasive test, requiring a long time, expensive and painful or the patient. Serious complications, such as bleeding or perforation, even if rare, occur with higher frequency compared to flexible sigmoidoscopy (200, 201).

Thus, diagnosis tests used today are invasive or have low sensitivity. The identification of early genetic and epigenetic markers (in stool and blood) specific for adenomas or invasive tumors is still under investigation. It is known that mucosal cells physiologically shed from gastrointestinal epithelium, allowing to detect tumor DNA and epigenetic modifications in stool samples, which can

be used for diagnostic and prognostic analyses. Similarly, markers can be detected in blood: to date, CRC-associated markers include KRAS, APC, p53 and methylation markers, such as vimentin and septin-9; however, these markers are currently used only in clinical trials, not in CRC screening (201).

The first and most important therapeutic approach for colorectal cancer is represented by surgical resection of the affected colon tract; relapse probability is correlated with several factors, including depth of tumor penetration through the bowel wall and lymph node involvement. Surgical resection efficacy is high for stage I and II tumors (90% and 75% respectively); stage III tumors with lymph node involvement have 60% of recurrence (202).

Many therapeutic strategies are today used in CRC treatment; some of them aim to impair DNA replication and thus cell proliferation. A drug commonly administered in CRC therapy is 5-fluorouracil, a fluorinated pyrimidine which inhibits thymidylate synthase, the rate-limiting enzyme of pyrimidine nucleotide synthesis. Fluorouracil is usually administered with leucovorin (or folinic acid), a reduced folate stabilising the drug-enzyme interaction and increasing the inhibition of DNA replication: simultaneous administration of both molecules induces 50% tumor size reduction in 20% of patients and increases survival from 6 to 11 months. Side effects of fluorouracil depends on administration method (intravenous or bolus); they include neutropenia (reduced number of neutrophils), stomatitis and diarrhea. Irinotecan is a semisynthetic derivative of natural alkaloid camptothecin, extracted from the bark of *Camptotheca acuminata*; this drug exerts cytotoxic effect by interacting with topoisomerase I, the enzyme that uncoils DNA through single-stranded breaks during replication. These DNA breaks are normally repaired, but camptothecin stabilises them causing DNA fragmentation and cell death. Irinotecan is administered as a prodrug, which is successively hydrolysed to its active metabolite SN-38 by hepatic carboxylesterases. Side effects include diarrhea, myelotoxicity (bone marrow suppression), nausea, vomiting and alopecia. Oxaliplatin is a third-generation derivative of platinum which forms big DNA adducts, thus

inducing apoptosis. Oxaliplatin and fluorouracil act synergistically: since the former causes thymidylate synthase downregulation, it increases fluorouracil efficacy. Oxaliplatin toxicity provokes renal dysfunctions, alopecia, ototoxic effects and neuropathy (203).

Another common therapeutic approach is based on molecules specifically targeting the epidermal growth factor receptor (EGFR), which acts upstream of several signal transduction pathways associated with CRC, such as Wnt, MAPK, and PI3K. EFGR is a 170-kDa transmembrane glycoprotein characterized by an extracellular domain, containing the ligand-binding site and the sequences involved in dimerization, and an intracellular domain with tyrosine kinase activity. EGFR upregulation has been observed in many cancer types, including CRC; therefore, therapies based on inhibitors of the intracellular domain tyrosine kinase activity, or on monoclonal antibodies targeting the EGFR extracellular domain have been created. The higher the expression of EGFR, the better the response to the anti-EGFR therapy. Tyrosine kinase inhibitors include gefitinib and erlotinib, which reversibly inhibits EGFR autophosphorylation; among the best known monoclonal antibodies there are cetuximab, an IgG1 recombinant human/mouse chimeric antibody, and panitumumab, an IgG2 fully human antibody; cetuximab and panitumumab are the only monoclonal antibodies used in clinical practise. Both in normal and tumor cells, cetuximab binds the extracellular domain of EGFR, causing the internalization of the receptor and thus impeding receptor-ligand interaction; panitumumab also interacts with the EGFR extracellular domain in both cell types, thus inhibiting ligand binding, receptor dimerization, autophosphorylation and signal transduction (204). Monoclonal antibody side effects are mild, including acne-like rash and skin drying and fissuring among the most common (203). The anti-EGFR therapy is useless in patients bearing KRAS mutations: in fact, mutated KRAS is constitutively active, triggering MAPK pathway downstream of EGFR even in absence of EGFR ligands; therefore, KRAS mutations confer resistance to anti-EGFR therapies upon tumor cells

(205). For this reason, it is necessary to perform KRAS genotyping on surgically removed tumors before administering these drugs to the patient.

To promote its progression, tumor induces angiogenesis, that is the formation of new blood vessels in order to guarantee itself a constant supply of nutrients; given the importance of this phenomenon, several angiogenesis inhibitors are under investigation for therapeutic applications. One of the most important angiogenic factors is VEGF (vascular endothelial growth factor), associated with cellular proliferation and blood vessel maturation; it has been observed that VEGF expression correlates not only with angiogenesis, but also with tumor progression. Several studies demonstrated that anti-VEGF monoclonal antibodies associated with chemotherapy increase patient survival (206). Bevacizumab is an anti-VEGF monoclonal antibody; its action combined with chemotherapeutic agents has been evaluated in several clinical trials in patients affected by advanced CRC: results showed a considerable improvement in patient response to therapy and an increase in overall survival of 4-7 months; drug administration has been quite well tolerated and caused mild toxic effects, including reversible hypertension and proteinuria (203).

The possible use of vitamin A and its retinoid derivatives in therapeutic approaches is attracting growing interest; currently, retinoid are used in the treatment of epithelial tumors because they regulate several pathways involved in carcinogenesis, including those typically associated with CRC. Retinoic acid receptors (RARs), located within the nucleus, bind all-trans retinoid acid (ATRA), forming a complex which can bind DNA and promote transcription of genes regulating cell differentiation. Retinoid and vitamin A administration, even through nutrition, increases ATRA levels, reducing proliferation and increasing apoptosis; moreover, retinoids induce decreased levels of cyclin D1 and matrix metalloproteases (MMPs), enzymes involved in invasion, migration and metastasis (196).

2.6 NcRNAs as mediators of tumor progression

Since ncRNA discovery, a lot of scientific papers have been published about their functions; it was immediately clear that ncRNAs regulate a wide range of fundamental biological processes, such as cell proliferation, differentiation, migration, angiogenesis and apoptosis; therefore, ncRNAs play a crucial role in cell physiology. Since their fundamental role, the association between ncRNA deregulation and several diseases, including tumors, is not surprising: it has been shown that ncRNAs act as oncogenes or tumor suppressor genes, contributing to tumor onset and progression.

2.6.1 LncRNA role in CRC pathogenesis

Despite their recent discovery, in the last decade many studies proved lncRNA involvement in carcinogenesis, and in particular in CRC, through several molecular mechanisms: lncRNAs modulate numerous signal transduction pathways classically associated with CRC; moreover, they interact with other molecules, such as DNA, RNA and proteins. It has been demonstrated that expression levels of several lncRNAs are altered in CRC; however, to date the molecular mechanism of only a few lncRNAs has been understood.

One of most studied cancer-associated lncRNAs is MALAT1, upregulated in metastases of many tumors, including non-small cell lung cancer, hepatocellular carcinoma, and endometrial stromal sarcoma (207, 208, 209). Several MALAT1 functions have been hypothesized, including the regulation of pre-mRNA processing because of the association between the lncRNA and SC35 splicing domains within the nucleus (210); MALAT1 accumulates in the nucleolus, where it acts as "riboregulator" modulating the expression of target genes, including numerous tumor suppressors (72, 211). In addition, MALAT1 controls the activity of E2F1 (E2F transcription factor 1), a key transcription factor regulating cell cycle progression: MALAT1 interacts with non-methylated Polycomb 2 protein (Pc2), thus activating interchromatin granule and promoting E2F1 sumoylation and activation (212). *In vitro* silencing experiments in bladder cancer cells demonstrated the

involvement of MALAT1 in cell migration through EMT regulation (213); in addition, MALAT1 depletion causes aberrant mitosis, during which numerous cells accumulate at G2/M boundary, and cell death (214). Finally, the tRNA-like 61-nucleotide fragment produced by MALAT1 3'-end, called mascRNA, induces proliferation and invasion in CRC cells (215).

The lncRNA BANCR (BRAF-activated non-protein coding RNA) seems to be associated with the BRAF V600E mutation, one of the most frequent mutations observed in several tumors, including CRC (216). BANCR upregulation, frequently observed in CRC tissues, correlates with lymph node metastasis and tumor stage; moreover, BANCR modulates migration and EMT in CRC cell lines (217). LncRNA-ATB (long non-coding RNA-activated by TGF-β) is also upregulated in CRC and promotes tumor progression by inducing EMT: in particular, lncRNA-ATB high levels cause reduce the expression of E-cadherin, while increasing ZEB1 (zinc finger E-box binding homeobox 1), a mesenchymal cell marker (45).

Through genome wide association studies (GWASs) a region of approximately 2 Mb on human chromosome 8q24 has been associated with increased susceptibility to several diseases, including CRC (218); this region is located within a large protein-coding gene desert, but it includes many ncRNA genes, such as CCAT1, CCAT2, PCAT2, PRNCR1 (219). In particular, the rs6983267 SNP (single nucleotide polymorphism) has been associated with an increased risk of CRC (220); this SNP maps to a region containing DNA enhancer elements and affects the binding affinity of TCF7L2 (transcription factor 7 like 2), a transcription factor regulating the expression of Wnt pathway target genes (221). The rs6983267 SNP lies within the gene coding for CCAT2 (colon cancer associated transcript 2), affecting its expression: the allele G induces a higher CCAT2 expression, thus increasing CRC risk; CCAT2 itself interacts with TCF7L2, inducing the upregulation of MYC, a proto-oncogene involved in CRC, and overactivating Wnt pathway; also, CCAT2 is itself a Wnt downstream target, in a positive feedback loop regulation mechanism (222). The upregulation of CCAT1-L, the long isoform of CCAT1 (cancer associated transcript 1), is

associated with tumor stage and progression in CRC (223): CCAT1-L accumulates within the nucleus, where it interacts with a MYC enhancer (MYC-335), which in turn interacts with MYC promoter through chromatin looping, thus positively regulating MYC expression. Moreover, CCAT1-L knockdown reduces MYC mRNA levels, suggesting that this lncRNA regulates MYC expression in *cis* (224). Also PVT1 (plasmacytoma variant translocation 1) is associated with CRC and seems to be necessary to maintain high levels of MYC protein; PVT1 is upregulated in CRC because of a copy number amplification on chromosome 8q24; knockdown of PVT1 decreases proliferation and migration *in vitro* by reducing MYC tumorigenic activity (225). In addition, PVT1 is the precursor of several spliced ncRNAs, including a cluster of six miRNAs (miR-1204, miR-1205, miR-1206, miR-1207-5p, miR-1207-3p, miR-1208) involved in carcinogenesis (226). PVT1 is also a target of p53, which binds and activates a DNA canonical response element near to miR-1204 coding region, thus increasing miRNA expression; ectopic expression of miR-1204 increase p53 levels inducing cell death in CRC cell lines (227).

Another lncRNA widely studied in several tumors is HOTAIR, which regulates gene expression through PRC2-induced chromatin remodeling, thus promoting metastasis in breast cancer (228); HOTAIR and PRC2 levels are correlated also in CRC, suggesting an involvement of the lncRNA in maintaining the mesenchymal and undifferentiated status in CRC cells (229). CRNDE (colorectal neoplasia differentially expressed) also interacts with PRC2, thus modulating chromatin remodeling (60). CRNDE promoter is bound by pluripotency-related transcription factors, including MYC, and CRNDE knockdown induces decreased expression of many pluripotency markers, such as SOX2, KLF4, NANOG, and OCT4; indeed, CRNDE expression is highest during the early stages of mammals development, decreasing progressively thereafter (230).

The lncRNA TUSC7 is regulated by p53 and acts as tumor suppressor, impairing tumor growth both *in vivo* and *in vitro*. In particular, this function is due to TUSC7 fourth exon, which contains two binding sites for miR-211, involved in cell growth regulation: TUSC7 acts as miRNA

sponge, sequestering miR-211 and causing its downregulation. Moreover, TUSC7 is a miR-211 target, creating a negative feedback regulation mechanism (71). Similarly, FER1L4 (fer-1 like family member 4), a lncRNA downregulated in CRC, acts as tumor suppressor by sequestering miR-106a, a well known oncogene in colorectal cancer. Ectopic expression of FER1L4 decreases miR-106a levels and inhibits proliferation, invasion and migration in CRC cell lines (231).

H19 upregulation has been observed in several tumors, including CRC (232). H19 function is associated with miRNAs: it is the precursor of miR-675, which regulates cell cycle and EMT by targeting RB1 (69); on the other hand, H19 contains binding sites for let-7 miRNA family, key tumor suppressors in CRC (70).

A CRC-associated lncRNA recently discovered is UPAT (UHRF1 protein associated transcript), essential for surviving and tumorigenicity of CRC cell lines. UPAT interacts with UHRF1 (ubiquitin like with PHD and ring finger domains 1), an ubiquitin ligase regulating chromatin remodeling and DNA methylation through the degradation of DNA-methyltransferase 1 (DNMT1) and histons; the interaction with UPAT stabilizes UHRF1 and inhibits its degradation via ubiquitination. Moreover, UPAT interacts with the promoter of Sprouty 4 (SPRY), a key regulator of MAPK pathway, which expression is induced by high levels of the lncRNA (233). Among other lncRNA interacting with DNMT1, DACOR1 (DNMT1-associated colon cancer repressed lncRNA 1) is downregulated in CRC cell lines. Ectopic expression of DACOR1 induces a gain of DNA methylation, although it doesn't alter DNMT1 protein levels: this evidence suggests that DACOR1 recruits DNMT1 on specific *loci* of the genome. In addition, increased expression of DACOR1 inhibits cell growth through TGF-β (transforming growth factor beta 1) pathway, a tumor suppressive pathway associated with CRC, and induces the downregulation of CBS (cystathionine β-synthase) and the subsequent increment of SAM (S-adenosyl methionine), essential donor of methyl groups for the activity of DNA methyltransferases (234).

2.6.2 CircRNAs and colorectal cancer

CircRNAs represent a class of quite unknown gene expression regulators; they compete with other RNAs for the interaction with miRNAs and RNA-binding proteins (RBPs), probably modulating their concentration (91). Reduced circRNA expression has been often observed in several tumors compared to normal tissues: this could be due to the absence of linear ends in circRNAs, which are resistant to exonucleases and tend to accumulate within cells with low proliferative rate (79).

Few data have been reported on circRNA involvement in colorectal cancer carcinogenesis. A recent study showed that circRNA expression levels are globally reduced in CRC tissues compared to normal mucosa; the negative correlation observed between circRNA expression and cell proliferation rate could be explained by several factors: i) an altered backsplicing machinery in tumor cells; ii) an increased circRNA degradation mediated by miRNAs deregulated in tumors; iii) circRNA "dilution" provoked by high proliferative rate (235). According to this evidence, the downregulation of hsa_circ_001988, circRNA transcribed from the FBXW7 (F-box and WD repeat domain containing 7) gene, has been recently reported in CRC tissues compared to normal adjacent tissues (236). In addition, it has been reported that circRNA expression is globally reduced in CRC cell lines bearing KRAS mutations compared to KRAS wild type CRC cells, while the expression of circRNA host genes is unvaried; authors hypothesized that reduced circRNA levels could be explained by their increased export toward exosomes (237).

In contrast, another study showed that hsa_circ_001569, circRNA transcribed from the ABCC1 (ATP binding cassette subfamily C member 1) gene, is upregulated in CRC tissues and positively correlates with aggressive characteristics, such as tumor stage, lymph node involvement, distant metastasis and poor differentiation. *In vitro* experiments demonstrated that hsa_circ_001569 promotes cell proliferation and migration and inhibits apoptosis, thus fostering tumor progression.

In particular, hsa_circ_001569 plays its oncogenic action by sequestering miR-145, well known tumor suppressor miRNA in CRC, thus inducing upregulation of miR-145 targets (238).

2.6.3 MiRNAs involved in CRC carcinogenesis

It is now clear that miRNAs play a crucial role in the regulation of cell physiology by interacting with their mRNA targets; it has been widely demonstrated that these ncRNAs regulate key processes in carcinogenesis, such as proliferation, invasion, migration, EMT and apoptosis. A lot of miRNA-coding genes are located at genomic regions frequently rearranged in tumors, such as fragile sites, minimal regions of loss of heterozygosity (LOH), minimal regions of amplification (minimal amplicons), or common breakpoint regions, strengthening miRNA role in carcinogenesis (239, 240).

The altered expression of several miRNAs has been reported in many different tumors, including colorectal cancer (15, 240). Depending on biological function, determined by the functions of mRNA targets, miRNAs can be classified as oncogenes (oncomiRs) or tumor suppressors (241, 242): i) oncomiRs are transcribed from genes localized in genomic regions which are amplified or upregulated in the tumor, and induce increased cell proliferation, angiogenesis, invasiveness, and reduced apoptosis; ii) tumor suppressor miRNAs are transcribed from genes located in regions which are deleted or silenced in the tumor, having an opposite biological effect. Several miRNAs acting as oncogenes or tumor suppressor have been associated with CRC: among the most known oncogenic miRNAs there are the miR-17-92 cluster, miR-21, miR-31, miR-106a, while among the tumor suppressor miRNAs there are the let-7 and miR-34 miRNA families, miR-143 and miR-145 (15).

2.7 Exosomes in tumor progression

The role of mediators of intercellular communication makes exosomes fundamental tools used by the tumor to spread pro-tumor *stimuli*. The tumor promotes its own progression by secreting several signals into the extracellular environment, thus building around itself a niche suited to its development and survival; recently it has been demonstrated that many of these signals are delivered by exosomes. Exosome involvement in carcinogenesis is supported by numerous studies reporting a greater amount of circulating vesicles in patients affected by several tumors compared with healthy individuals (243, 244, 245).

Exosomes secreted by tumor cells promote tumor progression in many ways; one of the most important mechanisms is certainly represented by the suppression of the immune response. Tumor-secreted exosomes can interact with immune system cells causing their death through several molecules, such as FasL (Fas ligand) and TRAIL (TNF-related apoptosi-inducing ligand) expressed on vesicle membranes, which bind their receptor on CD8$^+$ T lymphocytes (246, 247); in addition, the direct fusion of exosomes secreted by skin melanoma and CRC cells with monocytes inhibits monocyte differentiation into dendritic cells, instead inducing the secretion of cytokines causing T lymphocyte death (248).

Exosomes can also foster tumor progression by spreading pro-proliferative, pro-metastatic and pro-angiogenic signals. Exosomes express on their surface growth factor receptors, such as HER (human epidermal growth factor receptor) receptor family, and carry their ligands, which can be transferred to recipient cells and promote proliferation (249). Exosomes secreted by breast cancer and CRC cells express on their surface amphiregulin, an EGFR ligand, which interacts with its receptor on the plasma membrane of tumor cells, thus increasing their invasiveness (250). Finally, EGFR molecules contained within tumor-derived exosomes can be transferred to endothelial cells, inducing VEGF expression and activating angiogenesis (251).

In addition, exosomes can confer upon tumor cells resistance to immunotherapy treatments. Standard treatment for HER2-overexpressing breast carcinoma is represented by a monoclonal antibody called trastuzumab: the drug binds the extracellular domain of HER2/ERBB2 (erb-b2 receptor tyrosine kinase 2) receptor, resulting in antitumor effects, such as the inhibition of mitogenic signal transduction, the induction of antibody-dependent cell-mediated cytotoxicity (ADCC) by natural killer (NK) cells, and probably the downregulation of the receptor itself through its internalization. Tumor-derived exosomes express on their surface HER2, which binds and sequesters trastuzumab molecules, thus reducing their availability and efficiency on tumor cells; it has been observed that exosome-mediated protective effect is higher in advanced tumors because they produce more vesicles, which strongly limit drug availability thus promoting tumor progression and metastatization (252).

2.8 NcRNAs in biological fluids

Although they play a crucial regulatory role within cells, ncRNAs are also secreted into the extracellular environment; frequently, ncRNAs are contained within exosomes, which make them resistant to degradation, more stable and able to reach distant body areas. NcRNAs have been found in all biological fluids analysed to date, first of all blood. Searching for non-invasive methods to diagnose tumor, blood represents the most interesting sample, because: i) it can be obtained in a simple and non-invasive way; ii) since it circulates throughout the entire body, blood collects exosomes produced by all cell hystotypes, including brain cells, because exosomes are small enough to cross the blood-brain barrier. Recently, several studies reported that altered serum levels of ncRNAs are associated with many tumors, suggesting that ncRNAs could be used as diagnostic markers; circulating miRNAs have been widely analysed, both in neoplastic and non-neoplastic diseases, while research on circulating lncRNAs and circRNAs is just beginning.

Few data have been produced on CRC and lncRNAs. The lncRNA CRNDE is upregulated in CRC tissues; serum levels of the splicing variant CRNDE-h are significantly higher in CRC patients compared to healthy individuals, and showed a good diagnostic accuracy (87% sensitivity and 93% specificity) (253). Another study identified through microarray experiments a molecular signature composed by three new lncRNAs (XLOC_006844, LOC152578 e XLOC_000303) in plasma of CRC patients (254). Successively, it has been reported that CCAT1 and HOTAIR levels are increased, while lincRNA-p21/TP53COR1 (tumor protein p53 pathway corepressor 1) levels are decreased in plasma of CRC patients compared to healthy individuals; the predictive power of these lncRNAs raises by combining HOTAIR and CCAT1. Moreover, HOTAIR and CCAT1 plasma levels decrease and lincRNA-p21 levels increase after surgical excision of the tumor (255).

No data are to date available about circulating circRNAs and CRC.

The study presented in this doctoral thesis aims to investigate lncRNA and circRNA involvement in colorectal cancer through gene expression analysis in tissues and serum exosomes of CRC patients compared to normal colon mucosa and healthy individuals respectively. Through ROC curve computation, we assessed the diagnostic accuracy of ncRNAs in serum exosomes to evaluate their possible application as CRC biomarkers. By treating CRC cell lines with MAPK inhibitors, we also investigated ncRNAs involvement in MAPK pathway, which regulates cell cycle and has been associated with CRC. Finally, we performed ASO-mediated silencing of UCA1 in HCT-116 cells and analysed the expression of its mRNA and miRNA targets, aiming to investigate the effect of lncRNA-mRNA/miRNA and circRNA-miRNA interactions on the stability and the eventual degradation of the same RNA molecules.

3. Materials and methods

3.1 Cell lines

HCT-116 and Caco-2 cell lines were obtained from the Interlab Cell Line Collection (ICLC), an International Repository Authority within the IRCCS Azienda Ospedaliera Universitaria San Martino-IST Istituto Nazionale per la Ricerca sul Cancro (Genova, Italy); both cell lines derive from primary tumor. HCT-116 cells, bearing KRAS mutation, were cultured in Roswell Park Memorial Institute-1640 (RPMI-1640) medium (Gibco), supplemented with 10% fetal bovine serum (FBS) (Gibco) and 2mM L-glutamine (Lonza); Caco-2 cells, bearing wild type KRAS, were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Sigma), supplemented with 20% FBS (Gibco), 2mM L-glutamine (Lonza) and 1% non-essential amino acids (NEAA) (Lonza); both medium were supplemented with penicillin/streptomycin (10.000 U/ml) (Gibco). Cells were cultured at 37°C with 5% CO_2.

3.2 Isolation of exosomes secreted by CRC cell lines

Exosomes secreted by both cell lines were isolated from culture medium 3 days after seeding in 75 cm^2 flasks: the medium was harvested and centrifuged at 300 *g* (4°C, 10 minutes) on a Beckman J-6M/E centrifuge (rotor TY.JS 5.2) to pellet debris. Supernatant was then centrifuged at 16,500 *g* (4°C, 30 minutes) on Beckman L8 70M ultracentrifuge (rotor SW28), and then filtered through a 0.2 μm filter. The final supernatant was ultracentrifuged at 120,000 g (4°C, 70 minutes) on Beckman L8-70M ultracentrifuge. Exosome pellet was directly lysed for RNA isolation. FBS used to culture the cells was deprived of exosomes through the same protocol, removing exosome pellet and harvesting FBS supernatant for subsequent use.

3.3 RNA isolation from cell lines and exosomes

Total RNA was extracted from cell lines and exosomes with TriZol (Invitrogen), according to manufacturer instructions, and quantified by GenQuant pro spectrophotometer (Biochrom) and Qubit fluorescence quantification system (Invitrogen).

3.4 RNA isolation from CRC formalin-fixed, paraffin-embedded samples

For our study, 20 CRC patients were recruited at Azienda Ospedaliero-Universitaria Policlinico Vittorio Emanuele (Catania, Italy); formalin-fixed, paraffin-embedded (FFPE) tumor tissues and distal normal mucosa were retrieved at Unit of Anatomical Pathology, Department G.F. Ingrassia, University of Catania (Catania, Italy).

RNA was extracted from FFPE samples through PureLink FFPE RNA Isolation Kit (Ambion), according to manufacturer instructions, and quantified by GenQuant pro spectrophotometer and Qubit.

3.5 Exosome and RNA isolation from CRC patient serum

For serum exosome analysis, 20 CRC patients and 20 healthy individuals were recruited at Istituto Oncologico del Mediterraneo (IOM) (Viagrande, Catania, Italy) and Azienda Ospedaliero-Universitaria Policlinico Vittorio Emanuele (Catania, Italy). Following formal written consent, patients underwent fasting venous blood sampling. To separate serum from cellular component, blood samples were centrifuged at 15,000 g (4°C, 15 minutes) on Beckman J-6M/E centrifuge; supernatant was harvested and centrifuged again under the same conditions to remove eventual circulating cells or debris which could alter subsequent analysis. Supernatant serum was harvested and stored at −80°C until analysis.

Patient serum circulating exosomes were isolated with ExoQuick (System Byosciences), a proprietary polymer that gently precipitates exosomes, according to manufacturer instructions.

Exosome pellet was directly lysed with TriZol for total RNA isolation; extracted RNA was quantified by GenQuant pro spectrophotometer and Qubit.

3.6 PCR primer design

By literature data mining, we selected for our study 17 lncRNAs involved in CRC or in other tumors, and regulating carcinogenesis processes, such as cell cycle, apoptosis, gene expression, splicing (Table 3.1).

CCAT1	LIT1	PTENP1
CCAT2	MALAT1	TUG1
CRNDE	MEG3	UCA1
H19	MIR17HG	WRAP53
HOTAIR	PCAT1	ZEB2AS1
LINK-ROR	PCAT6	

Table 3.1: cancer-related lncRNAs selected by literature data mining.

To analyse the expression of selected lncRNAs, we designed PCR primers; in addition, we designed specific primers for two housekeeping genes used for normalization: PPIA (peptidylprolyl isomerase A) and GAPDH (glyceraldehyde 3-phosphate dehydrogenase). Primer design was performed by using the online tool PrimerBlast (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). Primers are shown in table 3.2.

Gene	Forward primer	Reverse primer
CCAT1	GGAAAGGTGCCGAGACATGA	GCCATACAGAGCCAACCTGG
CCAT2	CCTGCAGAGGGCACTAGACT	CCAGGGTCAGGCAATTGGTC
CRNDE	TCGATCGCGCTATTGTCATGG	CGCCTCGCTTAGACATTGGC
H19	GAGTCGGCACACTATGGCTG	GTCCGGATTCAAAGGCCCAG
HOTAIR	GTGGGGGTAAGAGAGCACCA	GAGGGTCTAAGTCCCGGGTG
LINK-ROR	CTTCCAGGTCTCAGGAATGGG	AGGTGGCTGGTGAGAGATCC
LIT1	CAGGACATGCCCACACTAGC	CGTTTCCTCTTACAGTGGCGG
MALAT1	TAGCTTGGATCCTTGTGGGC	AACCCACCAAAGACCTCGAC
MEG3	GCCAGCTGTCCCTCTTACCT	GCATAGCAAAGGTCAGGGCT
MIR17HG	TGCCACGTGGATGTGAAGATT	GTTCTCCAGGAAGTTGCAGGC

PCAT1	AGGCCAAGCTACAGGGAGAAA	TGGCATTATCTTGGGAGGTTCC
PCAT6	CTCTTGGACAACACTCCGCC	ATCTGGCGCTGCACAAATGA
PTENP1	CATTCTTGCATGTATTTGGGTTAGG	GGTATATGGTCCAGAGTCCAGC
TUG1	CTTTCACCATGGGGTGCCTG	AGCTGCTTTACACTGGGTGC
UCA1	TGCACCCTAGACCCGAAACT	CAAGTGTGACCAGGGACTGC
WRAP53	ACCGGACTGTGCGTGTTTTT	CTGGGCTGAAGGCTATGCAG
GAPDH	GGCTAGCTGGCCCGATTT	AGGCGCCCAATACGACCAAA
PPIA	ACTGAGTGGTTGGATGGCAAGC	AGCGCTCCATGGCCTCCACA

Table 3.2: PCR primers designed for selected lncRNAs and housekeeping genes.

Primers for 31 circRNAs were retrieved by a paper published by Memczak and colleagues on Nature in 2013 (80) (Table 3.3).

Gene	Forward primer	Reverse primer
circRNA1	TGCCCGTCATGATTCTCACAGAGT	ATTCCCTCACAGCCTCATTAGGGT
circRNA2	TCTTGAAGTGAGCTGTGTCAGGCA	TGTGGGAAGCAAGGAGGAACTTGA
circRNA3	TCTAGACTGCAGAGATGAATGAAAA	ATCTTCATCAGTGATCTCACTTTCC
circRNA4	TGTGCTCCTGCTCATACTGGTCAA	TCAGTGCCTCGAAAGAACTTCCGT
circRNA5	TGGATGAAGATGGGGAAAAG	CGGTTGAACAGCTTTTGGTT
circRNA6	CGACTCTTCCTGGTGAAGCATCTGAT	TACTTGAACGAAGCTCCATGCAGC
circRNA7	AGCTATTGCCTTAGATCCTTCAAGT	TAAATTCTGTTCGTTTAATGGCTTC
circRNA8	ACAGACACTCCTGCAACCCAGTTT	CGTCACCTTCCCAGGATACCTCT
circRNA9	ACTATCGACAGCAGACCTTTGAA	GAATCTTTTCAGGAAATGCCTCT
circRNA10	GGTACACGAGCCTCTGAACTCTT	GAGATGAAAGTGTCGTCCATCTACT
circRNA11	GCTCAGCCTCTGTGAAGAGC	GGGAGGCTACTTCCTCATCC
circRNA12	AACTCGGCTCTATTTTCAGTTTCTT	CCATCTAGACAACCATACAGAGTCA
circRNA13	GTTGAAAGTAGCCCGAGCAG	GGGTCCTCAGGTGTTTCGTA
circRNA14	GTGGAGGTCGTGGAGATCAT	CCACCACCTTGTCTCCATCT
circRNA15	TATGGTGGACTACTTCTTTCAGAGG	CTTTTCTCTTCAAGACACATGCAAC
circRNA16	TGGTGGGTAGACCAAGACTTGTGA	TATGTTGGTGGATCCTGTTCGGCA
circRNA17	AGTCCCACTTCAAACATTCGTCT	CACCTTCACTATTACGATACCATCC
circRNA18	GGGAAGTGGAACCGGATATT	GGCTTGCTCCACTCTTTCAG
circRNA19	TATGATTACACAGAAAAGCAAGCAG	ACACTGTCAAGATCAATTCCTTCTC
circRNA20	CCTACAGAAACAAAATGTCAACAAA	GAACCTAAAAGGTCAAAACAGAACA
circRNA21	TTTCTTGGTTTCATGGATACTGATT	TAAGATGACAAAGAAGGTAGGTGCT
circRNA22	GATGTGATCATCTTACCCTAAAGGA	TGATTCTAGGTCTACGTTGTTTTCC
circRNA23	ACTCACTGCTCAGAGTTTCAAATGT	GCTGTCGAGGCTGTAGAAGTATAAG
circRNA24	ACTGAAAATTCTGGATAATGGTGAA	ATAAGTGTTGTCAGTCACAAAAGCA
circRNA25	ACTTCCATTTTTGATGACTTCTCAC	GCTCAGCTGTCCATAAGAAGTTAAG
circRNA26	TGAAAATATTGATACTCCAGGAACC	AAATTCATCATCATCATCATCCTCT
circRNA27	GAAGCAAGTGGCTATCATAGAAAAG	GTCTCCACCTCCTTCTGTAGCTT
circRNA28	TTTTCCTTAGAGATCCACATTCAAG	GTAAGCCCTGCTGAATTAGAAAAAC

circRNA29	CCAAGTTGAAAATGGACGAGTAAT	CAATGCCAGATACTTAGAGCTTTTC
circRNA30	CCAGTATCACAGTATTGTGCTGTTTA	TTTGGGACACTCAGTTTTATGC
CDR1as	ACGTCTCCAGTGTGCTGA	CTTGACACAGGTGCCATC

Table 3.3: circRNA primers retrieved by Memczak *et al.*, Nature 2013.

3.7 LncRNA and circRNA expression analysis by Real-Time PCR

We investigated the expression of selected lncRNAs and circRNAs in CRC cell lines and exosomes, and in CRC patient tissues and serum exosomes though Real-Time PCR by using Power SYBR Green RNA-to-CT™ 1-Step kit (Applied Biosystems), according to manufacturer instructions. All Real-Time PCR reactions were performed on a 7900HT Fast Real-Time PCR System (Applied Biosystems).

Differentially expressed (DE) lncRNAs and circRNAs were identified through SDS RQ Manager 1.2 software (Applied Biosystem); normalization was performed using PPIA and GAPDH for tumor tissues and serum exosomes respectively. Expression fold changes of DE ncRNAs were calculated by applying the $2^{-\Delta\Delta Ct}$ method. Statistical analysis performed by paired T-test was used to compare tissue sample ΔCts, while unpaired T test was used for serum exosome ΔCts; statistical significance was established at a p-value ≤ 0.05.

3.8 Computational analysis

To confirm our results, we analysed DE lncRNA expression in several datasets derived from microarray experiments deposited in GEO DataSets (https://www.ncbi.nlm.nih.gov/gds/) (Table 3.4).

GEO DataSets ID	Samples
GSE10715	peripheral blood (19 tumors vs 11 controls), blood smears (10 tumors vs 10 controls)
GSE13294	fresh-frozen primary tumor tissues (74)
GSE14333	fresh-frozen primary tumor tissues (290)
GSE17538	fresh-frozen primary tumor tissues (232)

GSE18088	frozen or formalin-fixed and paraffin-embedded tissues (53)
GSE18105	cancer tissues and normal adjacent tissues (17), laser capture microdissected primary tumors (77)
GSE20916	distant full-thickness normal colon, normal colon mucosa dissected from tumor, adenoma, and carcinoma (40)
GSE2109	tumor tissues (315)
GSE21510	laser capture microdissected and homogenized tumors (148)
GSE23878	tumor tissues (35 tumors vs 24 matched controls)
GSE33114	fresh frozen or formalin-fixed and paraffin-embedded tissues (90 tumors vs 6 controls)
GSE35896	primary tumor tissues (62)
GSE3629	53 ulcerative colitis tissues (10 UC-associated cancer, 43 non-cancer), 60 sporadic cancer tissues
GSE39582	fresh-frozen primary tumor tissues (566)
GSE41258	primary colon adenocarcinomas, adenomas, metastasis and corresponding normal mucosae (390)

Table 3.4: CRC GEO DataSets used to confirm DE lncRNA expression.

In addition, we performed a computational analysis to identify potential mRNA and miRNA targets of DE ncRNAs.

For DE lncRNAs, data about correlation of expression were retrieved in three databases containing microarray data: MEM (http://biit.cs.ut.ee/mem/), which shows data about positive and negative correlation, lncRNAtor (http://lncrnator.ewha.ac.kr/index.htm) and BioGPS (http://biogps.org/), which only provide data about positive correlation; through these databases we obtained three list of mRNAs (one from each database) with correlation of expression for each DE lncRNA. These lists were submitted to the database cBioPortal (http://www.cbioportal.org/): according to CRC data published on Nature in 2012 within the TCGA (The Cancer Genome Atlas) project (256), cBioPortal database allowed us to identify those genes within the lists which have been shown to be altered in CRC (because of point mutations, deletions, gene amplification,

upregulation or downregulation); we selected genes altered in a percentage of cases corresponding to the first and the second quartiles. Gene lists filtered through cBioPortal were then unified for each DE lncRNA; for each gene, we retrieved from Gene NCBI (https://www.ncbi.nlm.nih.gov/gene/) the RefSeq accessions, which were used to perform a sequence alignment between the DE lncRNA and the mRNA targets by using Nucleotide BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (settings: "Somewhat similar sequence" and "Species-specific repeats for Homo sapiens"). We selected only those genes showing a complementary (not homologous) region longer than 25 nucleotides (Figure 3.1). Therefore, for each DE lncRNA, we obtained a list of genes involved in CRC and characterized by correlation of expression and sequence complementarity with the DE lncRNA.
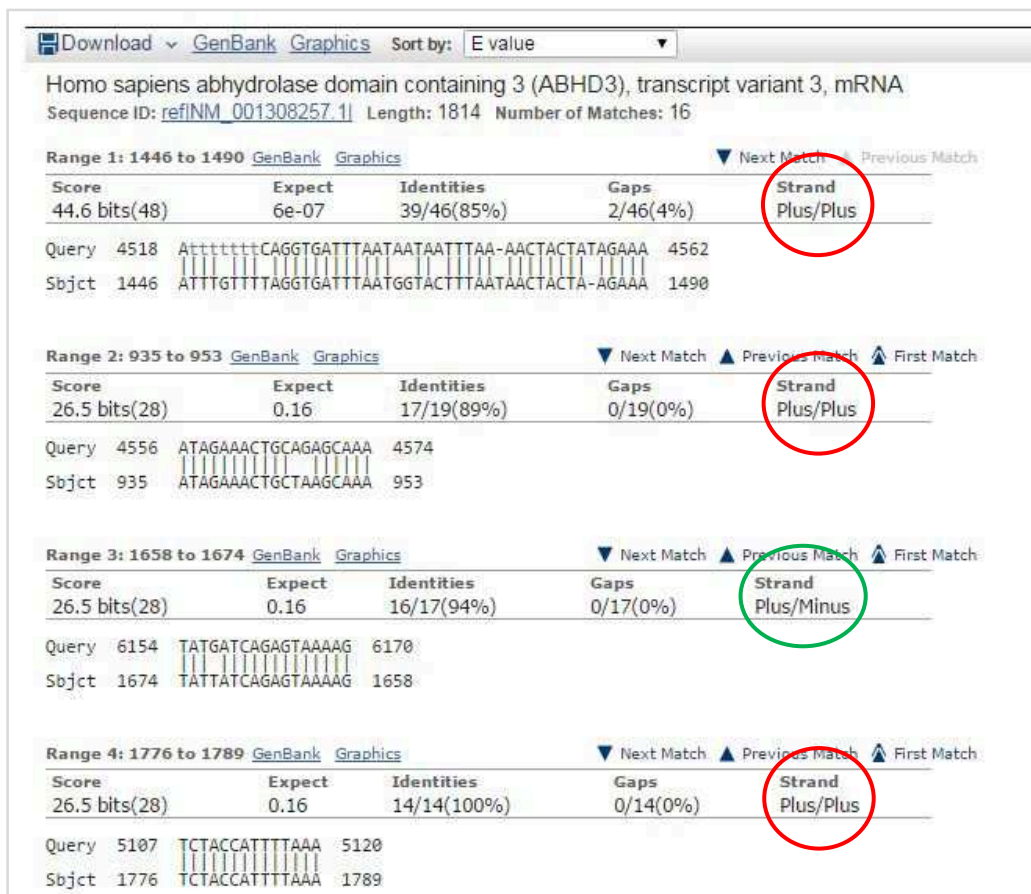
Figure 3.1: selection criteria for target genes of DE lncRNAs after sequence alignment: "strand plus/plus" indicates sequence homology, while "strand plus/minus" indicates sequence complementarity.

In addition, we aligned each DE lncRNA sequence with the whole genome sequence by using two different databases: NR and Human genomic and transcripts; this additional analysis allowed us to identify genomic regions characterized by sequence complementarity with each DE lncRNA, which could bind DNA modulating the expression of near target genes. Using again Nucleotide BLAST, we obtained genes, transcripts and chromosome tracts complementary to each DE lncRNA; again, we selected only those complementary tracts longer than 25-30 nucleotides. The list including complementary genes (or the genes nearest to the complementary region) was analysed to identify genes which expression was correlated to lncRNA expression and genes involved in CRC; for this purpose, we used again MEM, selecting only genes with p-values lower that 10^{-6}, and lncRNAtor, comparing the two gene lists; selected genes were then submitted to cBioPortal to investigate their involvement in CRC, again selecting only genes altered in a percentage of cases corresponding to the first and the second quartiles. This second analysis provided us a second list (for each DE lncRNA) of genes involved in CRC and characterized by correlation of expression and sequence complementarity with the DE lncRNA. The two final lists were unified for each DE lncRNA; finally, a literature search was performed to further filter the list, selecting only genes involved in carcinogenesis and cancer-related processes, such as regulation of splicing, chromatin status, cell cycle, apoptosis and metastasis.

Interactions with miRNAs have been investigated for both DE lncRNAs and circRNAs. First of all, miRNAs involved in CRC were identified by mining two databases, miRCancer (http://mircancer.ecu.edu/) and miR2Disease (http://www.mir2disease.org/); mature sequences of these miRNAs were retrieved from miRBase (http://www.mirbase.org/) and then used to perform a sequence alignment on each DE ncRNA again through Nucleotide BLAST (settings: "Word size 7", "Somewhat similar sequence" and "Species-specific repeats for Homo sapiens"). We selected only those miRNAs showing at least four complementary tracts located throughout the whole ncRNA sequence but always within the miRNA seed region. Successively, selected miRNAs were further

filtered selecting only those characterized by negative correlation of expression (anticorrelation) with each DE ncRNA.

3.9 Biological network construction and Gene ontology analysis

To understand the effects of DE ncRNA deregulation, we built two biological networks based on the interactions between all DE ncRNAs and their mRNA targets and all DE ncRNAs and their miRNA targets identified through the computational analysis; in addition, the software retrieved the nearest neighbours of target genes. Networks were built by using the software Cytoscape 2.8 (http://www.cytoscape.org/) and the MIMI Plugin. To build the network based on ncRNA-miRNA interactions, miRNA validated target retrieved from the database TarBase v7.0 (http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=tarbase/index) were used.

Functional classification of Gene ontologies and pathway enrichment analysis were performed using all genes included in the two networks by using: i) the FatiGO online tool available on Babelomics 4.3 (http://v4.babelomics.org/functional.html), applying the Fisher's exact test with false discovery rate (FDR) Correction (p-value < 0.005); ii) the GeneMania online tool (http://www.genemania.org/), calculating q-values with FDR corrected hypergeometric test.

3.10 ROC curve analysis

To evaluate the diagnostic accuracy of ncRNAs deregulated in patient serum exosomes, ROC (receiver operating characteristic) curves were computed through the software MedCalc v 15.11.4. The area under the curve (AUC) and 95% confidence intervals (95% CIs) were calculated to assess the accuracy of sensitivity and specificity parameters and to find an appropriate cut-off point. Statistical significance was established at a p-value ≤ 0.05.

3.11 CRC cell lines treatment with MAPK inhibitors

We investigated the effect of MAPK pathway inhibition on DE ncRNAs in HCT-116 cells, characterized by overactivation of the pathway because of KRAS mutation. MAPK inhibition was achieved through treatment with FR180204 and U0126. FR180204 is an ATP-competitive inhibitor of ERK1 and ERK2 and a weaker inhibitor of the related kinase p38α. $2x10^5$ cells per well were seeded in 12-well plates and cultured in serum starvation conditions (1% FBS) for 6 hours; cells were then treated with 5 μM, 10 μM, 25 μM, 50 μM FR180204 (ERK Inhibitor II, Calbiochem). U0126 is a highly selective ATP-non-competitive inhibitor of both MEK1 and MEK2 kinases, specifically preventing MEK1 and MEK2 phosphorylation and activation *in vitro* by binding the inactive enzymes and blocking the recruitment of Raf-1 kinase. $3x10^5$ cells per well were seeded in 6-well plates and cultured in serum starvation conditions (1% FBS) for 14 hours; successively, cells were treated with 5 μM, 10 μM, 20 μM U0126 (MEK1/2 inhibitor, Merck). Cells were exposed to inhibitors for 24 hours. Control samples were treated with an equivalent volume of DMSO (solvent of the drugs used for treatments). All experiments were performed in biological triplicates.

Cell viability was assessed through MTT assay at 24 hours post-treatment.

3.12 Protein extraction and Western Blot analysis

MAPK pathway inhibition was verified by analysing phospho-ERK (pERK) levels through Western Blot. Proteins were extracted with RIPA buffer [2% protease-inhibitor cocktail (Sigma), 1% of both phosphatase-inhibitor cocktail 2, phosphatase-inhibitor cocktail 3 (Sigma)] and quantified by Qubit (Invitrogen). Blotting was performed by iBlot Dry Blotting System (Invitrogen). Proteins were detected using the ECL Plus Western Blotting detection (Amersham). Membranes were bound with monoclonal antibodies to p44/42 MAPK (Erk1/2) and phospho-p44/p42 MAPK (Erk1/2) (Cell Signaling), using beta-actin (Abcam) as loading control. Bound antibodies were detected with peroxidase-labeled goat antirabbit IgG-HRP secondary antibody (Santa Cruz).

3.13 *In vitro* silencing of UCA1

We performed UCA1 silencing in HCT-116 cell line though antisense oligonucleotide (ASO) technology. $4x10^4$ cells per well were seeded in 24-well plates and simultaneously reverse-transfected with 50 pmoles of UCA1 specific Antisense LNA™ GapmeR (Exiqon), using Lipofectamine RNAiMAX Reagent (Invitrogen), according to manufacturer instruction. All experiments were performed in biological triplicates.24 and 48 hours after transfection cells were lysed with TriZol for RNA isolation. UCA1 mRNA and miRNA targets were selected through our computational analysis, the analysis of GEO DataSets (GSE15960, GSE33114, GSE39582, GSE8671), and literature search. Specific PCR primers were designed and used to evaluate the expression of mRNA targets by Real-Time PCR (Table 3.5), while miRNA expression was evaluated through TaqMan miRNA assays (Applied Biosystems). PPIA and miR-29a were used as endogenous controls.

Gene	Forward primer	Reverse primer
AREG	CCTCACAGCTGTTGCTGTTATT	TTCGTTCCTCAGCTTCTCCTT
ASCL2	CTTGGAGCTGGCCCCATAA	CACTGCAGTCGAGAAGCTGT
LINC01764	CTCCGGACTGCTTCAAGTGT	GTTATTCACCGTCACAAGTTCTCA
MACC1	GCCACACAGAGAGAAATACAAGG	CGTGCGACTAACTCTTGGCA
NCL	CCCAGGGGATCACCTAATGC	CAGTGGTATCCTCAGACAGGC
PDP1	GTAGCCACCAGCACATCACT	AAGAAGTACGATTCCCCGGC
SYNCRIP	TACAAAAGAAGCAGCTCAGGAGG	ATTGTTGGCAACTGAGATGCAG
UGDH	GCAACAGCGATTGGAATGGA	TTTGGAAACAGCTCCCACCA

Table 3.5: PCR primers for UCA1 mRNA targets.

4. Results

4.1 NcRNA expression in CRC cell lines and exosomes

The expression of selected lncRNAs and circRNAs was evaluated through Real-Time PCR in two CRC cell lines, HCT-116 and Caco-2. The results of this preliminary screening showed that almost all analysed molecules are expressed in both cell lines (Figure 4.1).
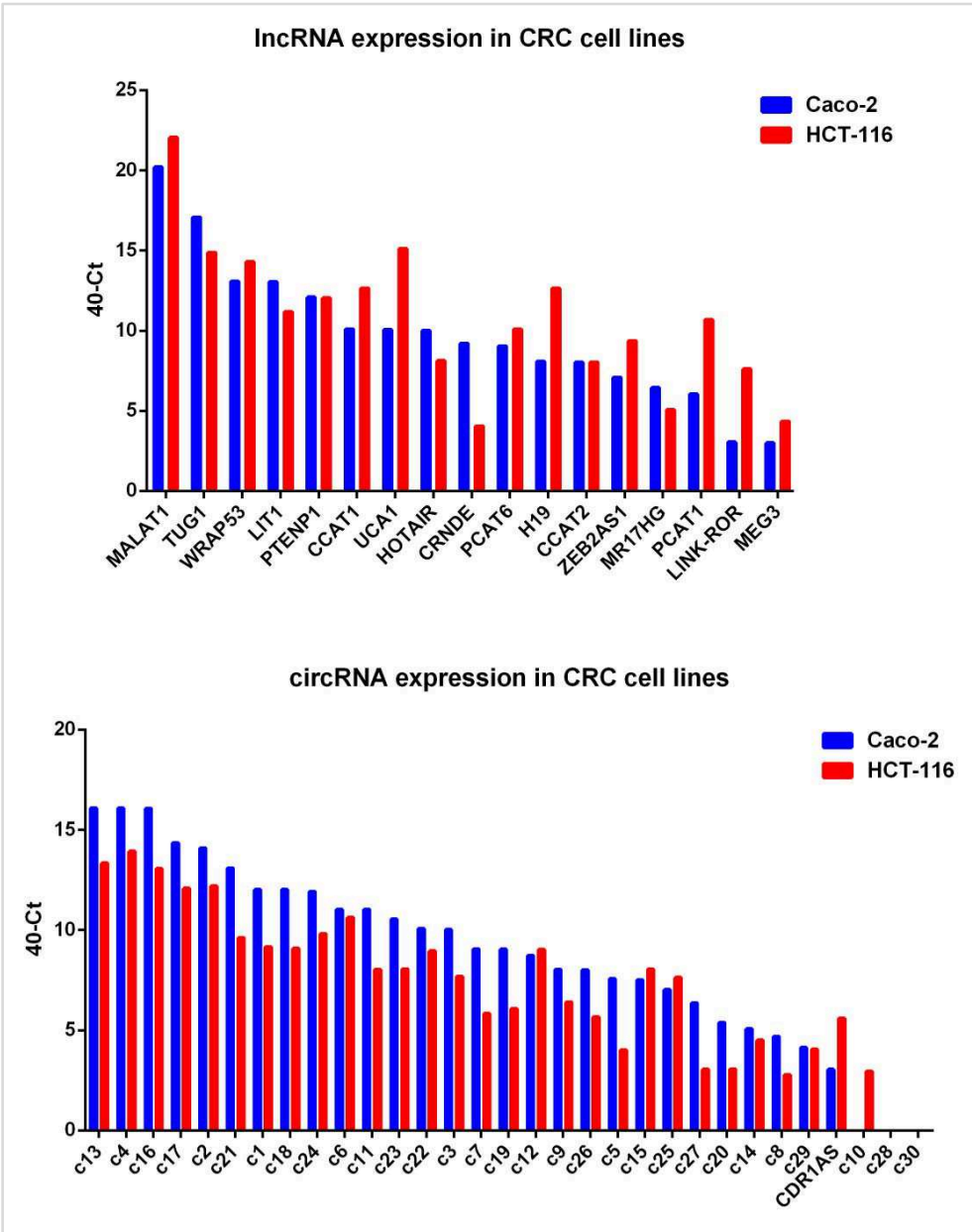
Figure 4.1: expression of selected lncRNAs and circRNAs in CRC cell lines HCT-116 and Caco-2; expression is shown as 40-Ct.

NcRNAs showing the highest expression in CRC cell lines were also tested in exosomes secreted by the same cells, aiming to evaluate their presence within vesicles. Except for circ2, all tested molecules were detected into exosomes secreted by both CRC cell lines (Figure 4.2).
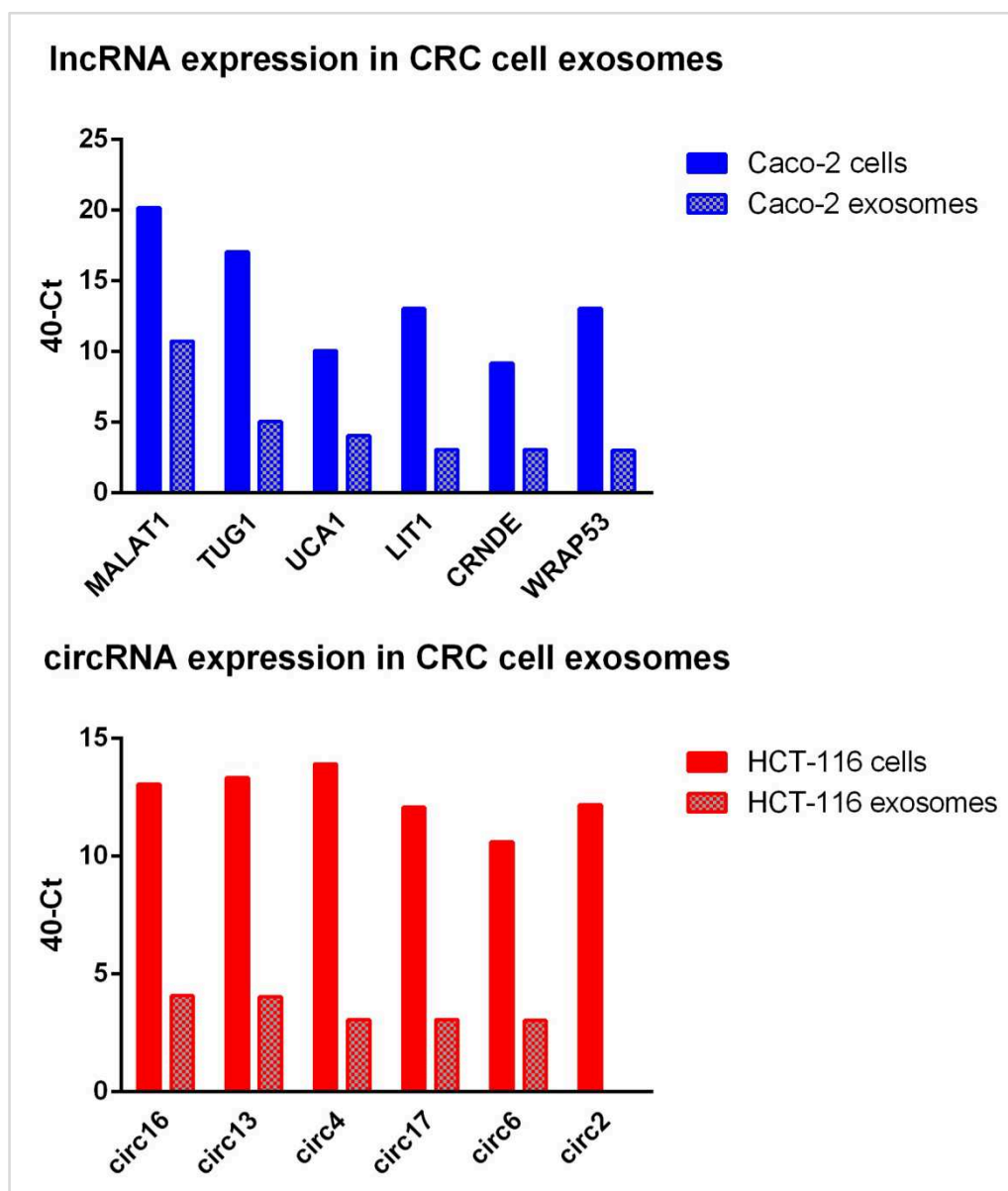
Figure 4.2: expression of lncRNAs and circRNAs in exosomes compared to donor cells; expression is shown as 40-Ct.

4.2 NcRNA expression in FFPE CRC biopsies

Through a trial experiment on three FFPE biopsy pairs, we selected 15 out of 17 lncRNAs (CCAT1, CCAT2, CRNDE, H19, HOTAIR, MALAT1, MEG3, MIR17HG, PCAT1, PCAT6,

PTENP1, TUG1, UCA1, WRAP53, ZEB2AS1) and 6 out of 31 circRNAs (circ2, circ4, circ13, circ17, circ22, CDR1AS) to be analysed in all FFPE samples. We performed Real-Time PCR analysis in 20 FFPE CRC biopsies and their normal adjacent tissues (NATs); results showed significant differential expression of 7 lncRNAs and 1 circRNA (Table 4.1).

	Fold Change	T test p-value
CCAT1	**8.37**	**0.0007**
CCAT2	**4.67**	**0.001**
CDR1AS	**-3.3**	**0.002**
circ13		0.31
circ17		0.23
circ2		0.31
circ22		0.28
circ4		0.46
CRNDE		0.31
H19		0.75
HOTAIR	**9.17**	**0.0008**
MALAT1	**-1.8**	**0.004**
MEG3		0.93
MR17HG		0.63
PCAT1		0.76
PCAT6		0.81
PTENP1		0.83
TUG1	**-2.04**	**0.036**
UCA1	**7.14**	**0.001**
WRAP53		0.83
ZEB2AS1	**12.13**	**0.002**

Table 4.1: statistical results of lncRNA and circRNA expression analysis in 20 tissue biopsies of CRC patients compared to normal adjacent tissues. For each ncRNA, paired T test p-value and fold change are shown.

Our results showed that 5 lncRNAs (CCAT1, CCAT2, HOTAIR, UCA1 and ZEB2AS1) were upregulated, while 2 lncRNAs (MALAT1 and TUG1) and 1 circRNA (CDR1AS) were downregulated in CRC tumor tissues compared to NATs (Figure 4.3).
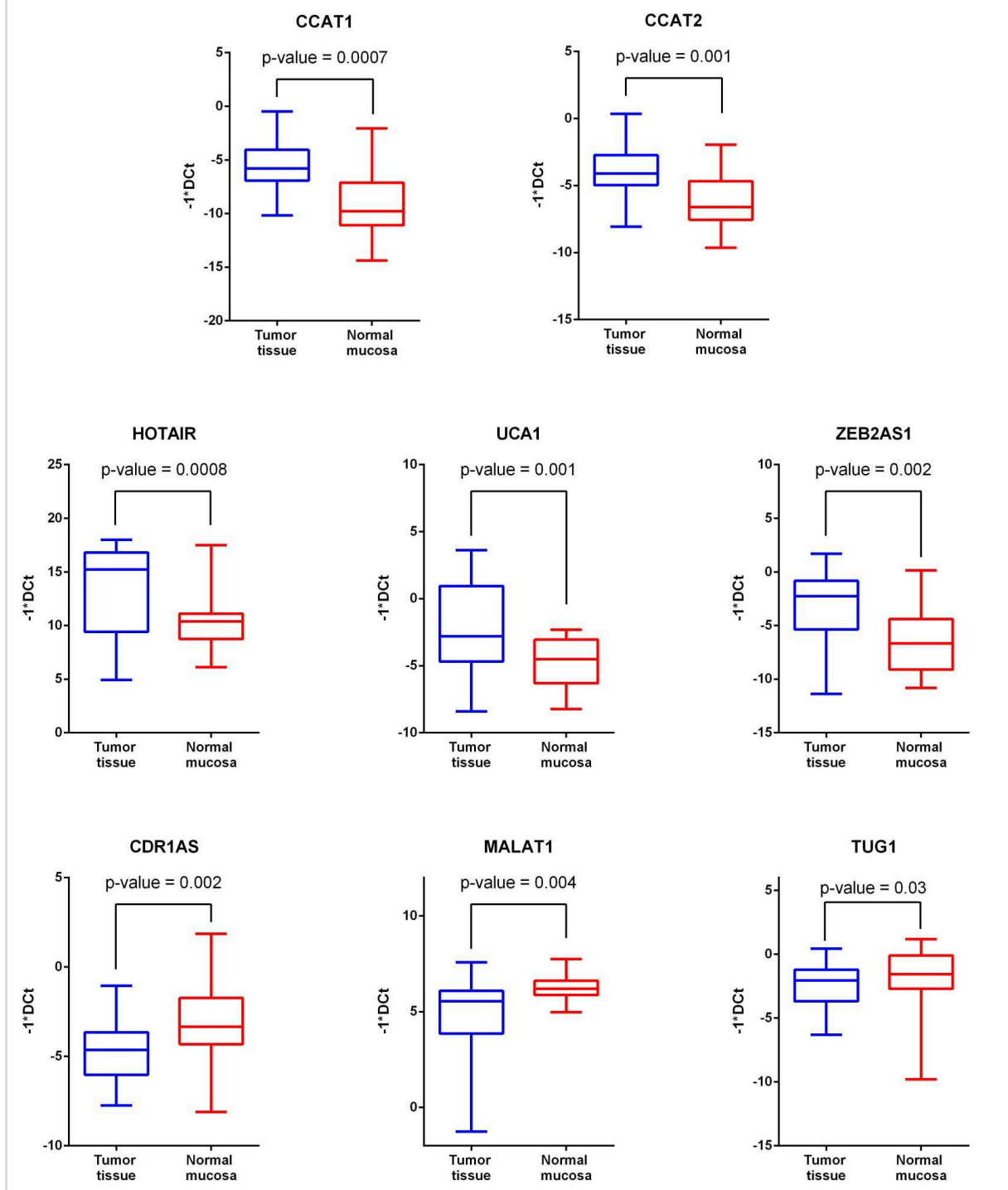
Figure 4.3: box plots showing differential expression of 7 lncRNAs and 1 circRNA in 20 CRC tumor tissues compared to normal adjacent tissues.

4.3 NcRNA expression in serum exosomes of CRC patients

We performed a trial experiment with a small group of samples to identify lncRNAs and circRNAs to be tested in the entire cohort of 20 CRC patients and 20 healthy individuals. The trial experiment allowed us to select 10 lncRNAs (CCAT2, CRNDE, HOTAIR, LIT1, MALAT1, PCAT1, PCAT6, TUG1, UCA1, ZEB2AS1) and 3 circRNAs (circ4, circ7, circ16), which expression was evaluated through Real-Time PCR. Our results showed that 3 ncRNAs were differentially expressed in serum exosomes of CRC patients compared to healthy individuals: in particular, UCA1 was downregulated, while circ16 and TUG1 were upregulated (Table 4.2) (Figure 4.4).

	Fold Change	T test p-value
CCAT2		0.88
circ16	**3.82**	**0.036**
circ4		0.42
circ7		0.21
CRNDE		0.61
HOTAIR		0.58
LIT1		0.31
MALAT1		0.35
PCAT1		0.31
PCAT6		0.77
TUG1	**1.69**	**0.029**
UCA1	**-4.22**	**0.03**
ZEB2AS1		0.95

Table 4.2: statistical results of lncRNA and circRNA expression analysis in serum exosomes of 20 CRC patients compared to 20 healthy individuals. For each ncRNA, unpaired T test p-value and fold change are shown.
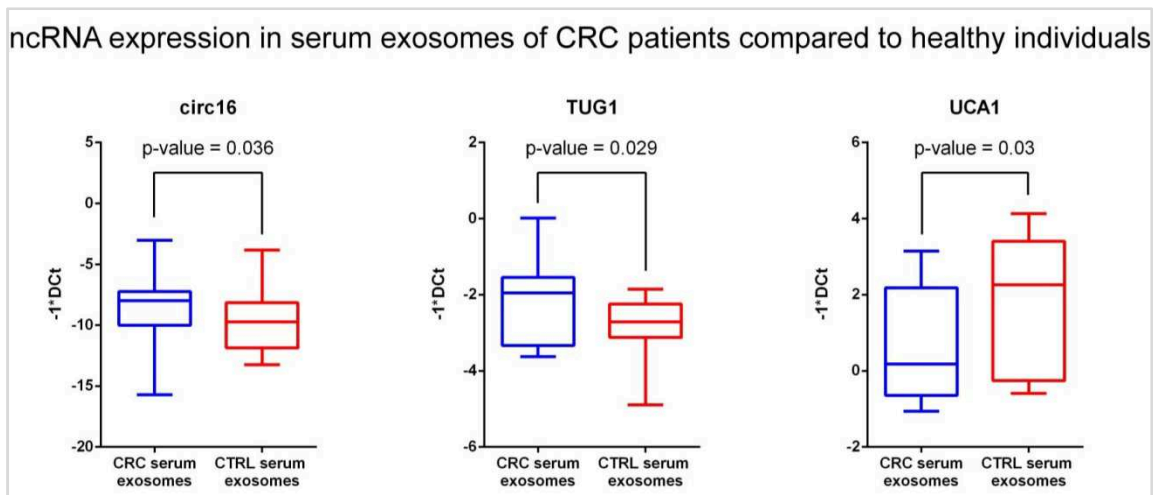
Figure 4.4: box plots showing differential expression of 2 lncRNAs and 1 circRNA in exosomes isolated from serum of 20 CRC patients and 20 healthy individuals.

We computed Pearson's correlation coefficient (r) to evaluate the correlation between TUG1 and UCA1 expression in serum exosomes of all CRC patients and healthy individuals, and in serum exosomes of CRC patients only. The r values showed in both cases a negative correlation, stronger in CRC patient group; T test revealed that this anticorrelation is strongly significant in both sample groups (Figure 4.5).
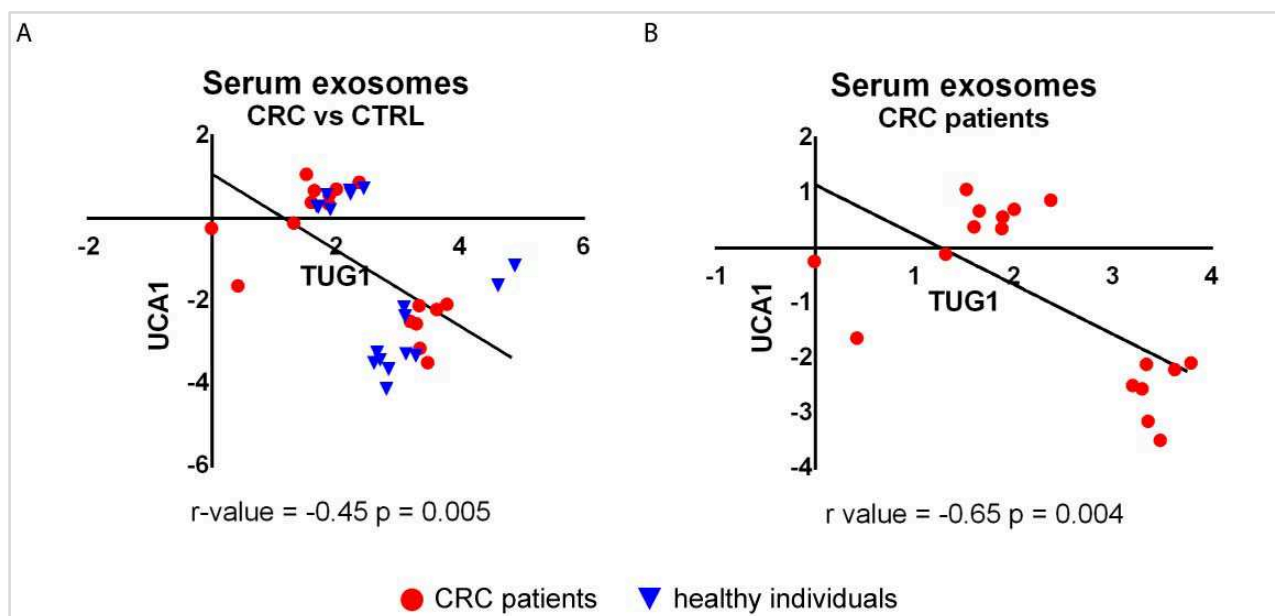
Figure 4.5: Pearson's correlation coefficient (r) showed a negative correlation between TUG1 and UCA1 expression in serum exosomes of CRC patients and healthy individuals (A), and in serum exosomes of CRC patients only (B).

No statistical correlation was observed between UCA1 and circ16 and between TUG1 and circ16.

4.4 Analysis of GEO DataSets and identification of mRNA and miRNA targets of DE ncRNAs

Overall, our study revealed the differential expression of 8 ncRNAs in CRC FFPE biopsies compared to NATs and serum exosomes of CRC patients compared to healthy individuals. To validate our results, we analysed the expression of these ncRNAs in several CRC GEO DataSets; however, we obtained no results for CCAT1, CCAT2, CDR1AS and circ16, because no probe specific for these ncRNAs was retrieved in the analysed datasets. Datasets based on CRC tissues and normal colon mucosa confirmed the same deregulation trends here reported and also revealed a positive correlation between the up- or downregulation of each DE ncRNA and CRC progression and features (Figure 4.6).
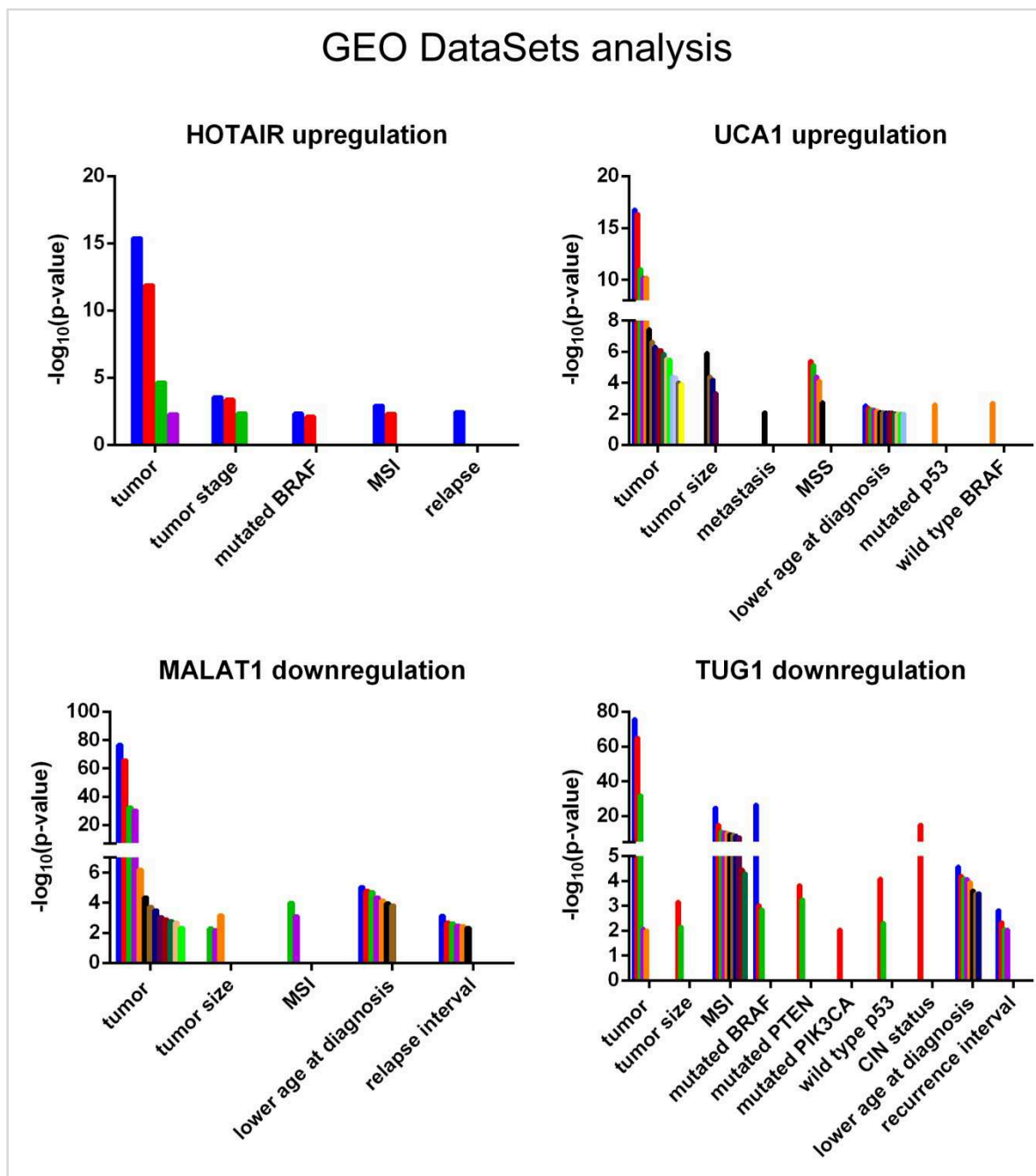
Figure 4.6: GEO DataSets analysis confirmed the upregulation of HOTAIR and UCA1 and the downregulation of MALAT1 and TUG1 in CRC tissues compared to NATs; moreover, the same analysis revealed a positive correlation between ncRNA up- or downregulation and CRC progression; statistical significance is shown as $-\log_{10}$(p-value). Each dataset is represented by a different colour. MSI: microsatellite instability; MSS: microsatellite stability; CIN: chromosomal instability.

In addition, we analysed a dataset based on a study upon peripheral blood of CRC patients compared to healthy individuals. Again, the analysis confirmed the upregulation of TUG1 and the downregulation of UCA1 that we observed in exosomes isolated from serum of CRC patients and

healthy individuals (Figure 4.7a). In particular, the same deregulation trend of both lncRNAs was observed in two different groups of CRC patients based on Dukes' classification: i) Dukes' stages A and B, ii) Dukes' stages C and D. Unfortunately, no clinical data was available to further analyse the association between ncRNA deregulation and CRC. Plus, TUG1 and UCA1 showed a significant negative correlation of expression (r-value=-0.439 p-value=0.02) (Figure 4.7b).
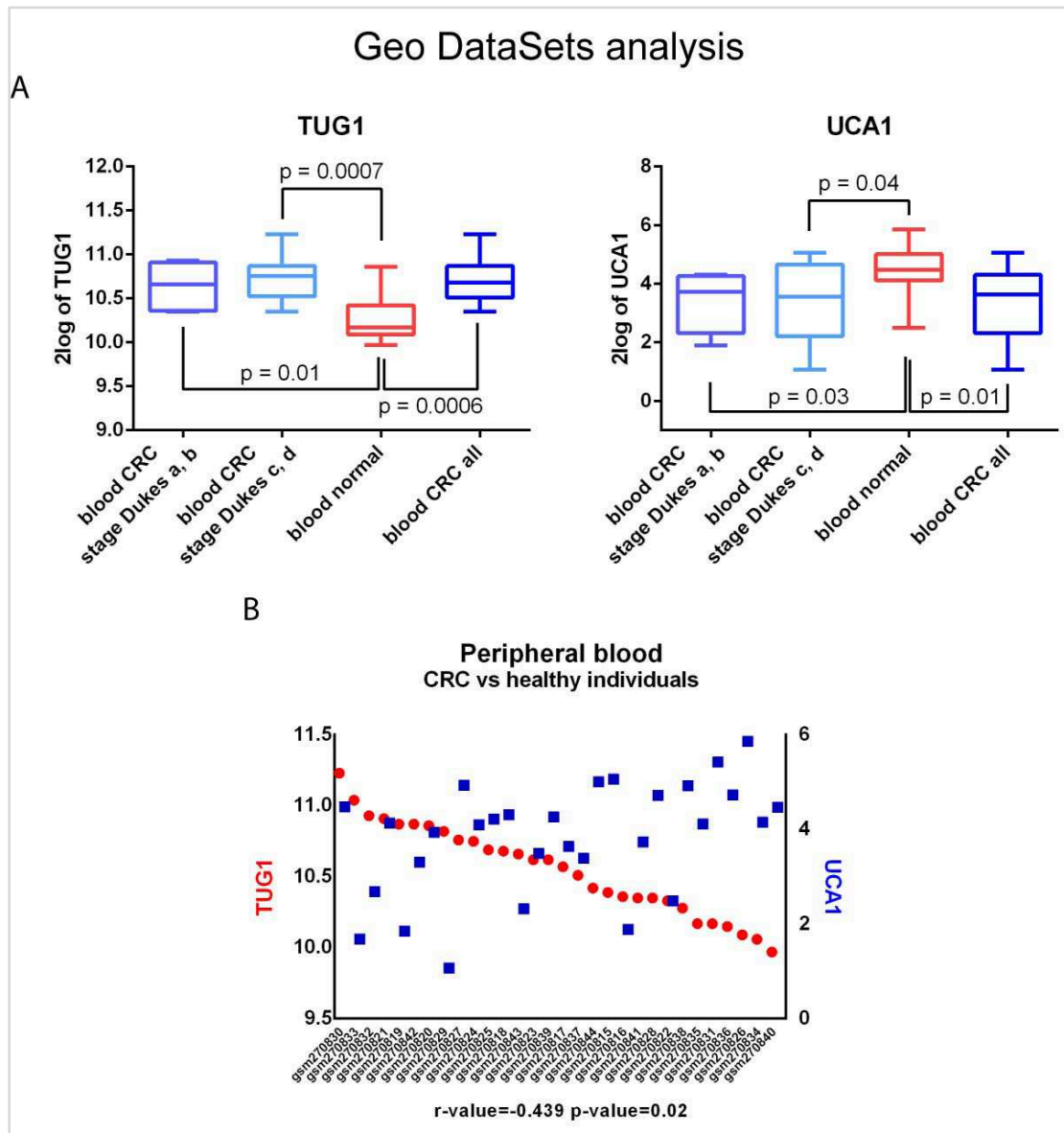
Figure 4.7: analysis of a GEO dataset reporting expression data of peripheral blood of CRC patients compared to healthy individuals (GSE10715). **A)** the dataset confirmed the upregulation of TUG1 and the downregulation of UCA1; the same trend is observed in two different groups of CRC patients, composed by i) Dukes' stages A and B, and ii) Dukes' stages C and D. **B)** TUG1 and UCA1 showed a significant negative correlation of expression.

We identified hypothetical mRNA targets of lncRNAs deregulated in tumor tissues using three databases (lncRNAtor, MEM, BioGPS); however, we obtained data about HOTAIR, MALAT1, TUG1 and UCA1 from all the three databases, while CCAT1 and ZEB2AS1 data were retrieved only from lncRNAtor; no data about CCAT1 were found in any analysed database. Genes with correlation of expression were submitted to cBioPortal database to verify their involvement in CRC; the sequences of selected genes were aligned with each DE lncRNA sequence through Nucleotide BLAST. Also, each lncRNA sequence was aligned with whole genome sequence; the existence of a correlation of expression of selected complementary genes with each DE lncRNA was investigated: i) genes complementary to HOTAIR, MALAT1, TUG1 and UCA1 were submitted to MEM; only genes with p-value inferior or equal to 10^{-6} were selected; ii) genes complementary to CCAT1 and ZEB2AS1 were compared to genes retrieved from lncRNAtor; only common genes were selected. Only the alignment with the whole genome provided data about CCAT2.

We also identified hypothetical miRNA targets of lncRNAs and circRNAs deregulated in tumor tissues through the alignment of mature sequences of miRNAs associated with CRC and the sequence of each DE ncRNA; only miRNAs showing sequence complementarity and negative correlation of expression with DE ncRNAs were selected.

Tables 4.3 and 4.4 show mRNA and miRNA targets selected for each DE ncRNA.

DE ncRNA	mRNA	gene name
CCAT1	ACTL6A	actin-like 6°
	PDZD2	PDZ domain containing 2
CCAT2	MKI67	marker of proliferation Ki-67
	PTPRU	protein tyrosine phosphatase, receptor type, U
	RASEF	RAS and EF-hand domain containing
HOTAIR	EDNRB	endothelin receptor type B
	SRC	SRC proto-oncogene, non-receptor tyrosine kinase
MALAT1	NCOA2	nuclear receptor coactivator 2
	PAX3	paired box 3
	RAC1	ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)

	ZEB2	zinc finger E-box binding homeobox 2
	HIVEP3	human immunodeficiency virus type I enhancer binding protein 3
TUG1	SEC24B	SEC24 homolog B, COPII coat complex component
	SIPA1L3	signal-induced proliferation-associated 1 like 3
	HEY2	hes-related family bHLH transcription factor with YRPW motif 2
UCA1	MACC1	metastasis associated in colon cancer 1
	MMP1	matrix metallopeptidase 1
ZEB2AS1	AIM2	absent in melanoma 2
	ZEB2	zinc finger E-box binding homeobox 2

Table 4.3: selected mRNA targets of DE ncRNAs identified through computational analysis.

DE ncRNA	miRNA
CCAT1	hsa-miR-145-5p
	hsa-miR-34b-5p
CCAT2	hsa-miR-15a-5p
	hsa-miR-320a
CDR1AS	hsa-miR-1246
	hsa-miR-499a-5p
HOTAIR	hsa-miR-148b-3p
	hsa-miR-328-3p
MALAT1	hsa-miR-133a-3p
	hsa-miR-135a-5p
	hsa-miR-1179
	hsa-miR-320a
TUG1	hsa-miR-223-3p
	hsa-miR-424-5p
UCA1	hsa-miR-143-3p
	hsa-miR-18a-5p
	hsa-miR-198

Table 4.4: selected miRNA targets of DE ncRNAs identified through computational analysis.

4.5 Analysis of gene ontologies

Gene ontology and pathway enrichment analyses were performed aiming to understand the effects of lncRNA and circRNA deregulation within tumor cells. We used mRNA and miRNA

targets of DE ncRNAs and their nearest neighbours to build two biological networks; all genes included in these network were then used to perform the analyses. The network based on miRNA targets comprised 3792 nodes and 26870 edges, while the network based on mRNA targets was made of 939 nodes and 12710 edges. Results of gene ontology and pathway enrichment analyses are shown in figure 4.8.
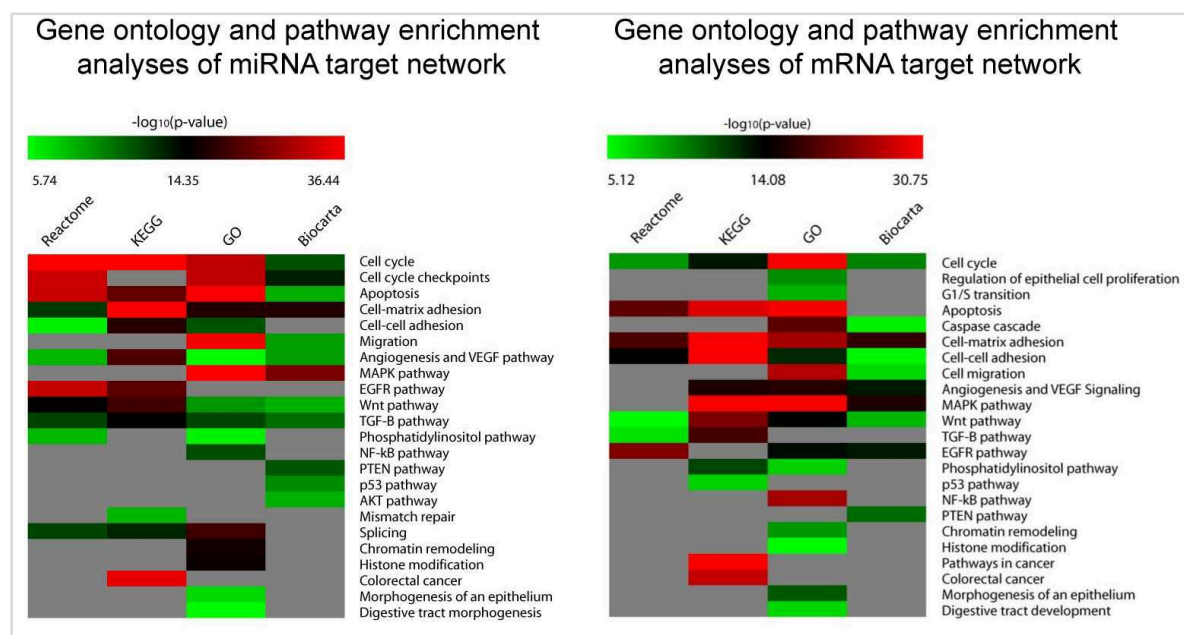
Figure 4.8: pathways regulated by miRNA and mRNA targets of DE ncRNAs in tumor tissues; chromatic scale represents -$\log_{10}$(p-value) of Fisher's exact test.

Gene ontology and pathway enrichment analyses were also performed for ncRNAs deregulated in serum exosomes of CRC patients; we performed different analysis for each DE ncRNA: FatiGO tool was used for circ16 and UCA1, while GeneMania was used for TUG1. Pathways regulated by ncRNAs deregulated in serum exosomes are shown in figure 4.9.
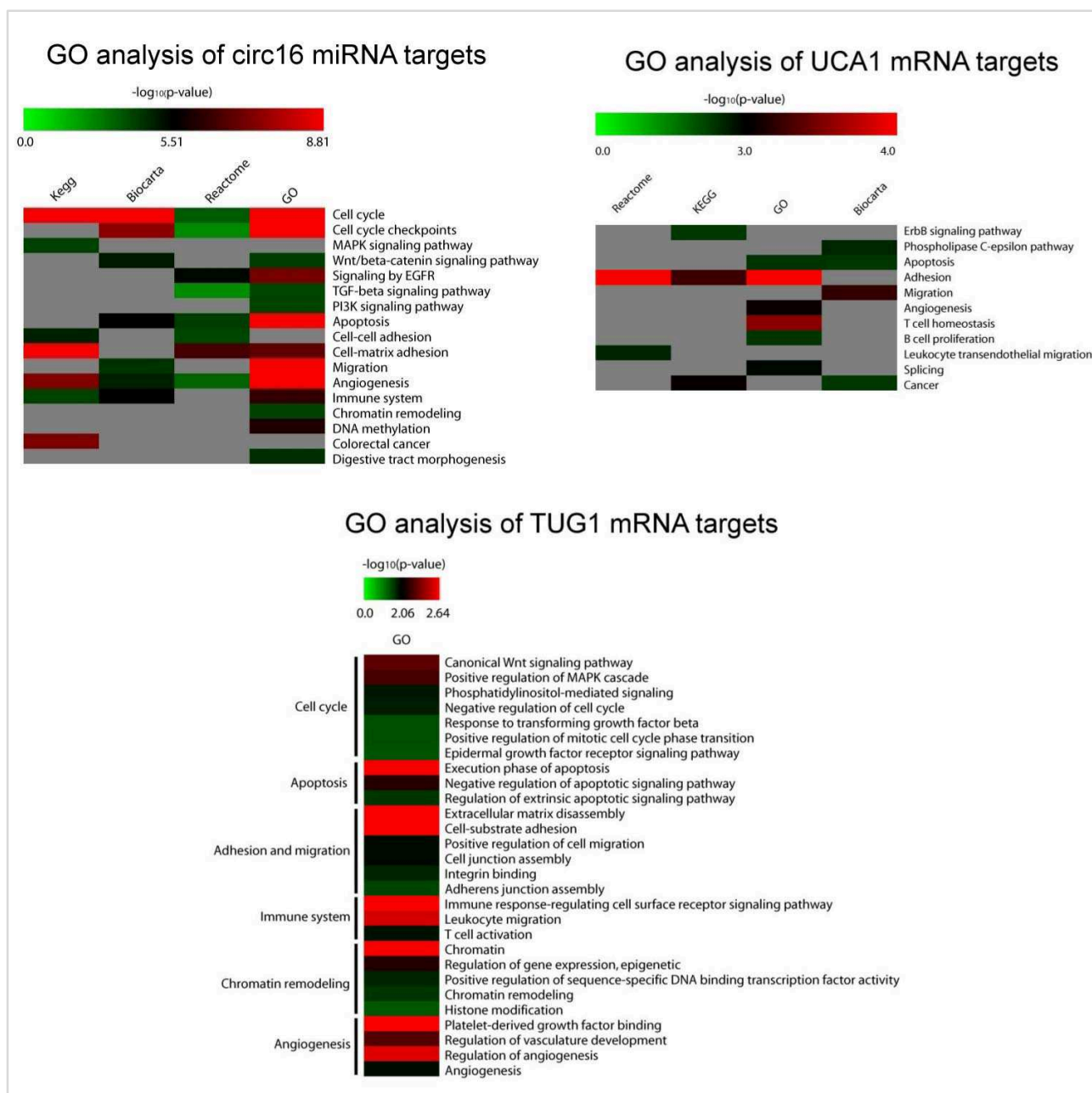
Figure 4.9: pathways regulated by miRNA targets of circ16 and mRNA targets of UCA1 and TUG1; chromatic scale represents -log$_{10}$(p-value) of Fisher's exact test.

4.6 Evaluation of the diagnostic accuracy of DE ncRNAs in serum exosomes

Differentially expressed ncRNAs in serum exosomes of CRC patients compared to healthy individuals represent potential tumor markers; to assess their diagnostic accuracy we computed ROC curves using ΔCt values respect to endogenous control GAPDH. Circ16 ROC curve showed an AUC of 0.771 (95% CIs, 0.508-0.936; p = 0.02) with 71% of sensitivity and 80% of specificity

(ΔCt cut-off value ≤ 8.88) (Figure 4.10a); TUG1 curve showed an AUC of 0.640 (95% CIs, 0.455-0.799) with 56% of sensitivity and 88% of specificity (ΔCt cut-off value > -2.22), but it was not statistically significant (p = 0.2) (Figure 4.10b); UCA1 curve showed an AUC of 0.719 (95% CIs, 0.533-0.863; p = 0.01) with 100% of sensitivity and 43% of specificity (ΔCt cut-off value ≤ 3.14) (Figure 4.10c).
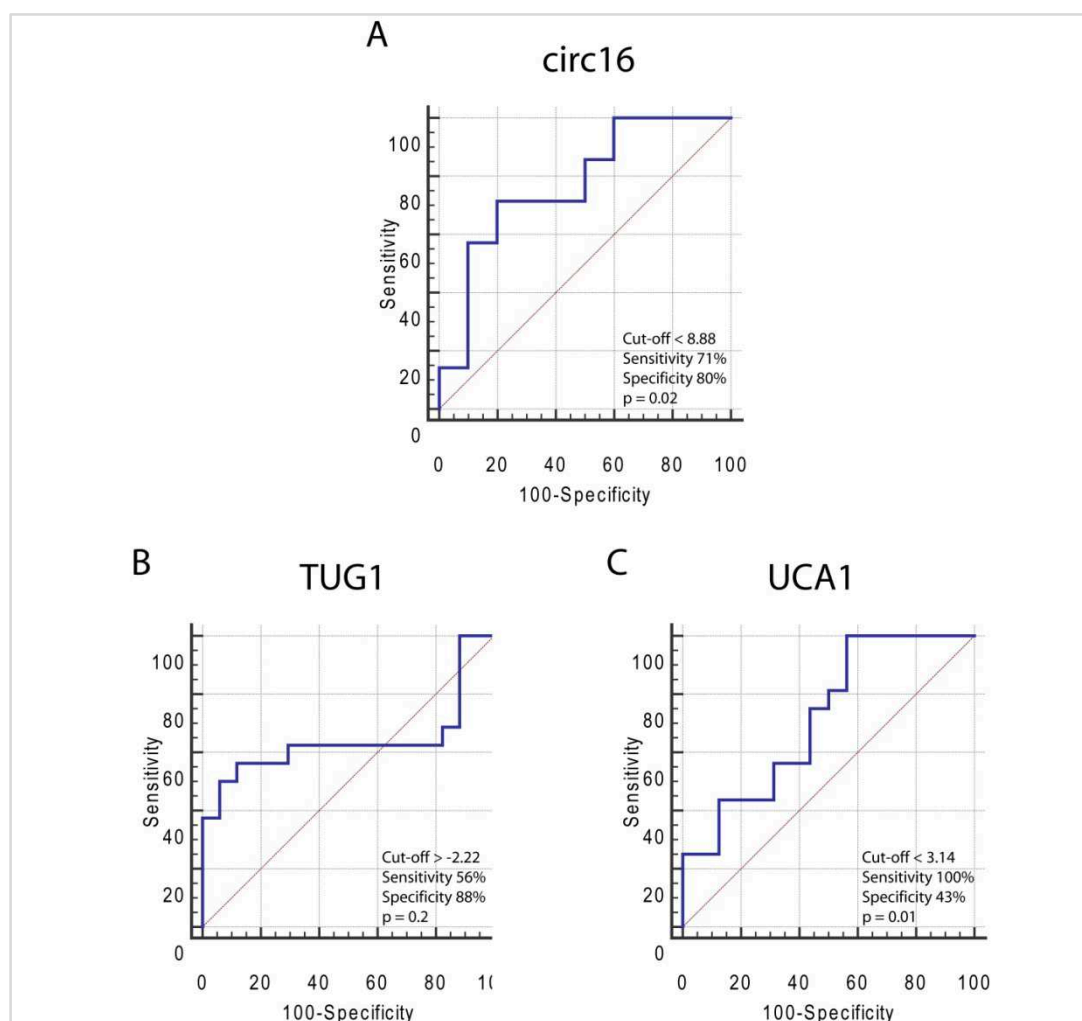
Figure 4.10: ROC curves of DE ncRNAs in serum exosomes of CRC patients compared to healthy individuals.

Although these results derived from a limited sample size and further analyses are needed, they suggest that ncRNAs in serum exosomes could be used as diagnostic markers for CRC with a good diagnostic accuracy.

4.7 *In vitro* inhibition of MAPKs

4.7.1 Cell viability assay

Cell viability was assessed through MTT assay 24 hours after treatment with MAPK inhibitors FR180204 and U0126. Results showed that both inhibitors significantly reduced cell proliferation of treated cells compared to negative controls treated with an equal volume of DMSO (Figure 4.11).
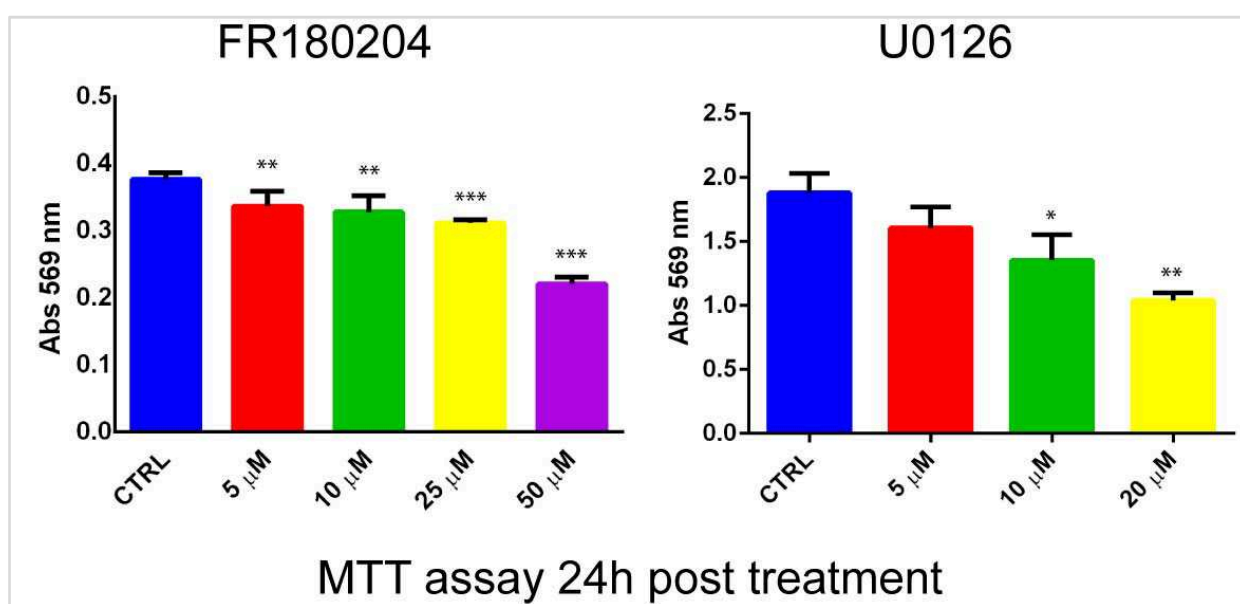
Figure 4.11: treatment with both MAPK inhibitors (FR180204 and U0126) significantly reduced cell viability, measured by MTT assay; cells treated with DMSO were used as controls. *: p-value < 0.05; **: p-value < 0.005; ***: p-value < 0.0005.

4.7.2 Confirmation of MAPK inhibition through Western Blot analysis

Western Blot analysis was performed 24 hours after treatment with MAPK inhibitors FR180204 and U0126 to confirm the inhibition of MAPK pathway. Results showed a significant reduction in phospho-ERK (pERK) levels in treated cells compared to control cells, treated with equal volume of DMSO (Figure 4.12).
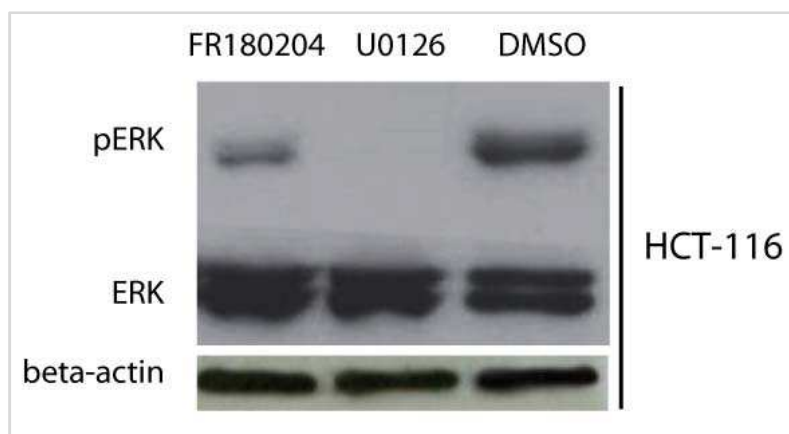
Figure 4.12: Western Blot showing ERK and pERK levels in HCT-116 cells treated with MAPK inhibitors FR180204 and U0126 compared to control cells treated with DMSO: MAPK inhibition is showed by reduced levels of pERK compared to ERK and beta-actin, used as loading control.

4.7.3 MAPK inhibition affects ncRNA expression

The expression of ncRNAs deregulated in CRC tissues and serum exosomes was investigated in HCT-116 cells 24 hours after treatment with U0126. Expression analysis was performed in Real-Time PCR using PPIA as endogenous control. Results showed that the expression of 4 out of 9 ncRNAs was affected by the inhibition of MAPK pathway (Table 4.5). The analysis will be also performed in HCT-116 cells treated with FR180204.

ncRNA	Fold Change	T test p-value
CCAT1		0.71
CCAT2		0.21
CDR1AS		0.31
circ16		0.74
HOTAIR	**-15**	**0.037**
MALAT1	**3.66**	**0.002**
TUG1	**5.85**	**0.00009**
UCA1	**3.14**	**0.033**
ZEB2AS1		0.59

Table 4.5: expression of ncRNAs in HCT-116 cells treated with U0126. For each ncRNA, median fold change and T test p-value are shown.

In particular, U0126 treatment induced the upregulation of MALAT1, TUG1 and UCA1 and the downregulation of HOTAIR (Figure 4.13).
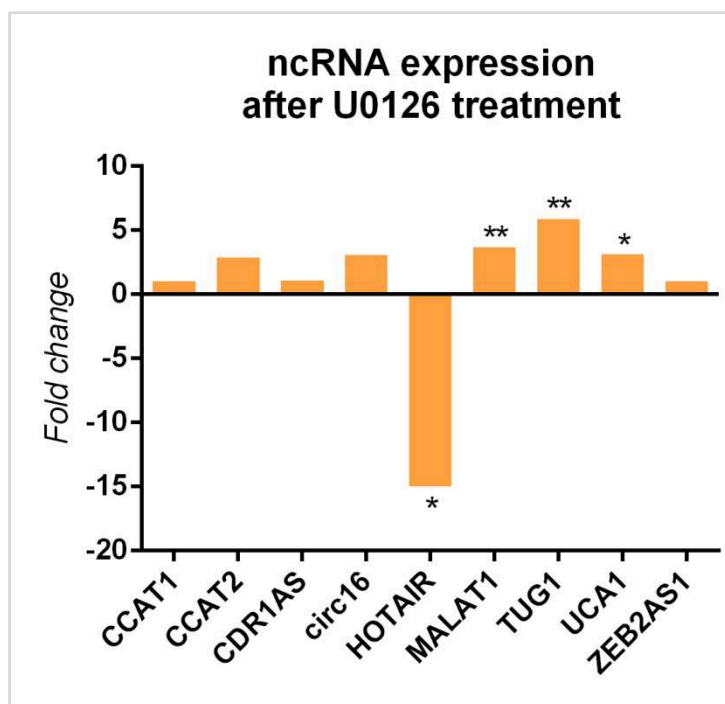
Figure 4.13: ncRNA expression in HCT-116 cells treated with U0126. *: p-value < 0.05; **: p-value < 0.01.

4.7.4 Evaluation of correlation between ncRNA expression and ERK activation

We analysed GEO DataSets to investigate the correlation between the expression of ncRNAs deregulated by MAPK inhibition and the activation of the pathway; the analysis was performed by retrieving from CRC datasets correlation data between each DE ncRNA and HSPA5 [heat shock protein family A (Hsp70) member 5], a transcriptional target of pERK (257). Results showed a positive correlation between HOTAIR and HSPA5, and a negative correlation between MALAT1, TUG1 or UCA1 and HSPA5, confirming our observation (Table 4.6).

	Geo DataSets ID	p-value	r-value
HOTAIR	GSE13294	3.00E-02	0.18
	GSE33114	3.00E-02	0.212
MALAT1	GSE37892	1.10E-08	-0.475
	GSE35896	1.30E-03	-0.4
	GSE8671	7.70E-03	-0.33
	GSE13294	2.70E-03	-0.24
	GSE4554	3.00E-02	-0.233
	GSE33114	2.00E-02	-0.217
TUG1	GSE18088	1.20E-05	-0.561
	GSE13294	4.00E-14	-0.559

	GSE33114	1.90E-09	-0.538
	GSE37892	1.30E-08	-0.473
	GSE35896	2.50E-04	-0.449
	GSE8671	2.20E-03	-0.376
	GSE4554	5.10E-03	-0.303
	GSE28702	5.80E-03	-0.301
	GSE2109	1.70E-07	-0.29
	GSE17538	1.50E-05	-0.28
	GSE18105	3.00E-03	-0.279
	GSE21510	2.90E-03	-0.243
	GSE41258	2.10E-05	-0.213
UCA1	GSE33114	4.40E-03	-0.272
	GSE14333	2.10E-03	-0.18

Table 4.6: analysis of CRC GEO DataSets showing correlation between ncRNAs deregulated by MAPK inhibition and HSPA5, a transcriptional target of pERK. Results confirmed the positive correlation observed between HOTAIR expression and MAPK activation, and the negative correlation between MALAT1, TUG1 or UCA1 and pathway activation.

4.8 UCA1 target expression after *in vitro* silencing

UCA1 transient silencing was achieved through transfection with ASOs; RNA was isolated from treated cells and analysed through Real-Time PCR to evaluate the expression of miRNA and mRNA targets of the lncRNA. Transfection efficiency is shown in table 4.7; only replicates with transfection efficiency higher than 70% were considered for target expression analysis.

	24h			48h		
	TREAT1	**TREAT2**	**TREAT3**	**TREAT1**	**TREAT2**	**TREAT3**
CTRL1	88.91443	95.28384	96.337	6.129288	23.58149	13.52776
CTRL2	51.28108	79.27341	83.90183	3.434258	2.200903	4.986811
CTRL3	28.71566	69.67335	76.44554	21.41815	36.02789	27.61162

Table 4.7: efficiency of UCA1 ASO transfection in HCT-116 cells at 24 and 48 hours.

Results showed that UCA1 transient silencing induced the upregulation of 3 mRNAs (MACC1, NCL, UGDH) and the downregulation of 1 miRNA (miR-328-3p) (Tables 4.8 and 4.9).

	24h		48h	
	Median fold change	**T test p-value**	**Median fold change**	**T test p-value**
AREG		0.72		0.35
ASCL2		0.06		0.89
LINC01764		0.06		0.69
MACC1	**1.74**	**0.04**		0.82
NCL	**2.11**	**0.02**		0.65
PDP1		0.92		0.69
SYNCRIP		0.22		0.44
UGDH	**2.89**	**0.031**		0.87

Table 4.8: expression of UCA1 mRNA targets after transient silencing of the lncRNA. Median fold change and T test

p-value are shown for each target.

	24h		48h	
	Median fold change	**T test p-value**	**Median fold change**	**T test p-value**
hsa-let-7b		0.83		0.59
hsa-miR-106a		0.81		0.28
hsa-miR-145		0.73		0.58
hsa-miR-16		0.36		0.17
hsa-miR-18a		0.56		0.21
hsa-miR-328-3p	**-2.91**	**0.05**		0.72
hsa-miR-339-5p		0.34		0.89

Table 4.9: expression of UCA1 miRNA targets after transient silencing of the lncRNA. Median fold change and T test

p-value are shown for each target.

5. Discussion

Despite their recent discovery, lncRNAs have been intensively studied to understand their functions in cellular physiology and their involvement in pathological processes, including carcinogenesis. The few characterized mechanisms of action suggest that lncRNAs play a crucial role in regulatory cellular networks, modulating fundamental biological processes such as proliferation, gene expression, apoptosis, migration. Less explored but equally fascinating and promising is the field of circRNAs, about which few data are to date available. The characterization of lncRNA and circRNA functions in physiological and pathological conditions may promote the development of new diagnostic and therapeutic approaches.

5.1 I NcRNAs contribute to CRC progression

Through Real-Time PCR we identified 7 lncRNAs (CCAT1, CCAT2, HOTAIR, MALAT1, TUG1, UCA1, ZEB2AS1) and 1 circRNA (CDR1AS) differentially expressed in tumor tissues compared to normal mucosa; in particular, we observed the upregulation of CCAT1, CCAT2, HOTAIR, UCA1 and ZEB2AS1, and the downregulation of CDR1AS, MALAT1 and TUG1.

Several studies have previously reported the upregulation of CCAT1 in CRC tissues compared to normal mucosa (223, 258, 259): CCAT1 levels are increased through all disease stages, and also in liver, peritoneal, and lymph node metastasis (223, 258); high levels of CCAT1 promote proliferation, invasion and migration in CRC cell lines (259). CCAT1 is transcribed from a *locus* near to MYC gene, and it also acts as super-enhancer of this oncogene: the CCAT1-L isoform fosters MYC transcription inducing the formation of a DNA loop that brings MYC promoter near to an enhancer located 335 kb (MYC-335) upstream and to a super-enhancer located 515 kb (MYC-515) upstream of MYC gene (224).

The upregulation of CCAT2 has also been reported by a previous study: higher levels of the lncRNA have been observed in CRC tissues characterized by microsatellite stability compared to CRC tissues with microsatellite instability and normal tissues. CCAT2 is also upregulated in primary tumors of metastatic patients compared to non-metastatic, suggesting a role of CCAT2 in metastatization: CCAT2 upregulation promotes tumor growth, invasion and metastasis both *in vitro* and *in vivo*. The CCAT2 gene is located within the region 8q24, the same of CCAT1 and MYC genes: CCAT2 is retained within the nucleus where it activates MYC transcription, thus regulating its expression and tumorigenic effects. CCAT2-induced expression of MYC could be mediated by TCF7L2, a transcription factor regulating the expression of Wnt target genes, including MYC itself; moreover, CCAT2 itself is a target of Wnt pathway. Therefore, this complex network represents a feedback regulation mechanism (222). CCAT2 upregulation and involvement in tumor progression have also been observed in other tumors, such as breast cancer (260), non-small lung cell carcinoma (261), and esophageal squamous cell carcinoma (262).

Few data are available about CDR1AS in CRC: its function of "molecular sponge" for miR-7 has been demonstrated in brain tissue (80, 91), and miR-7 downregulation has also been associated with CRC (263). However, our results showed the downregulation of CDR1AS in CRC tissues, suggesting a different function of this circRNA in CRC model; CDR1AS downregulation has also been observed in glioblastoma multiforme (264). A recent paper has reported a global downregulation of circRNAs in CRC, except for CDR1AS which showed increased levels (235). Further analyses on larger cohorts will be necessary to fully understand CDR1AS expression and function in CRC.

High HOTAIR levels have been observed in several tumors, including CRC; HOTAIR upregulation has also shown a positive correlation with poor prognosis, EMT and metastasis in CRC patients (229, 265, 266). *In vitro* knockout of HOTAIR affects migration and invasion of CRC cell lines, reduces vimentin and MMP9 metalloprotease expression and increases E-cadherin levels:

these evidences suggest a strong involvement of HOTAIR in metastatization (265). Recently, it has been suggested that HOTAIR contributes to apoptosis block in several tumors: HOTAIR knockdown *in vitro* reduces miR-125a-5p levels, inducing the consequent increase of one of its target, CASP2 (caspase 2), which in turn triggers apoptosis through the mitochondrial pathway; HOTAIR-mediated regulation of miR-125a-5p levels is based on epigenetic mechanisms modulated by EZH2 (enhancer of zeste 2 polycomb repressive complex 2 subunit) and LSD1, cofactors of the lncRNA in chromatin remodeling (267). A key role of HOTAIR has been reported in cancer stem cells (CSCs): TGF-β treatment of CD133[+] CSCs increases HOTAIR expression and triggers EMT (268); plus, HOTAIR knockdown impairs proliferation, migration, invasion and colony forming capacity in CD133[+] CSCs (268, 269).

Recent studies have reported the upregulation of MALAT1 in CRC primary tumors compared to normal mucosa and its association with metastasis, suggesting a possible application of MALAT1 as prognostic marker: MALAT1 levels are higher in tumor tissues of metastatic patients compared to non-metastatic and to normal mucosa (270, 271). Our analysis showed a reduced expression of MALAT1 in CRC tissues compared to normal adjacent tissues; similar evidences have been reported in other tumors. MALAT1 downregulation has been observed in glioma compared to normal brain tissues; *in vitro* silencing of MALAT1 promotes cell proliferation and invasion, increasing MKI67 (marker of proliferation Ki-67) and MMP2 (matrix metallopeptidase 2) expression respectively, while the overexpression of the lncRNA provokes cell cycle block at G0/G1 phase; similar results were obtained through *in vivo* experiments. Interestingly, MALAT1 upregulation *in vitro* reduces pERK levels, suppressing MAPK pathway, while low MALAT1 levels activate the same pathway (272), which is crucial in CRC carcinogenesis. MALAT1 downregulation has also been observed in breast cancer: low expression of the lncRNA is associated with axillary lymph node metastasis and shorter relapse-free survival. *In vitro* knockdown of MALAT1 promotes migration and invasion and triggers EMT through PI3K/AKT

pathway (273). Reduced levels of MALAT1 here reported differ from data available in literature about CRC, where MALAT1 upregulation is associated with poor prognosis and metastasis; to date, we know that only 3 out of 20 patients analysed developed lymph node metastases, and this limited number could explain this discordance. Moreover, it is noteworthy that literature reports a very complex *scenario* about MALAT1 expression in several tumors and that lncRNA mechanisms of action have not been fully understood yet; therefore, the downregulation observed in glioma and breast cancer, where MALAT1 participates in signaling pathways classically associated with colorectal cancer, could occur in CRC too. In addition, our analysis of several CRC GEO DataSets confirmed the downregulation of MALAT1 in CRC tissues compared to normal mucosa, and its association with CRC progression and features. Further studies will be necessary to understand MALAT1 involvement in carcinogenesis of CRC and other tumors.

Similarly, we reported the downregulation of TUG1 in CRC patients, while recent studies have reported the its upregulation and association with poor prognosis, EMT and metastasis in CRC (274, 275). The tumor suppressor activity of TUG1 has been proved in glioma and non-small lung cell cancer, where TUG1 has shown reduced levels. In glioma the downregulation of TUG1 is associated with tumor stage and poor prognosis; TUG1 overexpression in glioma cell lines significantly increases the percentage of apoptotic cells through caspase activation and Bcl-2 downregulation, while TUG1 knockdown promotes cell proliferation (276). Also in non-small lung cell cancer reduced levels of TUG1 are associated with poor prognosis, tumor stage and size; *in vitro* knockdown of TUG1 increases cell proliferation and colony forming capacity. In addition, TUG1 is a p53 target which preferentially accumulates within the nucleus, where it interacts with PRC2, thus epigenetically modulating the expression of HOXB7 (homeobox B7), a key transcription factor in cell proliferation belonging to HOXB gene cluster (277); HOXB7 promotes cell proliferation through AKT and MAPK pathways, also in CRC (278). Again, our data on TUG1 differs from CRC literature; this discordance could be explained by the limited cohort with a small

number of metastatic patients; again, GEO DataSets analysis confirmed the downregulation of TUG1 in CRC tissues compared to normal adjacent mucosa. Therefore, further analyses will be necessary.

The upregulation of UCA1 (urothelial cancer associated 1) here reported agrees with literature data. UCA1 shows higher levels in tumor tissues compared to normal mucosa, and its levels increase with tumor stage; UCA1 expression is also increased in tumor tissues of patients with lymph node metastasis compared to non-metastatic (279). UCA1 upregulation correlates with tumor stage, poor prognosis and metastasis; *in vitro* knockdown of UCA1 inhibits cell proliferation, invasion and migration and promotes apoptosis in CRC cell lines (280). High levels of UCA1 are also associated with resistance to 5-fluorouracil treatment through the inhibition of apoptosis (281).

No literature data are available about ZEB2AS1 in CRC or in other tumors compared to normal tissues. Recently it has been proposed a mechanism according to which the antisense transcript of ZEB2 gene regulates the translation of ZEB2 mRNA in several cell lines, including CRC cells. In epithelial cells, the main promoter of ZEB2 gene produces a transcript characterized by a 3-kb 5'-UTR region; after the removal of a 2.5-kb intron, only 481 nucleotides of the 5'-UTR region are kept. Within this residual 5'-UTR region lies a sequence that impairs ribosome scanning of the transcript and provokes its release, thus inhibiting ZEB2 mRNA translation. In mesenchymal cells, a second downstream promoter is activated, generating a greater amount of antisense transcript compared to ZEB2 mRNA; the antisense transcript binds the mRNA and masks its splicing sites, thus inhibiting intron removal; within the retained intron lies an internal ribosome entry site (IRES), close to the transcription start site: the antisense transcript thus modulates ZEB2 mRNA splicing, promoting its translation (20) (see section 2.1.2, figure 2.8). Therefore, the upregulation of ZEB2AS1 induces increased levels of Zeb2 protein. ZEB2 upregulation has been observed in primary tumor and metastases of CRC patients, and at invasion front compared to tumor centre; high levels of ZEB2 protein have also been associated with tumor stage and reduced

survival. *In vitro* experiments have proved that ZEB2 silencing reduces migration and invasion in CRC cell lines (282). Several studies have also demonstrated that some ZEB2-targeting miRNAs are downregulated in CRC (283, 284, 285).

It is clear that ncRNAs deregulated in CRC tissues contribute to carcinogenesis in several tumors, including CRC. Gene ontology and pathway enrichment analyses performed on the 8 ncRNAs deregulated in CRC tissues showed the involvement of these molecules and of their miRNA and mRNA targets in numerous processes associated with carcinogenesis, such as cell cycle and apoptosis regulation, adhesion and migration, angiogenesis; ncRNA deregulation also affects pathways classically associated with CRC, such as MAPK, Wnt, EGFR, PI3K, and mismatch repair pathways, and processes regulated by lncRNAs, such as splicing, chromatin remodeling and histone modifications, all associated with tumor progression.

Hence, both literature data and gene ontology analysis suggest a strong involvement of the DE ncRNAs in CRC progression.

5.2 NcRNAs in serum exosomes of CRC patients: new diagnostic biomarkers?

Expression analysis of ncRNAs in serum exosomes of CRC patients compared to healthy individuals showed altered expression of 2 lncRNAs (TUG1 and UCA1) and 1 circRNA (circ16); in particular, we observed the upregulation of circ16 and TUG1, and the downregulation of UCA1.

The presence of circ16 in serum exosomes has been previously reported (286); circ16, identified in circBase as hsa_circ_0000284, is a 1099 bp circular transcript produced from the second exon HIPK3 (homeodomain interacting protein kinase 3) gene, which is flanked by two long introns in the linear transcript. HIPK3 gene codes for a 170-kDa protein kinase which phosphorylates FADD (Fas associated via death domain) thus increasing resistance of prostate carcinoma cells to Fas-mediated apoptosis (287). Circ16 expression has been investigated in several cancer cell lines, including CRC cells, showing higher levels compared to the linear HIPK3

transcripts and a typically cytoplasmic localization; it has been demonstrated that circ16 knockdown significantly reduces cell growth in cell lines of several cancer histotypes, including CRC; in addition, this circRNA has Ago2-binding sites and acts as miRNA sponge for some miRNAs acting as tumor suppressor in several tumors (288); among them, miR-124, miR-193a, miR-29b, and miR-338 are downregulated in CRC (http://mircancer.ecu.edu/).

We reported for the first time the presence of TUG1 within exosomes isolated from human serum; previously, high TUG1 levels have been shown in plasma of patients affected by chronic lymphocytic leukemia and multiple myeloma (289), and in serum of multiple sclerosis patients (290); also, it has been reported that TUG1 is upregulated in serum of osteosarcoma patients, and that its levels significantly decrease after surgical resection of the tumor (291). Interestingly, our data showed an opposite deregulation trend in CRC tissues and in serum exosomes for TUG1, which is downregulated within tumor cells and upregulated in serum exosomes: this observation could suggest that tumor cells protect themselves from TUG1 tumor suppressive activity by secreting it through exosomes. Similar results were previously reported by our research group for miRNAs: we observed that exosomes secreted by Caco-2 and HCT-116 cells were enriched in tumor suppressive miRNAs, involved in blocking proliferation and immune escape (170).

Several studies have demonstrated the presence of UCA1 in different biological fluids, such as urine, gastric juice, plasma and serum of oncologic patients, but no one ever reported UCA1 presence within serum exosomes. High levels of UCA1 have been observed in plasma of gastric cancer patients (292) and in serum of hepatocellular carcinoma patients (293). Moreover, a signature of three lncRNAs (UCA1/CUDR, LSINCT-5 and PTENP1) was identified as diagnostic marker in serum of patients affected by gastric cancer compared to healthy individuals (294). UCA1 upregulation has been reported in plasma of CRC patients; the same study has shown that its levels decreased 14 days after tumor resection (279). It is noteworthy that UCA1 showed asymmetric distribution between tumor tissues, where the lncRNA is upregulated, and serum exosomes, where

it is downregulated; a similar observation was reported in this study for TUG1, which showed an opposite asymmetry. In this case, this observation could suggest that the oncogenic function of UCA1 is crucial for tumor progression, inducing tumor cells to retain the lncRNA by limiting its secretion through exosomes.

Hence, literature is enriched of evidences supporting the involvement of these ncRNAs in tumor progression and their possible application as disease biomarkers. We evaluated the diagnostic accuracy of DE ncRNAs through ROC curves: our data showed good accuracy for 2 out of 3 DE ncRNAs (circ16: 71% sensitivity, 80% specificity; UCA1: 100% sensitivity, 43% specificity), with significant p-values. On the contrary, TUG1 ROC curve showed no statistical significance, likely because of the small number of recruited patients. Surely further analyses will be necessary to evaluate the diagnostic accuracy of these molecules as CRC biomarkers.

Gene ontology and pathway enrichment analyses confirmed that ncRNAs deregulated in serum exosomes are involved in cancer-related processes and in pathways associated with CRC onset and progression. It has been shown that exosomes secreted by tumor cells can target several cell histotypes, including immune cells; exosomal pro-cancer effect may be accomplished by ncRNAs, as suggested by their involvement in pathways related to immune escape and suppression of immune response.

5.3 Identification of miRNA and mRNA targets of DE ncRNAs: future perspectives for the understanding of miRNA-ncRNA and mRNA-ncRNA interaction effects

Several evidences have proved that lncRNAs and circRNAs act through a plethora of molecular mechanisms to regulate various and fundamental cellular processes; on the contrary, miRNAs perform their function of post-transcriptional repressors through a unique mechanism, widely studied, which provokes mRNA target degradation or translation block. We wondered what effects could derive from lncRNA/circRNA-miRNA or lncRNA-mRNA interactions. According to

miRNA degradative function on mRNAs, we could hypothesize that the interaction between ncRNAs and miRNAs induces ribonuclease activation, with the consequent degradation of the ncRNA or the miRNA: it is likely that the miRNA, together with RISC complex and Ago proteins, binds the ncRNA as an mRNA target, causing its degradation; alternatively, the ncRNA could decrease miRNA levels sequestering it or inducing its degradation. Effects of lncRNA-mRNA interactions also need to be investigated: it has been demonstrated that this interaction may be involved in splicing regulation, but other effects may occur.

We performed a computational analysis to identify miRNA and mRNA targets of ncRNA deregulated in tumor tissues aiming to evaluate the effects of their interactions: we foresee to perform functional assays to silence lncRNAs or their mRNA targets (by using siRNA) in CRC cell lines aiming to evaluate the effects on mRNA and lncRNA expression respectively; also, the same approach will be used for miRNAs, modulating their levels with miRNA mimics/inhibitor and analysing the effects on lncRNA and circRNA expression.

Interestingly, no miRNA satisfying our selection criteria was identified for ZEB2AS1; this lncRNA acts as splicing regulator for ZEB2 mRNA, modulating ZEB2 protein levels. Therefore, we hypothesized that ZEB2AS1 doesn't act as miRNA sponge, as suggested by the lack of miRNA-binding sites within its sequence.

Several studies investigated the effect of miRNA-lncRNA/circRNA and mRNA-lncRNA interactions reporting many possible effects, including the degradation of one of the two ncRNAs: miR-9 binds MALAT1 causing its Ago2-mediated degradation (295), as observed for miR-671 and CDR1AS (296); the lncRNA HULC (hepatocellular carcinoma up-regulated long non-coding RNA) sequesters miR-372, reducing its activity and its expression post-transcriptionally (297); HULC is also destabilized by the interaction with IGF2BP1 (insulin like growth factor 2 mRNA binding protein 1), which recruits the cytoplasmic RNA decay machinery (298); GAS5 (growth arrest specific 5) binds c-Myc mRNA impairing its association with polysomes (299), while GHET1

(gastric carcinoma proliferation enhancing transcript 1) stabilizes the interaction between c-Myc mRNA and IGF2BP1, increasing mRNA stability and translation (300); gadd7 (growth-arrested DNA damage-inducible gene 7) interacts with TDP-43 (TAR DNA binding protein), thus causing the degradation of CDK6 (cyclin-dependent kinase 6) mRNA (301).

5.4 Altered expression of ncRNAs after treatment with MAPK inhibitors

MAPK pathway is crucial in CRC onset and progression; it has been observed that CRC patients frequently bare point mutations in KRAS, the G protein acting upstream the pathway, causing its constitutive activation and the consequent iperactivation of signal transduction.

We investigated the effects of MAPK pathway inhibition on DE ncRNA expression; we treated with two inhibitors HCT-116 cells, characterized by mutated KRAS and MAPK iperactivation. We used FR180204, ERK1/2 inhibitor with weak activity on p38α, and U0126, MEK1/2 specific inhibitor.

As showed by MTT assay and Western Blot, both inhibitors reduced cell viability and pERK levels in treated cells compared to control cells. Therefore, both inhibitors showed their efficiency in pathway inhibition.

Expression analysis was performed through Real-Time PCR; results showed that U0126 treatment altered the expression of 4 out of 9 ncRNAs (HOTAIR, MALAT1, TUG1 and UCA1); in particular, MALAT1, TUG1 and UCA1 were upregulated, while HOTAIR was strongly downregulated.

The association between HOTAIR and MAPK pathway has been previously reported in breast cancer: cells treated with SB203580, a specific inhibitor of p38/MAPK14 (mitogen-activated protein kinase 14), showed reduced expression of HOTAIR, suggesting an involvement of p38 in the regulation of its expression (302); our data also showed a strong downregulation of HOTAIR after MAPK inhibition, supporting the same hypothesis. A few evidences suggested that MALAT1

may be involved in MAPK pathway activation, although two different studies reported opposite roles of this lncRNA in glioma and gallbladder cancer cell lines: MALAT1 acts as tumor suppressor in glioma, and its upregulation causes reduced pERK levels, suppressing MAPK pathway (272); on the contrary, in gallbladder cancer, where MALAT1 acts as oncogene, its knockdown inactivates the pathway (303); therefore, despite the connection between MALAT1 and MAPKs seems to be a fact, further analysis will be necessary to understand the molecular mechanism. Plus, our data showed that MAPK inhibition induced increased expression of MALAT1, suggesting the involvement of the pathway in the regulation of MALAT1 expression. TUG1 has been suggested as MAPK activator in non-small cell lung cancer cells: the lncRNA epigenetically regulates HOXB7 expression, which in turn activates AKT and MAPK pathways; conversely, TUG1 knockdown induced increased levels of pERK (277). Our data confirmed the negative correlation between TUG1 and pERK levels, suggesting that MAPKs may regulate TUG1 expression in a feedback mechanism. A recent study reported that UCA1 knockdown significantly reduced pERK levels, but did not affect p38 phosphorylation, in hepatocarcinoma cell lines; UCA1 sequestered miR-216b, increasing FGFR1 (fibroblast growth factor receptor 1) protein levels and thus inducing ERK activation (304). However, our data showed a negative correlation between UCA1 levels and MAPK activation, suggesting that UCA1 expression is not regulated by MAPK pathway; we could speculate that UCA1 significant upregulation after MAPK inhibition is due to increased stability, and thus reduced degradation, of the lncRNA.

Interestingly, U0126 treatment induced an opposite alteration of HOTAIR, MALAT1 and TUG1 expression compared to that observed in CRC tumor tissues: HOTAIR, upregulated in CRC, was downregulated by U0126, while MALAT1 and TUG1, downregulated in tumor tissues, were upregulated after MAPK inhibition; on the contrary, UCA1 upregulation occurred both in tumor tissues compared to normal mucosa and in cells treated with U0126 compared to controls. These data suggest an involvement of HOTAIR, MALAT1 and TUG1 in cell cycle regulation mediated by

MAPKs; instead, UCA1 may perform its oncogenic function through other pathways. Moreover, our data confirmed the anticorrelation reported in literature between MALAT1 or TUG1 knockdown and MAPK pathway activation: the inhibition of the pathway caused increased expression of both lncRNAs, suggesting a feedback regulation mechanism between these lncRNAs and a still unknown member of the MAPK pathway.

5.5 *In vitro* silencing of UCA1 affects the expression of its targets

Our results showed that UCA1 was upregulated in tumor cells and downregulated in serum exosomes of CRC patients; because of this asymmetric distribution, we chose to focus our attention on this lncRNA to perform *in vitro* functional assays. Through ASO technology, we performed UCA1 transient silencing in HCT-116 cells aiming to investigate the effects of lncRNA downregulation on the expression of its targets; among them, we analysed miRNAs and mRNAs characterized by positive or negative correlation of expression with UCA1, retrieved by computational analysis and literature data. Also, we analysed the expression of LINC01764, a ncRNA transcribed from the antisense strand of UCA1 gene, hypothesizing that the two ncRNAs could be involved in a reciprocal mechanism of expression regulation. Results showed the upregulation of 3 mRNA targets (MACC1, NCL, UGDH) and the downregulation of 1 miRNA target (miR-328-3p).

Several studies reported the upregulation of MACC1 (metastasis associated in colon cancer 1) in CRC tissues and its association with metastatization, suggesting its application as a prognostic marker (305, 306, 307). MACC1 acts as transcriptional factor inducing the expression of MET (MET proto-oncogene, receptor tyrosine kinase), the hepatocyte growth factor (HGF) receptor; thus, MACC1 participates in various signal transduction pathways, including MAPK pathway, and regulates proliferation, migration and metastasis in CRC cells (305, 308). NCL (nucleolin) is an abundant non-ribosomal nucleolar protein, crucial in many steps of ribosome biogenesis; it is also

located in nucleus and cytoplasm, and on cell membrane. The localization of NCL within cells is associated with different functions, performed through the interaction with nucleic acids (both RNA and DNA): i) the nuclear NCL pool controls the expression of several oncogenes, genome stability and splicing; ii) within the cytoplasm, NCL binds various mRNAs coding for proteins involved in cell proliferation and apoptosis, thus regulating their translation; among NCL target there is also p53, which translation is inhibited; iii) on cell surface, NCL interacts with several ligands regulating proliferation, apoptosis and angiogenesis, such as Fas, Ras, HGF, VEGF (309, 310). Several studies recently reported NCL deregulation in many tumors, including its upregulation in CRC (311). UGDH (UDP-glucose 6-dehydrogenase) catalyzes the conversion of uridine diphosphate glucose (UDP-glucose) to uridine diphosphate glucuronic acid (UDP-glucuronic acid), which is a key precursor for the synthesis of glycosaminoglycans, common components of extracellular matrix. Among glycosaminoglycans, hyaluronic acid (HA) is involved in physiological cell growth and migration, but its increased levels have been associated with many tumors; including CRC, where it is an unfavorable prognostic factor (312). The incremented synthesis of HA is induced by UGDH upregulation: indeed, UGDH knockdown affects HA production, resulting in reduced motility of colon cancer cells (313). Among miRNAs, miR-328-3p showed reduced levels in CRC tissues and correlation with patient survival (314, 315). Experiment on CRC side population cells showed that miR-328-3p is critical in the maintenance of cancer stem-like phenotype of SP cells: the upregulation of the miRNA impairs drug resistance and invasion. Among its target, MMP16 (matrix metallopeptidase 16) plays a crucial role in metastatization (315). MACC1, NCL and UGDH all showed increased expression after UCA1 transient silencing; miR-328-3p levels were decreased.

However, it has been recently shown that lncRNA cellular localization affects the silencing efficiency of different methods: indeed, ASOs showed to be more suitable for nuclear lncRNA silencing, while cytoplasmic lncRNAs were more effectively silenced by RNAi (316). For this

reason, we will confirm these preliminary results with further experiments, performed by using siRNAs to achieve transient silencing of UCA1 in HCT-116 cells.

6. Conclusions and future perspectives

The onset and progression of a tumor require mutations causing inactivation or aberrant expression of various molecules; in recent years it has become clear that these molecules are represented not only by proteins, but also by non-coding RNAs. Research on ncRNAs, already advanced in the case of miRNAs but still at the beginning for lncRNAs and circRNAs, proved that altered expression of specific molecules can be associated with one or more tumors, promoting their progression through several molecular mechanisms. To understand these mechanisms and evaluate the possible application of ncRNAs as diagnostic markers or new therapeutic targets it will be necessary to investigate the involvement of these molecules in carcinogenesis.

Our study showed the aberrant expression of 8 ncRNAs (CCAT1, CCAT2, CDR1AS, HOTAIR, MALAT1, TUG1, UCA1 and ZEB2AS1) in CRC tumor tissues compared to normal mucosa and their involvement in cancer-related pathways; our results also suggested a connection between MAPK pathway, altered in CRC, and some of these ncRNAs (HOTAIR, MALAT1, TUG1 and UCA1) after the inhibition of MAPKs in HCT-116 cells by U0126 treatment. These data will be confirmed in the same cells treated with FR180204. We also demonstrated the deregulation of 3 ncRNAs (circ16, TUG1 and UCA1) in exosomes isolated from serum of CRC patients compared to healthy individuals, and evaluated their possible application as diagnostic biomarkers through ROC curve computation. Finally, through ASO-mediated silencing of UCA1 in HCT116, we investigated the effects of ncRNA-miRNA/mRNA interactions, aiming to understand if these interactions trigger degradation mechanisms of any of the interacting molecules. Preliminary results showed that MACC1, NCL, UGDH were upregulated by UCA1 silencing, while miR-328-3p was downregulated.

7. Bibliography

1. Crick FH. *On protein synthesis.* **Symp Soc Exp Biol.** 1958; 12:138-63.

2. Jacob F, Monod J. *Genetic regulatory mechanisms in the synthesis of proteins.* **J Mol Biol.** 1961; 3:318-56.

3. International Human Genome Sequencing Consortium. *Initial sequencing and analysis of the human genome.* **Nature.** 2001; 409:860-921.

4. International Human Genome Mapping Consortium. *A physical map of the human genome.* **Nature.** 2001; 409:934-41.

5. International Human Genome Sequencing Consortium. *Finishing the euchromatic sequence of the human genome.* **Nature.** 2004; 431:931-45.

6. Ezkurdia I, Juan D, Rodriguez JM, Frankish A, Diekhans M, Harrow J, Vazquez J, Valencia A, Tress ML. *Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes.* **Hum Mol Genet.** 2014; 23:5866-78.

7. Pertea M, Salzberg SL. *Between a chicken and a grape: estimating the number of human genes.* **Genome Biol.** 2010; 11:206.

8. Sana J, Faltejskova P, Svoboda M, Slaby O. *Novel classes of non-coding RNAs and cancer.* **J Transl Med.** 2012; 10:103.

9. ENCODE Project Consortium. *The ENCODE (ENCyclopedia Of DNA Elements) Project.* **Science.** 2004; 306:636-40.

10. ENCODE Project Consortium. *Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project.* **Nature.** 2007; 447:799-816.

11. Hangauer MJ, Vaughn IW, McManus MT. *Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs.* **PLoS Genet.** 2013; 9:e1003569.

12. Jacquier A. *The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs.* **Nat Rev Genet.** 2009; 10:833-44.

13. Kristiaan Vanhercke. *The power of RNA: discover next generation biomarkers and therapeutic targets.* www.biogazelle.com.

14. Mercer TR, Dinger ME, Mattick JS. *Long non-coding RNAs: insights into functions.* **Nat Rev Genet.** 2009; 10:155-9.

15. Ragusa M, Barbagallo C, Statello L, Condorelli AG, Battaglia R, Tamburello L, Barbagallo D, Di Pietro C, Purrello M. *Non-coding landscapes of colorectal cancer.* **World J Gastroenterol.** 2015; 21:11709-39.

16. Lee JT. *Epigenetic regulation by long noncoding RNAs.* **Science.** 2012; 338:1435-9.

17. Morris KV, Mattick JS. *The rise of regulatory RNA.* **Nat Rev Genet.** 2014; 15:423-37.

18. Popadin K, Gutierrez-Arcelus M, Dermitzakis ET, Antonarakis SE. *Genetic and epigenetic regulation of human lincRNA gene expression.* **Am J Hum Genet.** 2013; 93:1015-26.

19. Tahira AC, Kubrusly MS, Faria MF, Dazzani B, Fonseca RS, Maracaja-Coutinho V, Verjovski-Almeida S, Machado MC, Reis EM. *Long noncoding intronic RNAs are differentially expressed in primary and metastatic pancreatic cancer.* **Mol Cancer.** 2011; 10:141.

20. Beltran M, Puig I, Peña C, García JM, Alvarez AB, Peña R, Bonilla F, de Herreros AG. *A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition.* **Genes Dev.** 2008; 22:756-69.

21. Song X, Wang X, Arai S, Kurokawa R. *Promoter-associated noncoding RNA from the CCND1 promoter.* **Methods Mol Biol.** 2012; 809:609-22.

22. Chan WL, Chang JG. *Pseudogene-derived endogenous siRNAs and their function.* **Methods Mol Biol.** 2014; 1167:227-39.

23. Korneev SA, Korneeva EI, Lagarkova MA, Kiselev SL, Critchley G, O'Shea M. *Novel noncoding antisense RNA transcribed from human anti-NOS2A locus is differentially regulated during neuronal differentiation of embryonic stem cells.* **RNA.** 2008; 14:2030-7.

24. Peng JC, Shen J, Ran ZH. *Transcribed ultraconserved region in human cancers.* **RNA Biol.** 2013; 10:1771-7.

25. Rackham O, Shearwood AM, Mercer TR, Davies SM, Mattick JS, Filipovska A. *Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins.* **RNA.** 2011; 17:2085-93.

26. Clark MB, Amaral PP, Schlesinger FJ, Dinger ME, Taft RJ, Rinn JL, Ponting CP, Stadler PF, Morris KV, Morillon A, Rozowsky JS, Gerstein MB, Wahlestedt C, Hayashizaki Y, Carninci P, Gingeras TR, Mattick JS. *The reality of pervasive transcription.* **PLoS Biol.** 2011; 9:e1000625.

27. van Bakel H, Nislow C, Blencowe BJ, Hughes TR. *Most "dark matter" transcripts are associated with known genes.* **PLoS Biol.** 2010; 8:e1000371.

28. Kour S, Rath PC. *Age-Related Expression of a Repeat-Rich Intergenic Long Noncoding RNA in the Rat Brain.* **Mol Neurobiol.** 2017; 54:639-660.

29. You LH, Zhu LJ, Yang L, Shi CM, Pang LX, Zhang J, Cui XW, Ji CB, Guo XR. *Transcriptome analysis reveals the potential contribution of long noncoding RNAs to brown adipocyte differentiation.* **Mol Genet Genomics.** 2015; 290:1659-71.

30. Luan A, Paik KJ, Li J, Zielins ER, Atashroo DA, Spencley A, Momeni A, Longaker MT, Wang KC, Wan DC. *RNA Sequencing for Identification of Differentially Expressed Noncoding Transcripts during Adipogenic Differentiation of Adipose-Derived Stromal Cells.* **Plast Reconstr Surg.** 2015; 136:752-63.

31. Yu Q, Guo W, Shen J, Lv Y. *Effect of glucocorticoids on lncRNA and mRNA expression profiles of the bone microcirculatory endothelial cells from femur head of Homo sapiens.* **Genom Data.** 2015; 4:140-2.

32. Melo CA, Léveillé N, Rooijers K, Wijchers PJ, Geeven G, Tal A, Melo SA, de Laat W, Agami R. *A p53-bound enhancer region controls a long intergenic noncoding RNA required for p53 stress response.* **Oncogene.** 2016; 35:4399-406.

33. Kitagawa M, Kitagawa K, Kotake Y, Niida H, Ohhata T. *Cell cycle regulation by long non-coding RNAs.* **Cell Mol Life Sci.** 2013; 70:4785-94.

34. Nan A, Zhou X, Chen L, Liu M, Zhang N, Zhang L, Luo Y, Liu Z, Dai L, Jiang Y. *A transcribed ultraconserved noncoding RNA, Uc.173, is a key molecule for the inhibition of lead-induced neuronal apoptosis.* **Oncotarget.** 2016; 7:112-24.

35. He Q, Tan J, Yu B, Shi W, Liang K. *Long noncoding RNA HIF1A-AS1A reduces apoptosis of vascular smooth muscle cells: implications for the pathogenesis of thoracoabdominal aorta aneurysm.* **Pharmazie.** 2015; 70:310-5.

36. Wang C, Li G, Wu Y, Xi J, Kang J. *LincRNA1230 inhibits the differentiation of mouse ES cells towards neural progenitors.* **Sci China Life Sci.** 2016; 59:443-54.

37. Rosa A, Ballarino M. *Long Noncoding RNA Regulation of Pluripotency.* **Stem Cells Int.** 2016; 2016:1797692.

38. Kim J, Kim KM, Noh JH, Yoon JH, Abdelmohsen K, Gorospe M. *Long noncoding RNAs in diseases of aging.* **Biochim Biophys Acta.** 2016; 1859:209-21.

39. Frade AF, Laugier L, Ferreira LR, Baron MA, Benvenuti LA, Teixeira PC, Navarro IC, Cabantous S, Ferreira FM, da Silva Cândido D, Gaiotto FA, Bacal F, Pomerantzeff P, Santos RH, Kalil J, Cunha-Neto E, Chevillard C. *Myocardial Infarction-Associated Transcript, a Long Noncoding RNA, Is Overexpressed During Dilated Cardiomyopathy Due to Chronic Chagas Disease.* **J Infect Dis.** 2016; 214:161-5.

40. Széll M, Danis J, Bata-Csörgő Z, Kemény L. *PRINS, a primate-specific long non-coding RNA, plays a role in the keratinocyte stress response and psoriasis pathogenesis.* **Pflugers Arch.** 2016; 468:935-43.

41. Messemaker TC, Frank-Bertoncelj M, Marques RB, Adriaans A, Bakker AM, Daha N, Gay S, Huizinga TW, Toes RE, Mikkers HM, Kurreeman F. *A novel long non-coding RNA in the rheumatoid arthritis risk locus TRAF1-C5 influences C5 mRNA levels.* **Genes Immun.** 2016; 17:85-92.

42. Li CY, Zhu L, Liu Z, Zhou J, Shu Y. *Aberrant expression of long noncoding RNA PVT1 and its diagnostic and prognostic significance in patients with gastric cancer.* **Neoplasma.** 2016; 63:442-9.

43. Pan W, Zhou L, Ge M, Zhang B, Yang X, Xiong X, Fu G, Zhang J, Nie X, Li H, Tang X, Wei J, Shao M, Zheng J, Yuan Q, Tan W, Wu C, Yang M, Lin D. *Whole exome sequencing identifies lncRNA GAS8-AS1 and LPAR4 as novel papillary thyroid carcinoma driver alternations.* **Hum Mol Genet.** 2016; 25:1875-84.

44. He C, Jiang B, Ma J, Li Q. *Aberrant NEAT1 expression is associated with clinical outcome in high grade glioma patients.* **APMIS.** 2016; 124:169-74.

45. Yue B, Qiu S, Zhao S, Liu C, Zhang D, Yu F, Peng Z, Yan D. *LncRNA-ATB mediated E-cadherin repression promotes the progression of colon cancer and predicts poor prognosis.* **J Gastroenterol Hepatol.** 2016; 31:595-603.

46. Wu J, Li X, Xu Y, Yang T, Yang Q, Yang C, Jiang Y. *Identification of a long non-coding RNA NR_026689 associated with lung carcinogenesis induced by NNK.* **Oncotarget.** 2016; 7:14486-98.

47. Tordonato C, Di Fiore PP, Nicassio F. *The role of non-coding RNAs in the regulation of stem cells and progenitors in the normal mammary gland and in breast tumors.* **Front Genet.** 2015; 6:72.

48. Dieci G, Fiorino G, Castelnuovo M, Teichmann M, Pagano A. *The expanding RNA polymerase III transcriptome.* **Trends Genet.** 2007; 23:614-22.

49. Goodrich JA, Kugel JF. *Non-coding-RNA regulators of RNA polymerase II transcription.* **Nat Rev Mol Cell Biol.** 2006; 7:612-6.

50. Gardini A, Shiekhattar R. *The many faces of long noncoding RNAs.* **FEBS J.** 2015; 282:1647-57.

51. Wilusz JE. *Long noncoding RNAs: Re-writing dogmas of RNA processing and stability.* **Biochim Biophys Acta.** 2016; 1859:128-38.

52. Gomes AQ, Nolasco S, Soares H. *Non-coding RNAs: multi-tasking molecules in the cell.* **Int J Mol Sci.** 2013; 14:16010-39.

53. Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, Rastan S. *The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus.* **Cell.** 1992; 71:515-26.

54. Brown CJ, Hendrich BD, Rupert JL, Lafrenière RG, Xing Y, Lawrence J, Willard HF. *The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus.* **Cell.** 1992; 71:527-42.

55. van Bemmel JG, Mira-Bontenbal H, Gribnau J. *Cis- and trans-regulation in X inactivation.* **Chromosoma.** 2016; 125:41-50.

56. Wu Y, Zhang L, Wang Y, Li H, Ren X, Wei F, Yu W, Wang X, Zhang L, Yu J, Hao X. *Long noncoding RNA HOTAIR involvement in cancer.* **Tumour Biol.** 2014; 35:9531-8.

57. Bhan A, Mandal SS. *LncRNA HOTAIR: A master regulator of chromatin dynamics and cancer.* **Biochim Biophys Acta.** 2015; 1856:151-64.

58. Li L, Liu B, Wapinski OL, Tsai MC, Qu K, Zhang J, Carlson JC, Lin M, Fang F, Gupta RA, Helms JA, Chang HY. *Targeted disruption of Hotair leads to homeotic transformation and gene derepression.* **Cell Rep.** 2013; 5:3-12.

59. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY. *Long noncoding RNA as modular scaffold of histone modification complexes.* **Science.** 2010; 329:689-93.

60. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES, Rinn JL. *Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression.* **Proc Natl Acad Sci U S A.** 2009; 106:11667-72.

61. Hung T, Wang Y, Lin MF, Koegel AK, Kotake Y, Grant GD, Horlings HM, Shah N, Umbricht C, Wang P, Wang Y, Kong B, Langerød A, Børresen-Dale AL, Kim SK, van de Vijver M, Sukumar S, Whitfield ML, Kellis M, Xiong Y, Wong DJ, Chang HY. *Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters.* **Nat Genet.** 2011; 43:621-9.

62. Vance KW, Ponting CP. *Transcriptional regulatory functions of nuclear long noncoding RNAs.* **Trends Genet.** 2014; 30:348-55.

63. Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A. *Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript.* **Nature.** 2007; 445:666-70.

64. Liu T, Huang Y, Chen J, Chi H, Yu Z, Wang J, Chen C. *Attenuated ability of BACE1 to cleave the amyloid precursor protein via silencing long noncoding RNA BACE1‑AS expression.* **Mol Med Rep.** 2014; 10:1275-81.

65. Wilusz JE, Sunwoo H, Spector DL. *Long noncoding RNAs: functional surprises from the RNA world.* **Genes Dev.** 2009; 23:1494-504.

66. Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB. *An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles.* **Mol Cell.** 2009; 33:717-26.

67. Brannan CI, Dees EC, Ingram RS, Tilghman SM. *The product of the H19 gene may function as an RNA.* **Mol Cell Biol.** 1990; 10:28-36.

68. Rachmilewitz J, Goshen R, Ariel I, Schneider T, de Groot N, Hochberg A. *Parental imprinting of the human H19 gene.* **FEBS Lett.** 1992; 309:25-8.

69. Tsang WP, Ng EK, Ng SS, Jin H, Yu J, Sung JJ, Kwok TT. *Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer.* **Carcinogenesis.** 2010; 31:350-8.

70. Kallen AN, Zhou XB, Xu J, Qiao C, Ma J, Yan L, Lu L, Liu C, Yi JS, Zhang H, Min W, Bennett AM, Gregory RI, Ding Y, Huang Y. *The imprinted H19 lncRNA antagonizes let-7 microRNAs.* **Mol Cell.** 2013; 52:101-12.

71. Liu Q, Huang J, Zhou N, Zhang Z, Zhang A, Lu Z, Wu F, Mo YY. *LncRNA loc285194 is a p53-regulated tumor suppressor.* **Nucleic Acids Res.** 2013; 41:4976-87.

72. Lin R, Maeda S, Liu C, Karin M, Edgington TS. *A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas.* **Oncogene.** 2007; 26:851-8.

73. Gast M, Schroen B, Voigt A, Haas J, Kuehl U, Lassner D, Skurk C, Escher F, Wang X, Kratzer A, Michalik K, Papageorgiou A, Peters T, Loebel M, Wilk S, Althof N, Prasanth KV, Katus H, Meder B, Nakagawa S, Scheibenbogen C, Schultheiss HP, Landmesser U, Dimmeler S, Heymans S, Poller W. *Long noncoding RNA MALAT1-derived mascRNA is involved in cardiovascular innate immunity.* **J Mol Cell Biol.** 2016; 8:178-81.

74. Sanger HL, Klotz G, Riesner D, Gross HJ, Kleinschmidt AK. *Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures.* **Proc Natl Acad Sci U S A.** 1976; 73:3852-6.

75. Cocquerelle C, Mascrez B, Hétuin D, Bailleul B. *Mis-splicing yields circular RNA molecules.* **FASEB J.** 1993; 7:155-60.

76. Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. *Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types.* **PLoS One.** 2012; 7:e30733.

77. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE. *Circular RNAs are abundant, conserved, and associated with ALU repeats.* **RNA.** 2013; 19:141-57.

78. Matera AG, Wang Z. *A day in the life of the spliceosome.* **Nat Rev Mol Cell Biol.** 2014; 15:108-21.

79. Chen I, Chen CY, Chuang TJ. *Biogenesis, identification, and function of exonic circular RNAs.* **Wiley Interdiscip Rev RNA.** 2015; 6:563-79.

80. 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; 495:333-8.

81. Kameyama T, Suzuki H, Mayeda A. *Re-splicing of mature mRNA in cancer cells promotes activation of distant weak alternative splice sites.* **Nucleic Acids Res.** 2012; 40:7896-906.

82. Zhang XO, Wang HB, Zhang Y, Lu X, Chen LL, Yang L. *Complementary sequence-mediated exon circularization.* **Cell.** 2014; 159:134-47.

83. Ivanov A, Memczak S, Wyler E, Torti F, Porath HT, Orejuela MR, Piechotta M, Levanon EY, Landthaler M, Dieterich C, Rajewsky N. *Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals.* **Cell Rep.** 2015; 10:170-7.

84. Liang D, Wilusz JE. *Short intronic repeat sequences facilitate circular RNA production.* **Genes Dev.** 2014; 28:2233-47.

85. Salzman J, Chen RE, Olsen MN, Wang PL, Brown PO. *Cell-type specific features of circular RNA expression.* **PLoS Genet.** 2013; 9:e1003777.

86. Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, Zhu S, Yang L, Chen LL. *Circular intronic long noncoding RNAs.* **Mol Cell.** 2013; 51:792-806.

87. Conn SJ, Pillman KA, Toubia J, Conn VM, Salmanidis M, Phillips CA, Roslan S, Schreiber AW, Gregory PA, Goodall GJ. *The RNA binding protein quaking regulates formation of circRNAs.* **Cell.** 2015; 160:1125-34.

88. Wang Y, Wang Z. *Efficient backsplicing produces translatable circular mRNAs.* **RNA.** 2015; 21:172-9.

89. Glažar P, Papavasileiou P, Rajewsky N. *circBase: a database for circular RNAs.* **RNA.** 2014; 20:1666-70.

90. Guo JU, Agarwal V, Guo H, Bartel DP. *Expanded identification and characterization of mammalian circular RNAs.* **Genome Biol.** 2014; 15:409.

91. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J. *Natural RNA circles function as efficient microRNA sponges.* **Nature.** 2013; 495:384-8.

92. Thomas LF, Sætrom P. *Circular RNAs are depleted of polymorphisms at microRNA binding sites.* **Bioinformatics.** 2014; 30:2243-6.

93. You X, Vlatkovic I, Babic A, Will T, Epstein I, Tushev G, Akbalik G, Wang M, Glock C, Quedenau C, Wang X, Hou J, Liu H, Sun W, Sambandan S, Chen T, Schuman EM, Chen W. *Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity.* **Nat Neurosci.** 2015; 18:603-10.

94. Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, Yu B, Hu W, Dai L, Zhu P, Chang Z, Wu Q, Zhao Y, Jia Y, Xu P, Liu H, Shan G. *Exon-intron circular RNAs regulate transcription in the nucleus.* **Nat Struct Mol Biol.** 2015; 22:256-64.

95. Wang Z. *Not just a sponge: new functions of circular RNAs discovered.* **Sci China Life Sci.** 2015; 58:407-8.

96. Zlotorynski E. *Non-coding RNA: Circular RNAs promote transcription. Nat Rev Mol Cell Biol.* 2015; 16:206.

97. Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S. *circRNA biogenesis competes with pre-mRNA splicing.* **Mol Cell.** 2014; 56:55-66.

98. Granados-Riveron JT, Aquino-Jarquin G. *The complexity of the translation ability of circRNAs.* **Biochim Biophys Acta.** 2016; 1859:1245-51.

99. Kim VN, Nam JW. *Genomics of microRNA.* **Trends Genet.** 2006; 22:165-73.

100. Londin E, Loher P, Telonis AG, Quann K, Clark P, Jing Y, Hatzimichael E, Kirino Y, Honda S, Lally M, Ramratnam B, Comstock CE, Knudsen KE, Gomella L, Spaeth GL, Hark L, Katz LJ, Witkiewicz A, Rostami A, Jimenez SA, Hollingsworth MA, Yeh JJ, Shaw CA, McKenzie SE, Bray P, Nelson PT, Zupo S, Van Roosbroeck K, Keating MJ, Calin GA, Yeo C, Jimbo M, Cozzitorto J, Brody JR, Delgrosso K, Mattick JS, Fortina P, Rigoutsos I. *Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs.* **Proc Natl Acad Sci U S A.** 2015; 112:E1106-15.

101. Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN. *MicroRNA genes are transcribed by RNA polymerase II.* **EMBO J.** 2004; 23:4051-60.

102. Cai X, Hagedorn CH, Cullen BR. *Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs.* **RNA.** 2004; 10:1957-66.

103. Borchert GM, Lanier W, Davidson BL. *RNA polymerase III transcribes human microRNAs.* **Nat Struct Mol Biol.** 2006; 13:1097-101.

104. Lund E, Güttinger S, Calado A, Dahlberg JE, Kutay U. *Nuclear export of microRNA precursors.* **Science.** 2004; 303:95-8.

105. Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R. *Human RISC couples microRNA biogenesis and posttranscriptional gene silencing.* **Cell.** 2005; 123:631-40.

106. Haase AD, Jaskiewicz L, Zhang H, Lainé S, Sack R, Gatignol A, Filipowicz W. *TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing.* **EMBO Rep.** 2005; 6:961-7.

107. Lee Y, Hur I, Park SY, Kim YK, Suh MR, Kim VN. *The role of PACT in the RNA silencing pathway.* **EMBO J.** 2006; 25:522-32.

108. MacRae IJ, Ma E, Zhou M, Robinson CV, Doudna JA. *In vitro reconstitution of the human RISC-loading complex.* **Proc Natl Acad Sci U S A.** 2008; 105:512-7.

109. Filipowicz W, Bhattacharyya SN, Sonenberg N. *Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?* **Nat Rev Genet.** 2008; 9:102-14.

110. Kim VN. *MicroRNA biogenesis: coordinated cropping and dicing.* **Nat Rev Mol Cell Biol.** 2005; 6:376-85.

111. Yang JS, Phillips MD, Betel D, Mu P, Ventura A, Siepel AC, Chen KC, Lai EC. *Widespread regulatory activity of vertebrate microRNA* species.* **RNA.** 2011; 17:312-26.

112. Ro S, Park C, Young D, Sanders KM, Yan W. *Tissue-dependent paired expression of miRNAs.* **Nucleic Acids Res.** 2007; 35:5944-53.

113. Okamura K, Phillips MD, Tyler DM, Duan H, Chou YT, Lai EC. *The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution.* **Nat Struct Mol Biol.** 2008; 15:354-63.

114. Ruby JG, Jan CH, Bartel DP. *Intronic microRNA precursors that bypass Drosha processing.* **Nature.** 2007; 448:83-6.

115. Carthew RW, Sontheimer EJ. *Origins and Mechanisms of miRNAs and siRNAs.* **Cell.** 2009; 136:642-55.

116. Nilsen TW. *Mechanisms of microRNA-mediated gene regulation in animal cells.* **Trends Genet.** 2007; 23:243-9.

117. Ragusa M, Barbagallo D, Purrello M. *Exosomes: nanoshuttles to the future of BioMedicine.* **Cell Cycle.** 2015; 14:289-90.

118. Akers JC, Gonda D, Kim R, Carter BS, Chen CC. *Biogenesis of extracellular vesicles (EV): exosomes, microvesicles, retrovirus-like vesicles, and apoptotic bodies.* **J Neurooncol.** 2013 May; 113:1-11.

119. Urbanelli L, Magini A, Buratta S, Brozzi A, Sagini K, Polchi A, Tancini B, Emiliani C. *Signaling pathways in exosomes biogenesis, secretion and fate.* **Genes (Basel).** 2013; 4:152-70.

120. Théry C, Boussac M, Véron P, Ricciardi-Castagnoli P, Raposo G, Garin J, Amigorena S. *Proteomic analysis of dendritic cell-derived exosomes: a secreted subcellular compartment distinct from apoptotic vesicles.* **J Immunol.** 2001; 166:7309-18.

121. Kahlert C, Melo SA, Protopopov A, Tang J, Seth S, Koch M, Zhang J, Weitz J, Chin L, Futreal A, Kalluri R. *Identification of double-stranded genomic DNA spanning all chromosomes with mutated KRAS and p53 DNA in the serum exosomes of patients with pancreatic cancer.* **J Biol Chem.** 2014; 289:3869-75.

122. Thakur BK, Zhang H, Becker A, Matei I, Huang Y, Costa-Silva B, Zheng Y, Hoshino A, Brazier H, Xiang J, Williams C, Rodriguez-Barrueco R, Silva JM, Zhang W, Hearn S, Elemento O, Paknejad N, Manova-Todorova K, Welte K, Bromberg J, Peinado H, Lyden D. *Double-stranded DNA in exosomes: a novel biomarker in cancer detection.* **Cell Res.** 2014; 24:766-9.

123. Johnstone RM, Adam M, Hammond JR, Orr L, Turbide C. *Vesicle formation during reticulocyte maturation. Association of plasma membrane activities with released vesicles (exosomes).* **J Biol Chem.** 1987; 262:9412-20.

124. Raposo G, Nijman HW, Stoorvogel W, Liejendekker R, Harding CV, Melief CJ, Geuze HJ. *B lymphocytes secrete antigen-presenting vesicles.* **J Exp Med.** 1996; 183:1161-72.

125. Bodey B, Bodey B Jr, Kaiser HE. *Dendritic type, accessory cells within the mammalian thymic microenvironment. Antigen presentation in the dendritic neuro-endocrine-immune cellular network.* **In Vivo.** 1997; 11:351-70.

126. Skokos D, Le Panse S, Villa I, Rousselle JC, Peronet R, David B, Namane A, Mécheri S. *Mast cell-dependent B and T lymphocyte activation is mediated by the secretion of immunologically active exosomes.* **J Immunol.** 2001; 166:868-76.

127. Heijnen HF, Schiel AE, Fijnheer R, Geuze HJ, Sixma JJ. *Activated platelets release two types of membrane vesicles: microvesicles by surface shedding and exosomes derived from exocytosis of multivesicular bodies and alpha-granules.* **Blood.** 1999; 94:3791-9.

128. van Niel G, Raposo G, Candalh C, Boussac M, Hershberg R, Cerf-Bensussan N, Heyman M. *Intestinal epithelial cells secrete exosome-like vesicles.* **Gastroenterology.** 2001; 121:337-49.

129. Wolfers J, Lozier A, Raposo G, Regnault A, Théry C, Masurier C, Flament C, Pouzieux S, Faure F, Tursz T, Angevin E, Amigorena S, Zitvogel L. *Tumor-derived exosomes are a source of shared tumor rejection antigens for CTL cross-priming.* **Nat Med.** 2001; 7:297-303.

130. Andre F, Schartz NE, Movassagh M, Flament C, Pautier P, Morice P, Pomel C, Lhomme C, Escudier B, Le Chevalier T, Tursz T, Amigorena S, Raposo G, Angevin E, Zitvogel L. *Malignant effusions and immunogenic tumour-derived exosomes.* **Lancet.** 2002; 360:295-305.

131. Lopez-Verrilli MA, Caviedes A, Cabrera A, Sandoval S, Wyneken U, Khoury M. *Mesenchymal stem cell-derived exosomes from different sources selectively promote neuritic outgrowth.* **Neuroscience.** 2016; 320:129-39.

132. Ibrahim A, Marbán E. *Exosomes: Fundamental Biology and Roles in Cardiovascular Physiology.* **Annu Rev Physiol.** 2016; 78:67-83.

133. Verkhratsky A, Matteoli M, Parpura V, Mothet JP, Zorec R. *Astrocytes as secretory cells of the central nervous system: idiosyncrasies of vesicular secretion.* **EMBO J.** 2016; 35:239-57.

134. Hoorn EJ, Pisitkun T, Zietse R, Gross P, Frokiaer J, Wang NS, Gonzales PA, Star RA, Knepper MA. *Prospects for urinary proteomics: exosomes as a source of urinary biomarkers.* **Nephrology (Carlton).** 2005; 10:283-90.

135. Gatti JL, Métayer S, Belghazi M, Dacheux F, Dacheux JL. *Identification, proteomic profiling, and origin of ram epididymal fluid exosome-like vesicles.* **Biol Reprod.** 2005; 72:1452-65.

136. Keller S, Rupp C, Stoeck A, Runz S, Fogel M, Lugert S, Hager HD, Abdel-Bakky MS, Gutwein P, Altevogt P. *CD24 is a marker of exosomes secreted into urine and amniotic fluid.* **Kidney Int.** 2007; 72:1095-102.

137. Gonzalez-Begne M, Lu B, Han X, Hagen FK, Hand AR, Melvin JE, Yates JR. *Proteomic analysis of human parotid gland exosomes by multidimensional protein identification technology (MudPIT).* **J Proteome Res.** 2009; 8:1304-14.

138. Lässer C, Alikhani VS, Ekström K, Eldh M, Paredes PT, Bossios A, Sjöstrand M, Gabrielsson S, Lötvall J, Valadi H. *Human saliva, plasma and breast milk exosomes contain RNA: uptake by macrophages.* **J Transl Med.** 2011; 9:9.

139. Beach A, Zhang HG, Ratajczak MZ, Kakar SS. *Exosomes: an overview of biogenesis, composition and role in ovarian cancer.* **J Ovarian Res.** 2014; 7:14.

140. van Dommelen SM, Vader P, Lakhal S, Kooijmans SA, van Solinge WW, Wood MJ, Schiffelers RM. *Microvesicles and exosomes: opportunities for cell-derived membrane vesicles in drug delivery.* **J Control Release.** 2012; 161:635-44.

141. Savina A, Furlán M, Vidal M, Colombo MI. *Exosome release is regulated by a calcium-dependent mechanism in K562 cells.* **J Biol Chem.** 2003; 278:20083-90.

142. Krämer-Albers EM, Bretz N, Tenzer S, Winterstein C, Möbius W, Berger H, Nave KA, Schild H, Trotter J. *Oligodendrocytes secrete exosomes containing major myelin and stress-protective proteins: Trophic support for axons?* **Proteomics Clin Appl.** 2007; 1:1446-61.

143. Fauré J, Lachenal G, Court M, Hirrlinger J, Chatellard-Causse C, Blot B, Grange J, Schoehn G, Goldberg Y, Boyer V, Kirchhoff F, Raposo G, Garin J, Sadoul R. *Exosomes are released by cultured cortical neurones.* **Mol Cell Neurosci.** 2006; 31:642-8.

144. Trajkovic K, Hsu C, Chiantia S, Rajendran L, Wenzel D, Wieland F, Schwille P, Brügger B, Simons M. *Ceramide triggers budding of exosome vesicles into multivesicular endosomes.* **Science.** 2008; 319:1244-7.

145. Jaé N, McEwan DG, Manavski Y, Boon RA, Dimmeler S. *Rab7a and Rab27b control secretion of endothelial microRNA through extracellular vesicles.* **FEBS Lett.** 2015; 589:3182-8.

146. Li W, Hu Y, Jiang T, Han Y, Han G, Chen J, Li X. *Rab27A regulates exosome secretion from lung adenocarcinoma cells A549: involvement of EPI64.* **APMIS.** 2014; 122:1080-7.

147. Hsu C, Morohashi Y, Yoshimura S, Manrique-Hoyos N, Jung S, Lauterbach MA, Bakhti M, Grønborg M, Möbius W, Rhee J, Barr FA, Simons M. *Regulation of exosome secretion by Rab35 and its GTPase-activating proteins TBC1D10A-C.* **J Cell Biol.** 2010; 189:223-32.

148. Ostrowski M, Carmo NB, Krumeich S, Fanget I, Raposo G, Savina A, Moita CF, Schauer K, Hume AN, Freitas RP, Goud B, Benaroch P, Hacohen N, Fukuda M, Desnos C, Seabra MC, Darchen F, Amigorena S, Moita LF, Thery C. *Rab27a and Rab27b control different steps of the exosome secretion pathway.* **Nat Cell Biol.** 2010; 12:19-30.

149. Loomis RJ, Holmes DA, Elms A, Solski PA, Der CJ, Su L. *Citron kinase, a RhoA effector, enhances HIV-1 virion production by modulating exocytosis.* **Traffic.** 2006; 7:1643-53.

150. Chaineau M, Danglot L, Galli T. *Multiple roles of the vesicular-SNARE TI-VAMP in post-Golgi and endosomal trafficking.* **FEBS Lett.** 2009; 583:3817-26.

151. Fader CM, Sánchez DG, Mestre MB, Colombo MI. *TI-VAMP/VAMP7 and VAMP3/cellubrevin: two v-SNARE proteins involved in specific steps of the autophagy/multivesicular body pathways.* **Biochim Biophys Acta.** 2009; 1793:1901-16.

152. Llorente A, Skotland T, Sylvänne T, Kauhanen D, Róg T, Orłowski A, Vattulainen I, Ekroos K, Sandvig K. *Molecular lipidomics of exosomes released by PC-3 prostate cancer cells.* **Biochim Biophys Acta.** 2013; 1831:1302-9.

153. Laulagnier K, Motta C, Hamdi S, Roy S, Fauvelle F, Pageaux JF, Kobayashi T, Salles JP, Perret B, Bonnerot C, Record M. *Mast cell- and dendritic cell-derived exosomes display a specific lipid composition and an unusual membrane organization.* **Biochem J.** 2004; 380:161-71.

154. Fitzner D, Schnaars M, van Rossum D, Krishnamoorthy G, Dibaj P, Bakhti M, Regen T, Hanisch UK, Simons M. *Selective transfer of exosomes from oligodendrocytes to microglia by macropinocytosis.* **J Cell Sci.** 2011; 124:447-58.

155. Parolini I, Federici C, Raggi C, Lugini L, Palleschi S, De Milito A, Coscia C, Iessi E, Logozzi M, Molinari A, Colone M, Tatti M, Sargiacomo M, Fais S. *Microenvironmental pH is a key factor for exosome traffic in tumor cells.* **J Biol Chem.** 2009; 284:34211-22.

156. Huang X, Yuan T, Tschannen M, Sun Z, Jacob H, Du M, Liang M, Dittmar RL, Liu Y, Liang M, Kohli M, Thibodeau SN, Boardman L, Wang L. *Characterization of human plasma-derived exosomal RNAs by deep sequencing.* **BMC Genomics.** 2013; 14:319.

157. Ridder K, Keller S, Dams M, Rupp AK, Schlaudraff J, Del Turco D, Starmann J, Macas J, Karpova D, Devraj K, Depboylu C, Landfried B, Arnold B, Plate KH, Höglinger G, Sültmann H, Altevogt P, Momma S. *Extracellular vesicle-mediated transfer of genetic information between the hematopoietic system and the brain in response to inflammation.* **PLoS Biol.** 2014; 12:e1001874.

158. Subra C, Grand D, Laulagnier K, Stella A, Lambeau G, Paillasse M, De Medina P, Monsarrat B, Perret B, Silvente-Poirot S, Poirot M, Record M. *Exosomes account for vesicle-mediated transcellular transport of activatable phospholipases and prostaglandins.* **J Lipid Res.** 2010; 51:2105-20.

159. Christianson HC, Svensson KJ, van Kuppevelt TH, Li JP, Belting M. *Cancer cell exosomes depend on cell-surface heparan sulfate proteoglycans for their internalization and functional activity.* **Proc Natl Acad Sci U S A.** 2013; 110:17380-5.

160. Sinha A, Ignatchenko V, Ignatchenko A, Mejia-Guerrero S, Kislinger T. *In-depth proteomic analyses of ovarian cancer cell line exosomes reveals differential enrichment of functional categories compared to the NCI 60 proteome.* **Biochem Biophys Res Commun.** 2014; 445:694-701.

161. Tauro BJ, Greening DW, Mathias RA, Ji H, Mathivanan S, Scott AM, Simpson RJ. *Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes.* **Methods.** 2012; 56:293-304.

162. Fedele C, Singh A, Zerlanko BJ, Iozzo RV, Languino LR. *The αvβ6 integrin is transferred intercellularly via exosomes.* **J Biol Chem.** 2015; 290:4545-51.

163. Simpson RJ, Lim JW, Moritz RL, Mathivanan S. *Exosomes: proteomic insights and diagnostic potential.* **Expert Rev Proteomics.** 2009; 6:267-83.

164. Valadi H, Ekström K, Bossios A, Sjöstrand M, Lee JJ, Lötvall JO. *Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells.* **Nat Cell Biol.** 2007; 9:654-9.

165. Kogure T, Yan IK, Lin WL, Patel T. *Extracellular Vesicle-Mediated Transfer of a Novel Long Noncoding RNA TUC339: A Mechanism of Intercellular Signaling in Human Hepatocellular Cancer.* **Genes Cancer.** 2013; 4:261-72.

166. Takahashi K, Yan IK, Kogure T, Haga H, Patel T. *Extracellular vesicle-mediated transfer of long non-coding RNA ROR modulates chemosensitivity in human hepatocellular cancer.* **FEBS Open Bio.** 2014; 4:458-67.

167. Gezer U, Özgür E, Cetinkaya M, Isin M, Dalay N. *Long non-coding RNAs with low expression levels in cells are enriched in secreted exosomes.* **Cell Biol Int.** 2014; 38:1076-9.

168. Lasda E, Parker R. *Circular RNAs Co-Precipitate with Extracellular Vesicles: A Possible Mechanism for circRNA Clearance.* **PLoS One.** 2016; 11:e0148407.

169. Eldh M, Ekström K, Valadi H, Sjöstrand M, Olsson B, Jernås M, Lötvall J. *Exosomes communicate protective messages during oxidative stress; possible role of exosomal shuttle RNA.* **PLoS One.** 2010; 5:e15353.

170. Ragusa M, Statello L, Maugeri M, Barbagallo C, Passanisi R, Alhamdani MS, Li Destri G, Cappellani A, Barbagallo D, Scalia M, Valadi H, Hoheisel JD, Di Pietro C, Purrello M. *Highly skewed distribution of miRNAs and proteins between colorectal cancer cells and their exosomes following Cetuximab treatment: biomolecular, genetic and translational implications.* **Oncoscience.** 2014; 1:132-57.

171. Greening DW, Nguyen HP, Elgass K, Simpson RJ, Salamonsen LA. *Human Endometrial Exosomes Contain Hormone-Specific Cargo Modulating Trophoblast Adhesive Capacity: Insights into Endometrial-Embryo Interactions.* **Biol Reprod.** 2016; 94:38.

172. Choi DS, Kim DK, Kim YK, Gho YS. *Proteomics, transcriptomics and lipidomics of exosomes and ectosomes.* **Proteomics.** 2013; 13:1554-71.

173. Villarroya-Beltri C, Gutiérrez-Vázquez C, Sánchez-Cabo F, Pérez-Hernández D, Vázquez J, Martin-Cofreces N, Martinez-Herrera DJ, Pascual-Montano A, Mittelbrunn M, Sánchez-Madrid F. *Sumoylated hnRNPA2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs.* **Nat Commun.** 2013; 4:2980.

174. Yang JM, Gould SJ. *The cis-acting signals that target proteins to exosomes and microvesicles.* **Biochem Soc Trans.** 2013; 41:277-82.

175. Robbins PD, Morelli AE. *Regulation of immune responses by extracellular vesicles.* **Nat Rev Immunol.** 2014; 14:195-208.

176. Luga V, Zhang L, Viloria-Petit AM, Ogunjimi AA, Inanlou MR, Chiu E, Buchanan M, Hosein AN, Basik M, Wrana JL. *Exosomes mediate stromal mobilization of autocrine Wnt-PCP signaling in breast cancer cell migration.* **Cell.** 2012; 151:1542-56.

177. Zhu W, Huang L, Li Y, Zhang X, Gu J, Yan Y, Xu X, Wang M, Qian H, Xu W. *Exosomes derived from human bone marrow mesenchymal stem cells promote tumor growth in vivo.* **Cancer Lett.** 2012; 315:28-37.

178. Zduriencikova M, Gronesova P, Cholujova D, Sedlak J. *Potential biomarkers of exosomal cargo in endocrine signaling.* **Endocr Regul.** 2015; 49:141-50.

179. Feng D, Zhao WL, Ye YY, Bai XC, Liu RQ, Chang LF, Zhou Q, Sui SF. *Cellular internalization of exosomes occurs through phagocytosis.* **Traffic.** 2010; 11:675-87.

180. Hawari FI, Rouhani FN, Cui X, Yu ZX, Buckley C, Kaler M, Levine SJ. *Release of full-length 55-kDa TNF receptor 1 in exosome-like vesicles: a mechanism for generation of soluble cytokine receptors.* **Proc Natl Acad Sci U S A.** 2004; 101:1297-302.

181. Mathivanan S. *Exosomes and Shedding Microvesicles are Mediators of Intercellular Communication: How do they Communicate with the Target Cells?* **J Biotechnol Biomater.** 2012; 2:e110.

182. Muntasell A, Berger AC, Roche PA. *T cell-induced secretion of MHC class II-peptide complexes on B cell exosomes.* **EMBO J.** 2007; 26:4263-72.

183. Nolte-'t Hoen EN, Buschow SI, Anderton SM, Stoorvogel W, Wauben MH. *Activated T cells recruit exosomes secreted by dendritic cells via LFA-1.* **Blood.** 2009; 113:1977-81.

184. Admyre C, Johansson SM, Paulie S, Gabrielsson S. *Direct exosome stimulation of peripheral human T cells detected by ELISPOT.* **Eur J Immunol.** 2006; 36:1772-81.

185. Théry C, Duban L, Segura E, Véron P, Lantz O, Amigorena S. *Indirect activation of naïve CD4+ T cells by dendritic cell-derived exosomes.* **Nat Immunol.** 2002; 3:1156-62.

186. Segura E, Nicco C, Lombard B, Véron P, Raposo G, Batteux F, Amigorena S, Théry C. *ICAM-1 on exosomes from mature dendritic cells is critical for efficient naive T-cell priming.* **Blood.** 2005; 106:216-23.

187. Qazi KR, Gehrmann U, Domange Jordö E, Karlsson MC, Gabrielsson S. *Antigen-loaded exosomes alone induce Th1-type memory through a B-cell-dependent mechanism.* **Blood.** 2009; 113:2673-83.

188. Chairoungdua A, Smith DL, Pochard P, Hull M, Caplan MJ. *Exosome release of β-catenin: a novel mechanism that antagonizes Wnt signaling.* **J Cell Biol.** 2010; 190:1079-91.

189. Lachenal G, Pernet-Gallay K, Chivet M, Hemming FJ, Belly A, Bodon G, Blot B, Haase G, Goldberg Y, Sadoul R. *Release of exosomes from differentiated neurons and its regulation by synaptic glutamatergic activity.* **Mol Cell Neurosci.** 2011; 46:409-18.

190. Fröhlich D, Kuo WP, Frühbeis C, Sun JJ, Zehendner CM, Luhmann HJ, Pinto S, Toedling J, Trotter J, Krämer-Albers EM. *Multifaceted effects of oligodendroglial exosomes on neurons: impact on neuronal firing rate, signal transduction and gene regulation.* **Philos Trans R Soc Lond B Biol Sci.** 2014; 369. pii: 20130510.

191. Akkoca AN, Yanık S, Ozdemir ZT, Cihan FG, Sayar S, Cincin TG, Cam A, Ozer C. *TNM and Modified Dukes staging along with the demographic characteristics of patients with colorectal carcinoma.* **Int J Clin Exp Med.** 2014; 7:2828-35.

192. I numeri del cancro in Italia – 2015. http://www.registri-tumori.it/.

193. Markowitz SD, Bertagnolli MM. *Molecular origins of cancer: Molecular basis of colorectal cancer.* **N Engl J Med.** 2009; 361:2449-60.

194. Morán A, Ortega P, de Juan C, Fernández-Marcelo T, Frías C, Sánchez-Pernaute A, Torres AJ, Díaz-Rubio E, Iniesta P, Benito M. *Differential colorectal carcinogenesis: Molecular basis and clinical relevance.* **World J Gastrointest Oncol.** 2010; 2:151-8.

195. Fearon ER, Vogelstein B. *A genetic model for colorectal tumorigenesis.* **Cell.** 1990; 61:759-67.

196. Applegate CC, Lane MA. *Role of retinoids in the prevention and treatment of colorectal cancer.* **World J Gastrointest Oncol.** 2015; 7:184-203.

197. Zlobec I, Bihl MP, Schwarb H, Terracciano L, Lugli A. *Clinicopathological and protein characterization of BRAF- and K-RAS-mutated colorectal cancer and implications for prognosis.* **Int J Cancer.** 2010; 127:367-80.

198. Papadatos-Pastos D, Rabbie R, Ross P, Sarker D. *The role of the PI3K pathway in colorectal cancer.* **Crit Rev Oncol Hematol.** 2015; 94:18-30.

199. Chen HM, Fang JY. *Epigenetic Biomarkers for the Early Detection of Gastrointestinal Cancer.* **Gastrointest Tumors.** 2014; 1:201-8.

200. Hendon SE, DiPalma JA. *U.S. practices for colon cancer screening.* **Keio J Med.** 2005; 54:179-83.

201. Garborg K, Holme Ø, Løberg M, Kalager M, Adami HO, Bretthauer M. *Current status of screening for colorectal cancer.* **Ann Oncol.** 2013; 24:1963-72.

202. Markowitz SD, Dawson DM, Willis J, Willson JK. *Focus on colon cancer.* **Cancer Cell.** 2002; 1:233-6.

203. Meyerhardt JA, Mayer RJ. *Systemic therapy for colorectal cancer.* **N Engl J Med.** 2005; 352:476-87.

204. Kanthan R, Senger JL, Kanthan SC. *Molecular events in primary and metastatic colorectal carcinoma: a review.* **Patholog Res Int.** 2012; 2012:597497.

205. Chang DZ, Kumar V, Ma Y, Li K, Kopetz S. *Individualized therapies in colorectal cancer: KRAS as a marker for response to EGFR-targeted therapy.* **J Hematol Oncol.** 2009; 2:18.

206. Kimura Y, Morohashi S, Yoshizawa T, Suzuki T, Morohashi H, Sakamoto Y, Koyama M, Murata A, Kijima H, Hakamada K. *Clinicopathological significance of vascular endothelial growth factor, thymidine phosphorylase and microvessel density in colorectal cancer.* **Mol Med Rep.** 2016; 13:1551-7.

207. Ji P, Diederichs S, Wang W, Böing S, Metzger R, Schneider PM, Tidow N, Brandt B, Buerger H, Bulk E, Thomas M, Berdel WE, Serve H, Müller-Tidow C. *MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer.* **Oncogene.** 2003; 22:8031-41.

208. Lai MC, Yang Z, Zhou L, Zhu QQ, Xie HY, Zhang F, Wu LM, Chen LM, Zheng SS. *Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation.* **Med Oncol.** 2012; 29:1810-6.

209. Yamada K, Kano J, Tsunoda H, Yoshikawa H, Okubo C, Ishiyama T, Noguchi M. *Phenotypic characterization of endometrial stromal sarcoma of the uterus.* **Cancer Sci.** 2006; 97:106-12.

210. Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, Chess A. *A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains.* **BMC Genomics.** 2007; 8:39.

211. Garen A, Song X. *Regulatory roles of tumor-suppressor proteins and noncoding RNA in cancer and normal cell functions.* **Int J Cancer.** 2008; 122:1687-9.

212. Yang L, Lin C, Liu W, Zhang J, Ohgi KA, Grinstein JD, Dorrestein PC, Rosenfeld MG. *ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs.* **Cell.** 2011; 147:773-88.

213. Ying L, Chen Q, Wang Y, Zhou Z, Huang Y, Qiu F. *Upregulated MALAT-1 contributes to bladder cancer cell migration by inducing epithelial-to-mesenchymal transition.* **Mol Biosyst.** 2012; 8:2289-94.

214. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA, Blencowe BJ, Prasanth SG, Prasanth KV. *The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation.* **Mol Cell.** 2010; 39:925-38.

215. Xu C, Yang M, Tian J, Wang X, Li Z. *MALAT-1: a long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis.* **Int J Oncol.** 2011; 39:169-75.

216. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, Davis N, Dicks E, Ewing R, Floyd Y, Gray K, Hall S, Hawes R, Hughes J, Kosmidou V, Menzies A, Mould C, Parker A, Stevens C, Watt S, Hooper S, Wilson R, Jayatilake H, Gusterson BA, Cooper C, Shipley J, Hargrave D, Pritchard-Jones K, Maitland N, Chenevix-Trench G, Riggins GJ, Bigner DD, Palmieri G, Cossu A, Flanagan A, Nicholson A, Ho JW, Leung SY, Yuen ST, Weber BL, Seigler HF, Darrow TL, Paterson H, Marais R, Marshall CJ, Wooster R, Stratton MR, Futreal PA. *Mutations of the BRAF gene in human cancer.* **Nature.** 2002; 417:949-54.

217. Guo Q, Zhao Y, Chen J, Hu J, Wang S, Zhang D, Sun Y. *BRAF-activated long non-coding RNA contributes to colorectal cancer migration by inducing epithelial-mesenchymal transition.* **Oncol Lett.** 2014; 8:869-875.

218. Grisanzio C, Freedman ML. *Chromosome 8q24-Associated Cancers and MYC.* **Genes Cancer.** 2010; 1:555-9.

219. Huppi K, Pitt JJ, Wahlberg BM, Caplen NJ. *The 8q24 gene desert: an oasis of non-coding transcriptional activity.* **Front Genet.** 2012; 3:69.

220. Haiman CA, Le Marchand L, Yamamoto J, Stram DO, Sheng X, Kolonel LN, Wu AH, Reich D, Henderson BE. *A common genetic risk factor for colorectal and prostate cancer.* **Nat Genet.** 2007; 39:954-6.

221. Tuupanen S, Turunen M, Lehtonen R, Hallikas O, Vanharanta S, Kivioja T, Björklund M, Wei G, Yan J, Niittymäki I, Mecklin JP, Järvinen H, Ristimäki A, Di-Bernardo M, East P, Carvajal-Carmona L, Houlston RS, Tomlinson I, Palin K, Ukkonen E, Karhu A, Taipale J, Aaltonen LA. *The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling.* **Nat Genet.** 2009; 41:885-90.

222. Ling H, Spizzo R, Atlasi Y, Nicoloso M, Shimizu M, Redis RS, Nishida N, Gafà R, Song J, Guo Z, Ivan C, Barbarotto E, De Vries I, Zhang X, Ferracin M, Churchman M, van Galen JF, Beverloo BH, Shariati M, Haderk F, Estecio MR, Garcia-Manero G, Patijn GA, Gotley DC, Bhardwaj V, Shureiqi I, Sen S, Multani AS, Welsh J, Yamamoto K, Taniguchi I, Song MA, Gallinger S, Casey G, Thibodeau SN, Le Marchand L, Tiirikainen M, Mani SA, Zhang W, Davuluri RV, Mimori K, Mori M, Sieuwerts AM, Martens JW, Tomlinson I, Negrini M, Berindan-Neagoe I, Foekens JA, Hamilton SR, Lanza G, Kopetz S, Fodde R, Calin GA. *CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer.* **Genome Res.** 2013; 23:1446-61.

223. Nissan A, Stojadinovic A, Mitrani-Rosenbaum S, Halle D, Grinbaum R, Roistacher M, Bochem A, Dayanc BE, Ritter G, Gomceli I, Bostanci EB, Akoglu M, Chen YT, Old LJ, Gure AO. *Colon cancer associated transcript-1: a novel RNA expressed in malignant and pre-malignant human tissues.* **Int J Cancer.** 2012; 130:1598-606.

224. Xiang JF, Yin QF, Chen T, Zhang Y, Zhang XO, Wu Z, Zhang S, Wang HB, Ge J, Lu X, Yang L, Chen LL. *Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus.* **Cell Res.** 2014; 24:513-31.

225. Tseng YY, Moriarity BS, Gong W, Akiyama R, Tiwari A, Kawakami H, Ronning P, Reuland B, Guenther K, Beadnell TC, Essig J, Otto GM, O'Sullivan MG, Largaespada DA, Schwertfeger KL, Marahrens Y, Kawakami Y, Bagchi A. *PVT1 dependence in cancer with MYC copy-number increase.* **Nature.** 2014; 512:82-6.

226. Beck-Engeser GB, Lum AM, Huppi K, Caplen NJ, Wang BB, Wabl M. *Pvt1-encoded microRNAs in oncogenesis.* **Retrovirology.** 2008; 5:4.

227. Barsotti AM, Beckerman R, Laptenko O, Huppi K, Caplen NJ, Prives C. *p53-Dependent induction of PVT1 and miR-1204.* **J Biol Chem.** 2012; 287:2509-19.

228. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, Sukumar S, Chang HY. *Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis.* **Nature.** 2010; 464:1071-6.

229. Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, Tanaka F, Shibata K, Suzuki A, Komune S, Miyano S, Mori M. *Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers.* **Cancer Res.** 2011; 71:6320-6.

230. Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, Young G, Lucas AB, Ach R, Bruhn L, Yang X, Amit I, Meissner A, Regev A, Rinn JL, Root DE, Lander ES. *lincRNAs act in the circuitry controlling pluripotency and differentiation.* **Nature.** 2011; 477:295-300.

231. Yue B, Sun B, Liu C, Zhao S, Zhang D, Yu F, Yan D. *Long non-coding RNA Fer-1-like protein 4 suppresses oncogenesis and exhibits prognostic value by associating with miR-106a-5p in colon cancer.* **Cancer Sci.** 2015; 106:1323-32.

232. Cui H, Onyango P, Brandenburg S, Wu Y, Hsieh CL, Feinberg AP. *Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2.* **Cancer Res.** 2002; 62:6442-6.

233. Taniue K, Kurimoto A, Sugimasa H, Nasu E, Takeda Y, Iwasaki K, Nagashima T, Okada-Hatakeyama M, Oyama M, Kozuka-Hata H, Hiyoshi M, Kitayama J, Negishi L, Kawasaki Y, Akiyama T. *Long noncoding RNA UPAT promotes colon tumorigenesis by inhibiting degradation of UHRF1.* **Proc Natl Acad Sci U S A.** 2016; 113:1273-8.

234. Merry CR, Forrest ME, Sabers JN, Beard L, Gao XH, Hatzoglou M, Jackson MW, Wang Z, Markowitz SD, Khalil AM. *DNMT1-associated long non-coding RNAs regulate global gene expression and DNA methylation in colon cancer.* **Hum Mol Genet.** 2015; 24:6240-53.

235. Bachmayr-Heyda A, Reiner AT, Auer K, Sukhbaatar N, Aust S, Bachleitner-Hofmann T, Mesteri I, Grunt TW, Zeillinger R, Pils D. *Correlation of circular RNA abundance with proliferation--exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis, and normal human tissues.* **Sci Rep.** 2015; 5:8057.

236. Wang X, Zhang Y, Huang L, Zhang J, Pan F, Li B, Yan Y, Jia B, Liu H, Li S, Zheng W. *Decreased expression of hsa_circ_001988 in colorectal cancer and its clinical significances.* **Int J Clin Exp Pathol.** 2015; 8:16020-5.

237. Dou Y, Cha DJ, Franklin JL, Higginbotham JN, Jeppesen DK, Weaver AM, Prasad N, Levy S, Coffey RJ, Patton JG, Zhang B. *Circular RNAs are down-regulated in KRAS mutant colon cancer cells and can be transferred to exosomes.* **Sci Rep.** 2016; 6:37982.

238. Xie H, Ren X, Xin S, Lan X, Lu G, Lin Y, Yang S, Zeng Z, Liao W, Ding YQ, Liang L. *Emerging roles of circRNA_001569 targeting miR-145 in the proliferation and invasion of colorectal cancer.* **Oncotarget.** 2016; 7:26680-91.

239. Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M, Croce CM. *Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers.* **Proc Natl Acad Sci U S A.** 2004; 101:2999-3004.

240. Gaur A, Jewell DA, Liang Y, Ridzon D, Moore JH, Chen C, Ambros VR, Israel MA. *Characterization of microRNA expression levels and their biological correlates in human cancer cell lines.* **Cancer Res.** 2007; 67:2456-68.

241. Esquela-Kerscher A, Slack FJ. *Oncomirs - microRNAs with a role in cancer.* **Nat Rev Cancer.** 2006; 6:259-69.

242. Croce CM. *Causes and consequences of microRNA dysregulation in cancer.* **Nat Rev Genet.** 2009; 10:704-14.

243. Logozzi M, De Milito A, Lugini L, Borghi M, Calabrò L, Spada M, Perdicchio M, Marino ML, Federici C, Iessi E, Brambilla D, Venturi G, Lozupone F, Santinami M, Huber V, Maio M, Rivoltini L, Fais S. High levels of exosomes expressing CD63 and caveolin-1 in plasma of melanoma patients. **PLoS One.** 2009; 4:e5219.

244. Taylor DD, Lyons KS, Gerçel-Taylor C. *Shed membrane fragment-associated markers for endometrial and ovarian cancers.* **Gynecol Oncol.** 2002; 84:443-8.

245. Rabinowits G, Gerçel-Taylor C, Day JM, Taylor DD, Kloecker GH. *Exosomal microRNA: a diagnostic marker for lung cancer.* **Clin Lung Cancer.** 2009; 10:42-6.

246. Kim JW, Wieckowski E, Taylor DD, Reichert TE, Watkins S, Whiteside TL. *Fas ligand-positive membranous vesicles isolated from sera of patients with oral cancer induce apoptosis of activated T lymphocytes.* **Clin Cancer Res.** 2005; 11:1010-20.

247. Abusamra AJ, Zhong Z, Zheng X, Li M, Ichim TE, Chin JL, Min WP. *Tumor exosomes expressing Fas ligand mediate CD8+ T-cell apoptosis.* **Blood Cells Mol Dis.** 2005; 35:169-73.

248. Valenti R, Huber V, Filipazzi P, Pilla L, Sovena G, Villa A, Corbelli A, Fais S, Parmiani G, Rivoltini L. *Human tumor-released microvesicles promote the differentiation of myeloid cells with transforming growth factor-beta-mediated suppressive activity on T lymphocytes.* **Cancer Res.** 2006; 66:9290-8.

249. Marleau AM, Chen CS, Joyce JA, Tullis RH. *Exosome removal as a therapeutic adjuvant in cancer.* **J Transl Med.** 2012; 10:134.

250. Higginbotham JN, Demory Beckler M, Gephart JD, Franklin JL, Bogatcheva G, Kremers GJ, Piston DW, Ayers GD, McConnell RE, Tyska MJ, Coffey RJ. *Amphiregulin exosomes increase cancer cell invasion.* **Curr Biol.** 2011; 21:779-86.

251. Al-Nedawi K, Meehan B, Kerbel RS, Allison AC, Rak J. *Endothelial expression of autocrine VEGF upon the uptake of tumor-derived microvesicles containing oncogenic EGFR.* **Proc Natl Acad Sci U S A.** 2009; 106:3794-9.

252. Ciravolo V, Huber V, Ghedini GC, Venturelli E, Bianchi F, Campiglio M, Morelli D, Villa A, Della Mina P, Menard S, Filipazzi P, Rivoltini L, Tagliabue E, Pupa SM. *Potential role of HER2-overexpressing exosomes in countering trastuzumab-based therapy.* **J Cell Physiol.** 2012; 227:658-67.

253. Graham LD, Pedersen SK, Brown GS, Ho T, Kassir Z, Moynihan AT, Vizgoft EK, Dunne R, Pimlott L, Young GP, Lapointe LC, Molloy PL. *Colorectal Neoplasia Differentially Expressed (CRNDE), a Novel Gene with Elevated Expression in Colorectal Adenomas and Adenocarcinomas.* **Genes Cancer.** 2011; 2:829-40.

254. Shi J, Li X, Zhang F, Zhang C, Guan Q, Cao X, Zhu W, Zhang X, Cheng Y, Ou K, Chen Q, Hu S. *Circulating lncRNAs associated with occurrence of colorectal cancer progression.* **Am J Cancer Res.** 2015; 5:2258-65.

255. Zhao W, Song M, Zhang J, Kuerban M, Wang H. *Combined identification of long non-coding RNA CCAT1 and HOTAIR in serum as an effective screening for colorectal carcinoma.* **Int J Clin Exp Pathol.** 2015; 8:14131-40.

256. Cancer Genome Atlas Network. *Comprehensive molecular characterization of human colon and rectal cancer.* **Nature.** 2012; 487:330-7.

257. Mhaidat NM, Alali FQ, Matalqah SM, Matalka II, Jaradat SA, Al-Sawalha NA, Thorne RF. *Inhibition of MEK sensitizes paclitaxel-induced apoptosis of human colorectal cancer cells by downregulation of GRP78.* **Anticancer Drugs.** 2009; 20:601-6.

258. Alaiyan B, Ilyayev N, Stojadinovic A, Izadjoo M, Roistacher M, Pavlov V, Tzivin V, Halle D, Pan H, Trink B, Gure AO, Nissan A. *Differential expression of colon cancer associated transcript1 (CCAT1) along the colonic adenoma-carcinoma sequence.* **BMC Cancer.** 2013; 13:196.

259. He X, Tan X, Wang X, Jin H, Liu L, Ma L, Yu H, Fan Z. *C-Myc-activated long noncoding RNA CCAT1 promotes colon cancer cell proliferation and invasion.* **Tumour Biol.** 2014; 35:12181-8.

260. Cai Y, He J, Zhang D. *Long noncoding RNA CCAT2 promotes breast tumor growth by regulating the Wnt signaling pathway.* **Onco Targets Ther.** 2015; 8:2657-64.

261. Qiu M, Xu Y, Yang X, Wang J, Hu J, Xu L, Yin R. *CCAT2 is a lung adenocarcinoma-specific long non-coding RNA and promotes invasion of non-small cell lung cancer.* **Tumour Biol.** 2014; 35:5375-80.

262. Wang J, Qiu M, Xu Y, Li M, Dong G, Mao Q, Yin R, Xu L. *Long noncoding RNA CCAT2 correlates with smoking in esophageal squamous cell carcinoma.* **Tumour Biol.** 2015; 36:5523-8.

263. Zhang N, Li X, Wu CW, Dong Y, Cai M, Mok MT, Wang H, Chen J, Ng SS, Chen M, Sung JJ, Yu J. *microRNA-7 is a novel inhibitor of YY1 contributing to colorectal tumorigenesis.* **Oncogene.** 2013; 32:5078-88.

264. Barbagallo D, Condorelli A, Ragusa M, Salito L, Sammito M, Banelli B, Caltabiano R, Barbagallo G, Zappalà A, Battaglia R, Cirnigliaro M, Lanzafame S, Vasquez E, Parenti R, Cicirata F, Di Pietro C, Romani M, Purrello M. *Dysregulated miR-671-5p / CDR1-AS / CDR1 / VSNL1 axis is involved in glioblastoma multiforme.* **Oncotarget.** 2016; 7:4746-59.

265. Wu ZH, Wang XL, Tang HM, Jiang T, Chen J, Lu S, Qiu GQ, Peng ZH, Yan DW. *Long non-coding RNA HOTAIR is a powerful predictor of metastasis and poor prognosis and is associated with epithelial-mesenchymal transition in colon cancer.* **Oncol Rep.** 2014; 32:395-402.

266. Ren Z, Wang W, Li J. *Identifying molecular subtypes in human colon cancer using gene expression and DNA methylation microarray data.* **Int J Oncol.** 2016; 48:690-702.

267. Tang L, Shen H, Li X, Li Z, Liu Z, Xu J, Ma S, Zhao X, Bai X, Li M, Wang Q, Ji J. *MiR-125a-5p decreases after long non-coding RNA HOTAIR knockdown to promote cancer cell apoptosis by releasing caspase 2.* **Cell Death Dis.** 2016; 7:e2137.

268. Pádua Alves C, Fonseca AS, Muys BR, de Barros E Lima Bueno R, Bürger MC, de Souza JE, Valente V, Zago MA, Silva WA Jr. *Brief report: The lincRNA Hotair is required for epithelial-to-mesenchymal transition and stemness maintenance of cancer cell lines.* **Stem Cells.** 2013; 31:2827-32.

269. Dou J, Ni Y, He X, Wu D, Li M, Wu S, Zhang R, Guo M, Zhao F. *Decreasing lncRNA HOTAIR expression inhibits human colorectal cancer stem cells.* **Am J Transl Res.** 2016; 8:98-108.

270. Ji Q, Zhang L, Liu X, Zhou L, Wang W, Han Z, Sui H, Tang Y, Wang Y, Liu N, Ren J, Hou F, Li Q. *Long non-coding RNA MALAT1 promotes tumour growth and metastasis in colorectal cancer through binding to SFPQ and releasing oncogene PTBP2 from SFPQ/PTBP2 complex.* **Br J Cancer.** 2014; 111:736-48.

271. Zheng HT, Shi DB, Wang YW, Li XX, Xu Y, Tripathi P, Gu WL, Cai GX, Cai SJ. *High expression of lncRNA MALAT1 suggests a biomarker of poor prognosis in colorectal cancer.* **Int J Clin Exp Pathol.** 2014; 7:3174-81.

272. Han Y, Wu Z, Wu T, Huang Y, Cheng Z, Li X, Sun T, Xie X, Zhou Y, Du Z. *Tumor-suppressive function of long noncoding RNA MALAT1 in glioma cells by downregulation of MMP2 and inactivation of ERK/MAPK signaling.* **Cell Death Dis.** 2016; 7:e2123.

273. Xu S, Sui S, Zhang J, Bai N, Shi Q, Zhang G, Gao S, You Z, Zhan C, Liu F, Pang D. *Downregulation of long noncoding RNA MALAT1 induces epithelial-to-mesenchymal transition via the PI3K-AKT pathway in breast cancer.* **Int J Clin Exp Pathol.** 2015; 8:4881-91.

274. Sun J, Ding C, Yang Z, Liu T, Zhang X, Zhao C, Wang J. *The long non-coding RNA TUG1 indicates a poor prognosis for colorectal cancer and promotes metastasis by affecting epithelial-mesenchymal transition.* **J Transl Med.** 2016; 14:42.

275. Wang L, Zhao Z, Feng W, Ye Z, Dai W, Zhang C, Peng J, Wu K. *Long non-coding RNA TUG1 promotes colorectal cancer metastasis via EMT pathway.* **Oncotarget.** 2016; 7:51713-51719.

276. Li J, Zhang M, An G, Ma Q. *LncRNA TUG1 acts as a tumor suppressor in human glioma by promoting cell apoptosis.* **Exp Biol Med (Maywood).** 2016; 241:644-9.

277. Zhang EB, Yin DD, Sun M, Kong R, Liu XH, You LH, Han L, Xia R, Wang KM, Yang JS, De W, Shu YQ, Wang ZX. *P53-regulated long non-coding RNA TUG1 affects cell proliferation in human non-small cell lung cancer, partly through epigenetically regulating HOXB7 expression.* **Cell Death Dis.** 2014; 5:e1243.

278. Liao WT, Jiang D, Yuan J, Cui YM, Shi XW, Chen CM, Bian XW, Deng YJ, Ding YQ. *HOXB7 as a prognostic factor and mediator of colorectal cancer progression.* **Clin Cancer Res.** 2011; 17:3569-78.

279. Tao K, Yang J, Hu Y, Sun Y, Tan Z, Duan J, Zhang F, Yan H, Deng A. *Clinical significance of urothelial carcinoma associated 1 in colon cancer.* **Int J Clin Exp Med.** 2015; 8:21854-60.

280. Ni B, Yu X, Guo X, Fan X, Yang Z, Wu P, Yuan Z, Deng Y, Wang J, Chen D, Wang L. *Increased urothelial cancer associated 1 is associated with tumor proliferation and metastasis and predicts poor prognosis in colorectal cancer.* **Int J Oncol.** 2015; 47:1329-38.

281. Bian Z, Jin L, Zhang J, Yin Y, Quan C, Hu Y, Feng Y, Liu H, Fei B, Mao Y, Zhou L, Qi X, Huang S, Hua D, Xing C, Huang Z. *LncRNA-UCA1 enhances cell proliferation and 5-fluorouracil resistance in colorectal cancer by inhibiting miR-204-5p.* **Sci Rep.** 2016; 6:23892.

282. Kahlert C, Lahes S, Radhakrishnan P, Dutta S, Mogler C, Herpel E, Brand K, Steinert G, Schneider M, Mollenhauer M, Reissfelder C, Klupp F, Fritzmann J, Wunder C, Benner A, Kloor M, Huth C, Contin P, Ulrich A, Koch M, Weitz J. *Overexpression of ZEB2 at the invasion front of colorectal cancer is an independent prognostic marker and regulates tumor invasion in vitro.* **Clin Cancer Res.** 2011; 17:7654-63.

283. Sun Z, Zhang Z, Liu Z, Qiu B, Liu K, Dong G. *MicroRNA-335 inhibits invasion and metastasis of colorectal cancer by targeting ZEB2.* **Med Oncol.** 2014; 31:982.

284. Zheng YB, Luo HP, Shi Q, Hao ZN, Ding Y, Wang QS, Li SB, Xiao GC, Tong SL. *miR-132 inhibits colorectal cancer invasion and metastasis via directly targeting ZEB2.* **World J Gastroenterol.** 2014; 20:6515-22.

285. Geng L, Chaudhuri A, Talmon G, Wisecarver JL, Are C, Brattain M, Wang J. *MicroRNA-192 suppresses liver metastasis of colon cancer.* **Oncogene.** 2014; 33:5332-40.

286. Li Y, Zheng Q, Bao C, Li S, Guo W, Zhao J, Chen D, Gu J, He X, Huang S. *Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis.* **Cell Res.** 2015; 25:981-4.

287. Curtin JF, Cotter TG. *JNK regulates HIPK3 expression and promotes resistance to Fas-mediated apoptosis in DU 145 prostate carcinoma cells.* **J Biol Chem.** 2004; 279:17090-100.

288. Zheng Q, Bao C, Guo W, Li S, Chen J, Chen B, Luo Y, Lyu D, Li Y, Shi G, Liang L, Gu J, He X, Huang S. *Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs.* **Nat Commun.** 2016; 7:11215.

289. Isin M, Ozgur E, Cetin G, Erten N, Aktan M, Gezer U, Dalay N. *Investigation of circulating lncRNAs in B-cell neoplasms.* **Clin Chim Acta.** 2014; 431:255-9.

290. Santoro M, Nociti V, Lucchini M, De Fino C, Losavio FA, Mirabella M. *Expression Profile of Long Non-Coding RNAs in Serum of Patients with Multiple Sclerosis.* **J Mol Neurosci.** 2016; 59:18-23.

291. Ma B, Li M, Zhang L, Huang M, Lei JB, Fu GH, Liu CX, Lai QW, Chen QQ, Wang YL. *Upregulation of long non-coding RNA TUG1 correlates with poor prognosis and disease status in osteosarcoma.* **Tumour Biol.** 2015; 37:4445-55.

292. Gao J, Cao R, Mu H. *Long non-coding RNA UCA1 may be a novel diagnostic and predictive biomarker in plasma for early gastric cancer.* **Int J Clin Exp Pathol.** 2015; 8:12936-42.

293. Kamel MM, Matboli M, Sallam M, Montasser IF, Saad AS, El-Tawdi AH. *Investigation of long noncoding RNAs expression profile as potential serum biomarkers in patients with hepatocellular carcinoma.* **Transl Res.** 2016; 168:134-45.

294. Dong L, Qi P, Xu MD, Ni SJ, Huang D, Xu QH, Weng WW, Tan C, Sheng WQ, Zhou XY, Du X. *Circulating CUDR, LSINCT-5 and PTENP1 long noncoding RNAs in sera distinguish patients with gastric cancer from healthy controls.* **Int J Cancer.** 2015; 137:1128-35.

295. Leucci E, Patella F, Waage J, Holmstrøm K, Lindow M, Porse B, Kauppinen S, Lund AH. *microRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus.* **Sci Rep.** 2013; 3:2535.

296. Hansen TB, Wiklund ED, Bramsen JB, Villadsen SB, Statham AL, Clark SJ, Kjems J. *miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA.* **EMBO J.** 2011; 30:4414-22.

297. Wang J, Liu X, Wu H, Ni P, Gu Z, Qiao Y, Chen N, Sun F, Fan Q. *CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer.* **Nucleic Acids Res.** 2010; 38:5366-83.

298. Hämmerle M, Gutschner T, Uckelmann H, Ozgur S, Fiskin E, Gross M, Skawran B, Geffers R, Longerich T, Breuhahn K, Schirmacher P, Stoecklin G, Diederichs S. *Posttranscriptional destabilization of the liver-specific long noncoding RNA HULC by the IGF2 mRNA-binding protein 1 (IGF2BP1).* **Hepatology.** 2013; 58:1703-12.

299. Hu G, Lou Z, Gupta M. *The long non-coding RNA GAS5 cooperates with the eukaryotic translation initiation factor 4E to regulate c-Myc translation.* **PLoS One.** 2014; 9:e107016.

300. Yang F, Xue X, Zheng L, Bi J, Zhou Y, Zhi K, Gu Y, Fang G. *Long non-coding RNA GHET1 promotes gastric carcinoma cell proliferation by increasing c-Myc mRNA stability.* **FEBS J.** 2014; 281:802-13.

301. Liu X, Li D, Zhang W, Guo M, Zhan Q. *Long non-coding RNA gadd7 interacts with TDP-43 and regulates Cdk6 mRNA decay.* **EMBO J.** 2012; 31:4415-27.

302. Zhuang Y, Nguyen HT, Burow ME, Zhuo Y, El-Dahr SS, Yao X, Cao S, Flemington EK, Nephew KP, Fang F, Collins-Burow B, Rhodes LV, Yu Q, Jayawickramarajah J, Shan B. *Elevated expression of long intergenic non-coding RNA HOTAIR in a basal-like variant of MCF-7 breast cancer cells.* **Mol Carcinog.** 2015; 54:1656-67.

303. Wu XS, Wang XA, Wu WG, Hu YP, Li ML, Ding Q, Weng H, Shu YJ, Liu TY, Jiang L, Cao Y, Bao RF, Mu JS, Tan ZJ, Tao F, Liu YB. *MALAT1 promotes the proliferation and metastasis of gallbladder cancer cells by activating the ERK/MAPK pathway.* **Cancer Biol Ther.** 2014; 15:806-14.

304. Wang F, Ying HQ, He BS, Pan YQ, Deng QW, Sun HL, Chen J, Liu X, Wang SK. *Upregulated lncRNA-UCA1 contributes to progression of hepatocellular carcinoma through inhibition of miR-216b and activation of FGFR1/ERK signaling pathway.* **Oncotarget.** 2015; 6:7899-917.

305. Stein U, Walther W, Arlt F, Schwabe H, Smith J, Fichtner I, Birchmeier W, Schlag PM. *MACC1, a newly identified key regulator of HGF-MET signaling, predicts colon cancer metastasis.* **Nat Med.** 2009; 15:59-67.

306. Zhao Y, Dai C, Wang M, Kang H, Lin S, Yang P, Liu X, Liu K, Xu P, Zheng Y, Li S, Dai Z. *Clinicopathological and prognostic significance of metastasis-associated in colon cancer-1 (MACC1) overexpression in colorectal cancer: a meta-analysis.* **Oncotarget.** 2016; 7:62966-62975.

307. Ashktorab H, Hermann P, Nouraie M, Shokrani B, Lee E, Haidary T, Brim H, Stein U. *Increased MACC1 levels in tissues and blood identify colon adenoma patients at high risk.* **J Transl Med.** 2016; 14:215.

308. Tang J, Chen JX, Chen L, Tang JY, Cui Z, Liu CH, Wang Z. *Metastasis associated in colon cancer 1 (MACC1) promotes growth and metastasis processes of colon cancer cells.* **Eur Rev Med Pharmacol Sci.** 2016; 20:2825-34.

309. Tajrishi MM, Tuteja R, Tuteja N. *Nucleolin: The most abundant multifunctional phosphoprotein of nucleolus.* **Commun Integr Biol.** 2011; 4:267-75.

310. Berger CM, Gaume X, Bouvet P. *The roles of nucleolin subcellular localization in cancer.* **Biochimie.** 2015; 113:78-85.

311. Hammoudi A, Song F, Reed KR, Jenkins RE, Meniel VS, Watson AJ, Pritchard DM, Clarke AR, Jenkins JR. *Proteomic profiling of a mouse model of acute intestinal Apc deletion leads to identification of potential novel biomarkers of human colorectal cancer (CRC).* **Biochem Biophys Res Commun.** 2013; 440:364-70.

312. Ropponen K, Tammi M, Parkkinen J, Eskelinen M, Tammi R, Lipponen P, Agren U, Alhava E, Kosma VM. *Tumor cell-associated hyaluronan as an unfavorable prognostic factor in colorectal cancer.* **Cancer Res.** 1998; 58:342-7.

313. Wang TP, Pan YR, Fu CY, Chang HY. *Down-regulation of UDP-glucose dehydrogenase affects glycosaminoglycans synthesis and motility in HCT-8 colorectal carcinoma cells.* **Exp Cell Res.** 2010; 316:2893-902.

314. Xu M, Kuang Y, Wang M, Han X, Yang Q. *A microRNA expression signature as a predictor of survival for colon adenocarcinoma.* **Neoplasma.** 2017;64(1):56-64.

315. Xu XT, Xu Q, Tong JL, Zhu MM, Nie F, Chen X, Xiao SD, Ran ZH. *MicroRNA expression profiling identifies miR-328 regulates cancer stem cell-like SP cells in colorectal cancer.* **Br J Cancer.** 2012; 106:1320-30.

316. Lennox KA, Behlke MA. *Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides.* **Nucleic Acids Res.** 2016; 44:863-77.