



UNIVERSITÀ DEGLI STUDI DI CATANIA

DOTTORATO DI RICERCA IN “SISTEMI COMPLESSI PER LE SCIENZE FISICHE,
SOCIO-ECONOMICHE E DELLA VITA”

XXXII CICLO

DIPARTIMENTO DI FISICA E ASTRONOMIA “ETTORE MAJORANA”

DUILIA BREX

Identificazione di LINC00483 come potenziale oncosoppressore nel
carcinoma coloretale tramite un approccio combinato di biologia *in silico*
ed *in vitro*.

TESI DI DOTTORATO

Coordinatore del Dottorato: Prof. Andrea Rapisarda

Tutor: Prof. Marco Ragusa

Anno Accademico 2018 - 2019

Indice

Abstract	1
1. Introduzione.....	4
2. I Non-Coding RNA (ncRNA).....	6
2.1 I “Long Non-Coding RNA” (lncRNA)	8
2.1.1 Caratteristiche dei lncRNA	9
2.1.2 Classificazione dei lncRNA	11
2.1.3 Biogenesi dei lncRNA	13
2.1.4 Funzioni dei lncRNA.....	14
2.2 I “microRNA” (miRNA)	18
2.2.1 Biogenesi dei miRNA.....	19
2.2.2 Funzioni dei miRNA	22
2.3 I “Circular RNA” (circRNA).....	24
2.3.1 Proprietà dei circRNA	25
2.3.2 Biogenesi dei circRNA	25
2.3.3 Funzioni dei circRNA.....	27
3. Meccanismi di interazione tra ncRNA e loro influenza sulla biologia del cancro	30
3.1 I miRNA inducono la degradazione dei lncRNA	30
3.2 lncRNA come “esche” per i miRNA.....	35
3.3 CircRNA come “miRNA sponges”	41
4. Il Carcinoma del Colon-Retto (CRC).....	46
4.1 Patogenesi del CRC.....	49
4.2 Meccanismi epigenetici nel CRC: metilazione e modificazioni istoniche	60
4.3 Diagnosi e terapia.....	61
4.4 Ruolo dei miRNA nella patogenesi e nell’evoluzione del CRC.....	65
4.5 Impatto dei lncRNA sulla patobiologia del CRC.....	67
5. Il processo metastatico.....	78
5.1 Ruolo della Transizione Eepitelio-Mesenchima (EMT) nel processo metastatico.....	79
5.2 Coinvolgimento dei ncRNA nel processo metastatico	83
5.3 miRNA e lncRNA come potenziali candidati per la terapia anti-metastatica	85
6. Materiali e Metodi.....	89

6.1	Analisi Computazionale	89
6.1.1	Identificazione dei target dei lncRNA	91
6.2	Linee cellulari	94
6.3	Estrazione di RNA da cellule	95
6.4	Estrazione di RNA da tessuti FFPE.....	95
6.5	Estrazione di RNA da tessuti congelati inclusi in OCT e da biopsie fresche.....	96
6.6	Disegno dei primers per Real Time PCR.....	97
6.7	Analisi di espressione tramite Real-Time PCR	98
6.8	Trattamenti delle linee cellulari.....	99
6.8.1	Trasfezione delle HCT-116 con LINC00483.....	99
6.8.2	Saggio della vitalità cellulare.....	101
6.8.3	Saggio della migrazione cellulare.....	101
6.8.4	Trattamento delle HCT-116 con TGFβ-1.....	102
6.8.5	Trattamento delle HCT-116 con IL-6.....	103
6.8.6	Blocco del ciclo cellulare	104
6.8.7	Trattamento delle HCT-116 con l'inibitore delle MAPK U0126.....	104
7.	Risultati.....	106
7.1	Analisi Computazionale	106
7.2	Espressione dei lncRNA nelle biopsie CRC FFPE	111
7.3	Trattamenti delle linee cellulari.....	113
7.3.1	Il trattamento con TGFβ-1 induce la EMT in vitro nelle cellule HCT-116.....	113
7.3.2	Inibizione in vitro della pathway delle MAPK	116
7.4	Identificazione dei target di LINC00483.....	117
7.5	Analisi di espressione di LINC00483 e dei suoi geni target.....	129
7.5.1	Analisi di espressione di LINC00483 sui tessuti CRC inclusi in OCT e sulle biopsie fresche	129
7.5.2	Trasfezione delle HCT-116 con LINC00483 e analisi di espressione dei geni target.....	131
7.5.3	Analisi di espressione di LINC00483 e dei geni target sulle linee cellulari trattate con l'inibitore delle MAPK U0126	132
7.5.4	Blocco del ciclo cellulare: analisi di espressione di LINC00483 e dei geni target	135
7.5.5	Analisi di espressione di LINC00483 e dei geni target sulle linee cellulari trattate con IL-6	138
7.6	LINC00483 inibisce la capacità migratoria delle cellule HCT-116.....	141
7.7	Saggio della vitalità cellulare: LINC00483 non altera la vitalità delle cellule HCT-116.....	142

7.8 LINC00483 e HIGD2A: analisi di correlazione	143
8. Discussione.....	147
9. Bibliografia	160

Abstract

Nell'ultimo decennio si è assistito ad un *exploit* di studi di sequenziamento su larga scala del genoma umano che hanno portato ad importanti e inaspettate scoperte: solo il 2% del genoma dei mammiferi codifica mRNA, mentre circa il 70% di questo è pervasivamente trascritto in lunghi e piccoli "non-coding RNA" (ncRNA). I "long non-coding RNA" (lncRNA) sono molecole di RNA non codificanti lunghe più di 200 bp, con o senza 5'-cap e coda di poly(A). L'esatta funzione dei lncRNA e i loro meccanismi di azione necessitano ancora di ricerche approfondite. Tuttavia, è stato dimostrato che i lncRNA svolgono un importante ruolo nella regolazione dell'espressione genica in varie malattie, incluso il cancro. Nel lavoro presentato in questa tesi abbiamo utilizzato un approccio combinato di biologia computazionale e sperimentale per identificare lncRNA la cui alterata espressione fosse associata alle caratteristiche clinico-patologiche più gravi del cancro coloretale (CRC). L'analisi di espressione su biopsie di CRC dei lncRNA precedentemente selezionati mediante l'analisi computazionale ci ha permesso di identificare un lncRNA - LINC00483 - che risultava sotto-espresso in maniera statisticamente significativa nei tessuti tumorali rispetto alla adiacente mucosa sana. Inoltre, questo lncRNA risultava anche *down*-regolato nei tumori metastatici rispetto ai primitivi, suggerendo un'associazione tra la sotto-espressione di LINC00483 e il fenotipo metastatico. La funzione di LINC00483 sembra essere associata all'inibizione della proliferazione e della transizione epitelio - mesenchima (EMT). La sua espressione, infatti, diminuiva nella linea cellulare di CRC HCT-116 trattata con TGF β -1, noto induttore della EMT, mentre aumentava nelle cellule

trattate con l'inibitore delle MAPK U0126, suggerendo così un'associazione tra la *down*-regolazione di LINC00483 e la trasformazione fenotipica da cellula epiteliale a mesenchimale, e tra l'*up*-regolazione di questo lncRNA ed il blocco del *signalling* iperproliferativo controllato da ERK. Per comprendere quale fosse il ruolo ed i *target* molecolari di LINC00483 abbiamo eseguito esperimenti di sovra-espressione *in vitro* tramite vettore d'espressione. Il saggio della migrazione e il saggio di proliferazione hanno mostrato che la sovra-espressione di LINC00483 riduceva la migrazione nelle HCT-116, tuttavia non è stato osservato un suo effetto fenotipico diretto sulla proliferazione cellulare. Ipotizzando una funzione di "*miRNA sponge*" per LINC00483 abbiamo identificato, tramite approccio computazionale, i potenziali mRNA *target* indiretti di LINC00483 (la cui espressione fosse positivamente correlata con il lncRNA e che fossero legati dagli stessi miRNA) e saggiato la loro espressione nelle HCT-116 trasfettate con il vettore. Tutti i geni *target* saggiati risultavano *up*-regolati nelle cellule trasfettate, confermando un'associazione tra LINC00483 e questi trascritti. Analizzando l'espressione di questi potenziali *target* su linee cellulari nelle quali era stata indotta la EMT ed il blocco del ciclo cellulare abbiamo messo in evidenza un particolare gene *target* - HIGD2A (HIG1 hypoxia inducible domain family member 2A) - la cui espressione era strettamente associata a quella di LINC00483. Inoltre, l'espressione di HIGD2A era ridotta nei tessuti di CRC rispetto alla adiacente mucosa sana e correlava positivamente e in maniera statisticamente significativa con quella di LINC00483. In conclusione, i nostri risultati suggeriscono che LINC00483 possa essere identificato come un nuovo gene oncosoppressore nel CRC e possa esplicare la sua funzione tramite il controllo negativo di

processi legati alla EMT (*i.e.*, migrazione cellulare) e della proliferazione cellulare. LINC00483 potrebbe operare attraverso una complessa *network* di interazione tra molecole di RNA nella quale si trova coinvolto il gene HIGD2A (così come anche altri dei geni *target* analizzati in questo studio), sebbene i meccanismi molecolari alla base della sua azione debbano essere chiariti ulteriormente.

1. Introduzione

Nell'ultimo decennio si è assistito ad un *exploit* di studi di sequenziamento su larga scala del genoma umano che hanno portato ad importanti e inaspettate scoperte. Contrariamente all'aspettativa iniziale che gli organismi più complessi contenessero un numero maggiore di geni, è ormai chiaro che uomo e topo hanno approssimativamente lo stesso numero di geni codificanti proteine del piccolo nematode *Caenorhabditis elegans*, la maggior parte dei quali sono ortologhi, e che tutti gli organismi pluricellulari sequenziati ad oggi hanno meno geni codificanti proteine rispetto ad alcuni semplici eucarioti unicellulari¹. Sulla base di tali analisi il numero di geni non aumenta proporzionalmente alla complessità di un organismo tanto quanto ci si aspettava inizialmente. Ad esempio, il genoma di *Drosophila melanogaster* contiene solo il doppio dei geni di alcune specie batteriche, sebbene l'organizzazione del suo genoma sia molto più complessa rispetto a questi ultimi. Tali evidenze suggeriscono che il numero dei geni codificanti proteine da solo non è sufficiente a spiegare la complessità degli eucarioti superiori (Figura 1).

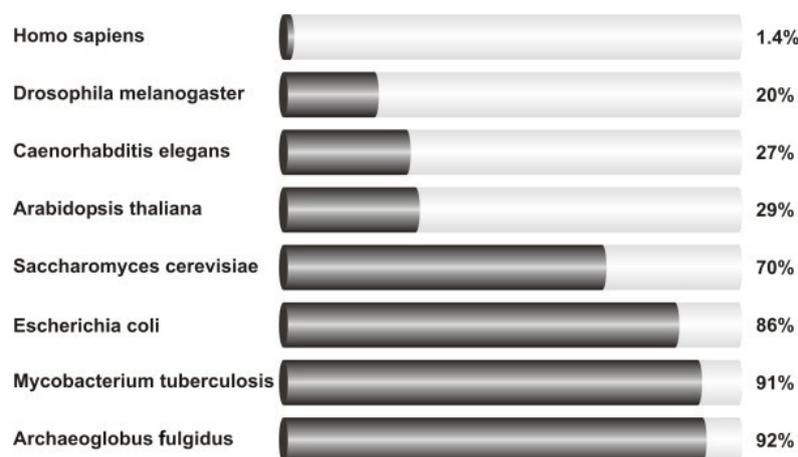


Figura 1. Percentuale di geni codificanti proteine in diversi genomi eucariotici e batterici (da Sana et al., J Transl Med. 2012).

Questo paradosso può essere, in parte, spiegato tramite lo *splicing* alternativo dei pre-mRNA, in cui diverse specie di mRNA, codificanti diverse isoforme proteiche, possono derivare da un singolo gene. In più, alcune modifiche post-traduzionali contribuiscono alla complessità e diversità delle specie proteiche². Una ulteriore spiegazione per questo apparente paradosso proviene da due inaspettate scoperte: 1) che la complessità biologica è correlata alla porzione del genoma non-codificante per proteine. Dalle analisi genomiche appare, infatti, evidente che all'aumentare della complessità di un organismo, il contributo della porzione codificante del suo genoma diminuisce (Figura 2); 2) che solo il 2% del genoma dei mammiferi codifica mRNA, mentre la gran parte è pervasivamente trascritto in lunghi e piccoli "non-protein-coding RNA" (ncRNA).

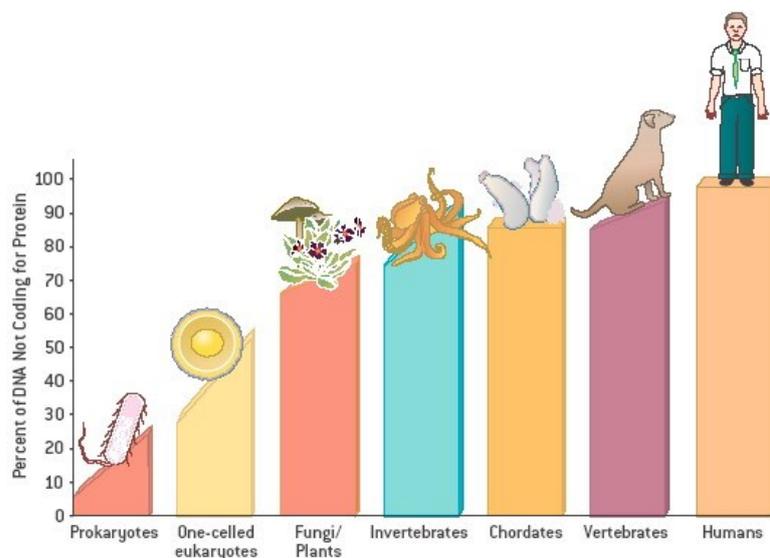


Figura 2. Correlazione tra complessità degli organismi e regioni non codificanti del genoma: le sequenze non codificanti proteine coprono solo una piccola frazione del DNA dei procarioti. Tra gli eucarioti, così come aumenta la loro complessità, allo stesso modo aumenta, proporzionalmente, la porzione del DNA non codificante (da Abdelmohsen K. *Encyclopedia of Cancer*, Springer.2015).

Queste scoperte hanno completamente rivoluzionato la tradizionale concezione dell'RNA come semplice intermedio tra DNA e proteine, ed implicano che la maggior parte del genoma - a lungo considerato come "*junk DNA*" - codifichi, invece, per RNA funzionali che dirigono lo sviluppo degli organismi complessi³. È ormai largamente accettato che gli RNA non codificanti costituiscono elementi molto importanti di tutti i meccanismi cellulari e che, insieme ai fattori di trascrizione, sono responsabili dello stretto controllo del *repertoire* di geni espressi in virtualmente tutti gli organismi viventi⁴. Appare, quindi, chiaro che il *signalling* degli RNA è centrale in tutti i meccanismi genetici negli eucarioti, inclusi il silenziamento trascrizionale e post-trascrizionale, la disgenesi degli ibridi, la compensazione del dosaggio del cromosoma-X e l'esclusione allelica, la riprogrammazione delle cellule germinali - che coinvolgono tutti processi epigenetici¹. I ncRNA contribuiscono, quindi, alle complesse *network* necessarie per regolare le funzioni cellulari e potrebbero rappresentare la risposta definitiva alla complessità del genoma³.

2. I Non-Coding RNA (ncRNA)

Mentre i primi metodi basati sul clonaggio hanno rivelato più di 7.000 geni umani tra il 1970 e il 1980, le analisi su larga scala delle sequenze ESTs (*Expressed Sequence Tags*) nel 1990 hanno suggerito che il numero stimato di geni umani risieda in un *range* tra 35.000 e 100.000. Il completamento del progetto genoma umano ha ridotto considerevolmente il campo evidenziando un numero sorprendentemente piccolo di geni codificanti proteine, che è oggi convenzionalmente indicato come inferiore ai 25.000. Mentre il numero di geni

codificanti proteine ha mantenuto un ampio consenso (20.000-22.000), recenti studi sul trascrittoma umano hanno rivelato un numero sbalorditivo di *non-coding RNA* (ncRNA). Questi elementi trascritti, che mancano della capacità di codificare per proteine, sono incredibilmente abbondanti in tutti gli organismi studiati ad oggi, dal lievito all'uomo. Inoltre, nell'ultimo decennio, numerosi studi hanno dimostrato che i ncRNA hanno distinte funzioni biologiche e operano attraverso precisi meccanismi. Nonostante si tratti di una classe di molecole ampiamente studiata non è ancora chiaro se la trascrizione dei ncRNA rifletta un fenomeno biologico funzionale o questi rappresentino sottoprodotti di un sistema trascrizionale caratterizzato da qualche falla⁵. Quasi il 98% del genoma umano non codifica per proteine. Questo DNA non codificante è stato a lungo considerato come un terreno arido, senza alcuna apparente funzionalità nella sintesi delle proteine e dunque indicato come "*junk DNA*". Tuttavia, il DNA intergenico, non-codificante, è stato rivalutato nel tempo in quanto tesoro di informazioni che possono essere decifrate nella forma di elementi nucleotidici "non genici" (e.g., sequenze ripetute, sequenze trasponibili, elementi interspersi) e di diversi *non-coding RNA*⁶. La più comune classificazione dei ncRNA è basata sulla loro lunghezza. Questa prevede la loro suddivisione in due classi: 1) *long non-coding RNA* (lncRNA), più lunghi di 200 nucleotidi; e 2) *small non-coding RNA*, la cui lunghezza è pari o inferiore ai 200 nucleotidi [*i.e.*, *microRNA* (miRNA), *small interfering RNA* (siRNA), *PIWI-interacting RNA* (piRNA), *competitive endogeneous RNA* (ceRNA)]. È stato proposto un ulteriore criterio per catalogare i ncRNA. Questi possono essere suddivisi in due classi sulla base delle loro caratteristiche funzionali: ncRNA costitutivi e ncRNA regolatori. I ncRNA costitutivi sono costitutivamente espressi in tutte le cellule a

causa del loro coinvolgimento nella fisiologia della cellula: questi includono *transfer RNA* (tRNA), *ribosomal RNA* (rRNA), *small nuclear RNA* (snRNA), *small nucleolar RNA* (snoRNA) e *telomerase RNA*⁷. D'altra parte, i ncRNA regolatori possono essere espressi in specifici tipi cellulari, o durante definiti stadi dello sviluppo e del differenziamento cellulare e, infine, in risposta a stimoli esterni^{8, 9}. Questa categoria comprende miRNA, siRNA, lncRNA (le molecole di RNA più strettamente coinvolte nella biologia del cancro⁸), e piRNA. Si deve tuttavia rilevare che molti ncRNA mostrano caratteristiche che potrebbero permetterne l'assegnazione a diverse categorie, eludendo quindi la classificazione sistematica¹⁰.

2.1 I "Long Non-Coding RNA" (lncRNA)

Sebbene molte molecole di RNA siano più lunghe di 200 bp, il termine lncRNA [anche riportato come lincRNA (*long intergenic ncRNA*)], non può essere uniformemente applicato a tutte queste. Mentre il sistema di nomenclatura è ancora in evoluzione, il termine lncRNA si riferisce tipicamente a lunghe molecole di *non-coding RNA* poliadenilate, trascritte dalla RNA polimerasi II ed associate a caratteristiche epigenetiche comuni ai geni codificanti proteine, come la trimetilazione della lisina 4 dell'istone 3 (H3K4me3) al sito di inizio della trascrizione (TSS) e la trimetilazione della lisina 36 dell'istone 3 (H3K36me3) in ogni parte del gene¹¹.

2.1.1 Caratteristiche dei lncRNA

▪ Localizzazione nel genoma

I *long non-coding RNA* costituiscono la più ampia ed eterogenea classe di RNA non codificanti proteine. Essi includono trascritti che possono: essere localizzati in regioni intergeniche [*long intergenic non-coding RNA* (lincRNA)]; giacere all'interno di introni di geni codificanti proteine; sovrapporsi parzialmente alle *UTRs* (*untranslated regions*) o ai promotori di geni codificanti proteine; essere trascritti da pseudogeni e controllare l'espressione dei loro paraloghi funzionali codificanti proteine; essere trascritti da Regioni Ultra Conservate Trascritte (tUCR), che sono evolutivamente altamente conservate e possono essere localizzate in regioni intra - o intergeniche¹². La localizzazione genomica dei lncRNA sembra presentare una diretta associazione con la loro conservazione evolutiva. Ricerche scientifiche suggeriscono che la maggior parte dei lncRNA sia evolutivamente conservata sebbene in misura minore rispetto ai geni codificanti per proteine¹³.

▪ Meccanismi di azione: regolazione in *cis* e *trans*

La maggior parte dei lncRNA meglio caratterizzati ad oggi svolgono un ruolo funzionale nella regolazione dell'espressione genica, tipicamente a livello trascrizionale piuttosto che post-trascrizionale. Questo effetto si verifica agendo sia su geni vicini sul genoma (*cis*-regolazione) sia su geni genomicamente distanti (*trans*-regolazione)⁵. I lncRNA che agiscono in *trans* (*trans-acting* lncRNA) e i loro rispettivi *target* sono fisicamente localizzati in differenti *loci* sul genoma. Al contrario, i lncRNA

che agiscono in *cis* (*cis-acting* lncRNA) e i loro *target* sono localizzati sullo stesso *locus*, ma su filamenti opposti del DNA. I *cis-acting* lncRNA sono stati identificati per la prima volta nei virus, poi nei procarioti ed infine, negli eucarioti¹⁴. I lncRNA che agiscono da elementi *cis* regolatori controllano localmente l'espressione genica reclutando i complessi di modificazione degli istoni su specifiche regioni del genoma. Questo effetto può essere altamente specifico per un particolare gene, come la regolazione di IGF2 da parte dei lncRNA, oppure, può coinvolgere un'ampia regione cromosomica, come l'inattivazione del cromosoma X nelle femmine tramite XIST. Molti *cis-acting* lncRNA, inclusi H19, AIR e KCNQ1OT1, sono anche funzionalmente correlati tramite il loro coinvolgimento nella regolazione delle regioni soggette ad *imprinting* genomico. Tale coinvolgimento è critico per mantenere l'espressione specifica del gene parentale. In particolare, una regione soggetta ad *imprinting* genomico sul cromosoma 11 umano è stata ampiamente studiata per il ruolo dei lncRNA. Nell'uomo, i lncRNA meglio noti sono H19 e KCNQ1OT1, che sono espressi sull'allele materno e paterno, rispettivamente, e silenziano i geni IGF2 e KCNQ1. Come la maggior parte dei *cis-acting* lncRNA, i *trans-acting* lncRNA tipicamente facilitano la regolazione epigenetica dell'espressione genica. Tuttavia, poiché i *trans-acting* lncRNA possono agire su regioni fisicamente distinte del genoma, generalmente si ritiene che il trascritto maturo sia l'attore principale in questi casi, a differenza dei *cis-acting* lncRNA, come XIST, che possono agire tramite l'atto stesso della trascrizione⁵.

2.1.2 Classificazione dei lncRNA

La nomenclatura e i simboli di 399 lncRNA umani sono stati approvati dalla *HUGO-Gene Nomenclature Committee* (HGNC) (<https://www.genenames.org/cgi-bin/genefamilies/set/788>). I lncRNA sono classificati sulla base della loro struttura, funzione, localizzazione ed interazione con i geni codificanti proteine o altri elementi del DNA. Le strutture secondarie e terziarie dei lncRNA sono molto conservate rispetto alle loro strutture primarie. Lo studio della relazione tra struttura e funzione di queste molecole ad alto peso molecolare è ancora oggi una sfida poiché sono difficili da cristallizzare⁶. In linea di massima, i lncRNA possono essere suddivisi in 5 categorie, sulla base dei *loci* genomici da cui vengono trascritti. Questi includono le regioni intergeniche e i filamenti senso e antisenso dei geni codificanti proteine con i siti di inizio della trascrizione situati in introni, esoni, regioni promotore, *enhancers* e 3' o 5'UTR.

I lncRNA possono, quindi, essere classificati in:

- 1) *Long Intergenic non-coding RNA* (lincRNA): i lincRNA giacciono all'interno dell'intervallo genomico tra due geni (Figura 3A);
- 2) *Long non-coding RNA* intronici: originano esclusivamente dagli introni di geni codificanti proteine e, normalmente, vengono generati come risultato dello *splicing* alternativo (Figura 3B);
- 3) *Long non-coding RNA* senso: come implica il termine, questi lncRNA, anche noti come lncRNA esonici, sono sintetizzati dalle porzioni codificanti proteine dei geni, con gli esoni che si sovrappongono a quelli degli mRNA (Figura 3C);

4) *Natural Antisense Transcripts (NATs)*: i NATs sono trascritti in direzione opposta rispetto ai geni codificanti proteine (Figure 3D);

5) *Long non-coding RNA bidirezionali*: si tratta di trascritti che iniziano in stretta vicinanza (<1000 bp) ai siti di inizio dei geni codificanti proteine, e poi procedono in direzione opposta⁶ (Figure 3E).

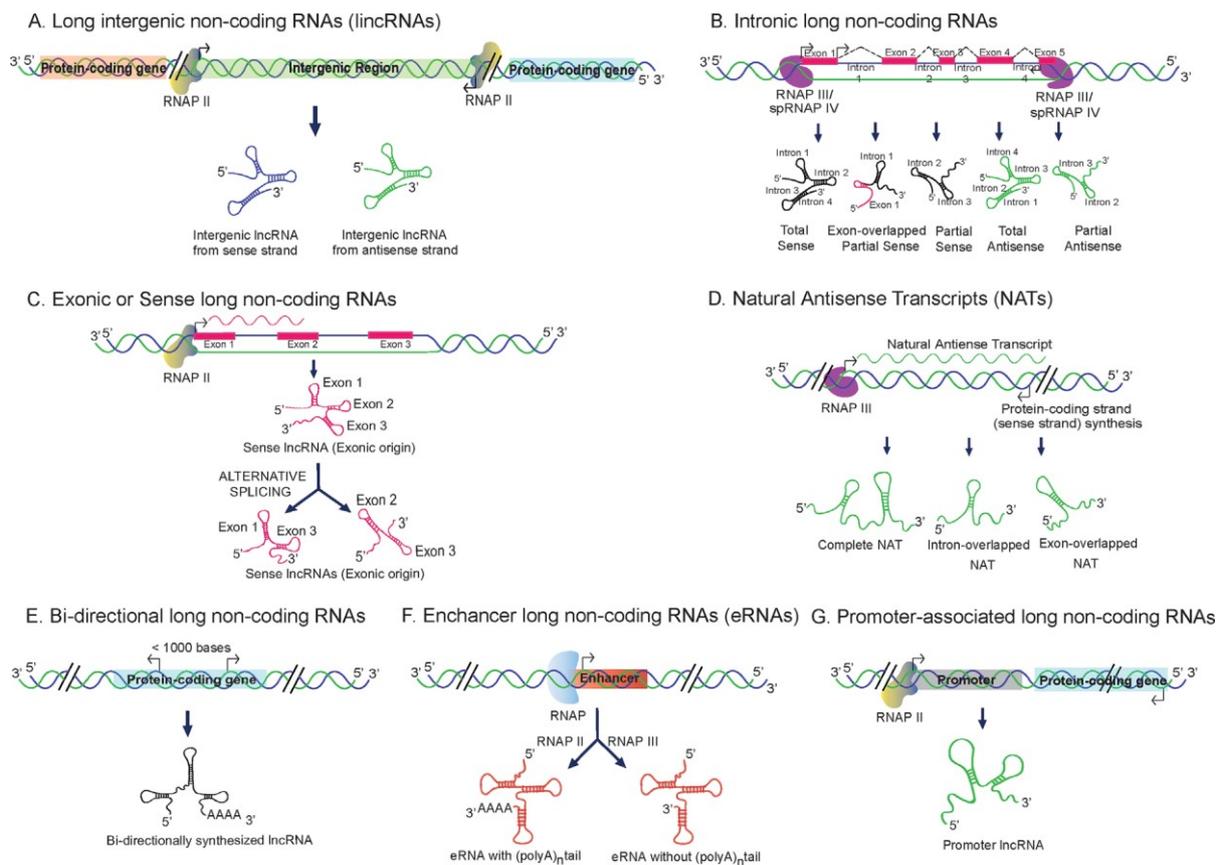


Figura 3. Classificazione dei lncRNA sulla base della loro localizzazione genomica. (A) Long intergenic non-coding RNA (lincRNA); (B) Intronic long non-coding RNA; (C) Exonic, or sense long non-coding RNA; (D) Natural antisense transcripts (NATs); (E) Bi-directional long non-coding; (F) Enhancer long non-coding RNA (eRNA); (G) Promoter long non-coding RNA (da Khandelwal et al., Mol Carcinog. 2015).

Le suddette classi possono essere ulteriormente raggruppate in base a 3 diversi meccanismi di regolazione. Primo, agendo da “competitors”, i lncRNA possono legare

le *Dna Binding Proteins* e impedire a queste ultime di legare il proprio *target*. Ad esempio, i lncRNA possono influenzare la metilazione del DNA inibendo il legame di DNMT1 [*DNA (cytosine-5) -methyltransferase 1*] al DNA che, in definitiva, influenza l'attivazione trascrizionale del gene *target*. Secondo, agendo come "recruiters", i lncRNA possono aumentare la metilazione del DNA reclutando regolatori epigenetici su alcuni siti *target*. In ultimo, i lncRNA possono agire da precursori dei miRNA tramite digestione da parte di specifiche RNAsi quali Dicer⁶.

2.1.3 Biogenesi dei lncRNA

I lncRNA possono essere trascritti dalle regioni intergeniche, esoniche o codificanti del genoma dall'enzima RNA polimerasi II. Soltanto una piccola parte di essi è trascritta dalle RNA polimerasi I e III^{15, 16}. Circa metà dei lncRNA viene processato a livello post-trascrizionale andando incontro alle medesime modifiche degli mRNA. Il pre-lncRNA presenta, infatti, l'estremità 3' poliadenilata, il *cap* di metil-guanosina all'estremità 5' e, spesso, subisce uno *splicing* alternativo¹⁷, sebbene alcuni lncRNA siano monoesonici e non poliadenilati^{18, 19}. La coda di Poly(A) protegge le estremità 3' dalla degradazione enzimatica; in sua assenza, è stata osservata la presenza di strutture alternative con funzione simile, come la tripla elica²⁰. In generale, i lncRNA di geni codificanti sono dotati di promotore e fattori di trascrizione propri oltre che motivi di DNA unici. Le modificazioni epigenetiche giocano un ruolo nella biogenesi dei lncRNA. La metilazione degli istoni svolge un ruolo importante nella regolazione

trascrizionale. La metilazione della lisina 4 dell'istone H3 (H3K4) è segno dell'attivazione trascrizionale, mentre la trimetilazione H3K27 indica il silenziamento genico. Vari lncRNA, inclusi HOTTIP, XIST, FIRRE, sono coinvolti nell'attivazione della trascrizione genica e nell'organizzazione della architettura nucleare⁶. Dall'altra parte, il reclutamento di lncRNA come i trascritti ALU o di *triplex* lncRNA-DNA può inibire la regolazione trascrizionale tramite il legame alla RNA polimerasi II²¹. Il legame di diversi fattori di trascrizione ai lncRNA forma un trascritto nascente che, infine, regola il processamento degli mRNA attraverso lo *splicing* alternativo. Questo legame dei lncRNA all'mRNA può aumentare o inibire la traduzione o può promuovere la degradazione dell'mRNA. I lncRNA maturi possono essere presenti nel nucleo o nel citoplasma. Sebbene i lncRNA citoplasmatici non vengano tradotti, sono stati identificati piccoli peptidi che sono generati dai lncRNA tramite la loro associazione ai ribosomi. Alcune ricerche indicano che gli pseudogeni trascrizionalmente attivi possono anche produrre queste molecole, oppure queste possono essere trascritte dal promotore o da regioni intergeniche²².

2.1.4 Funzioni dei lncRNA

La maggior parte dei lncRNA, sebbene non codificanti in natura, è coinvolta nei processi patologici e in diversi processi biologici. L'esatta funzione dei lncRNA e i loro meccanismi di azione necessitano ancora di ricerche approfondite. Tuttavia, in generale, è stato dimostrato che i lncRNA svolgono un importante ruolo nella

regolazione dell'espressione genica in varie malattie, incluso il cancro⁶. Diversamente dai miRNA, che reprimono l'espressione genica attraverso un comune meccanismo che coinvolge il complesso RISC, i lncRNA presentano un'ampia varietà di meccanismi di azione attraverso i quali sono implicati nella modulazione della regolazione epigenetica, nello *splicing* alternativo, nella localizzazione e attività delle proteine¹² (Figura 4). Il meccanismo dello *splicing* alternativo può essere ricondotto a tre modalità di azione. Anzitutto, i lncRNA interagiscono con specifici fattori di *splicing*, formano quindi un *duplex* RNA-RNA con le molecole di pre-mRNA e, infine, influenzano il rimodellamento della cromatina, completando lo *splicing* dei geni *target*²³.

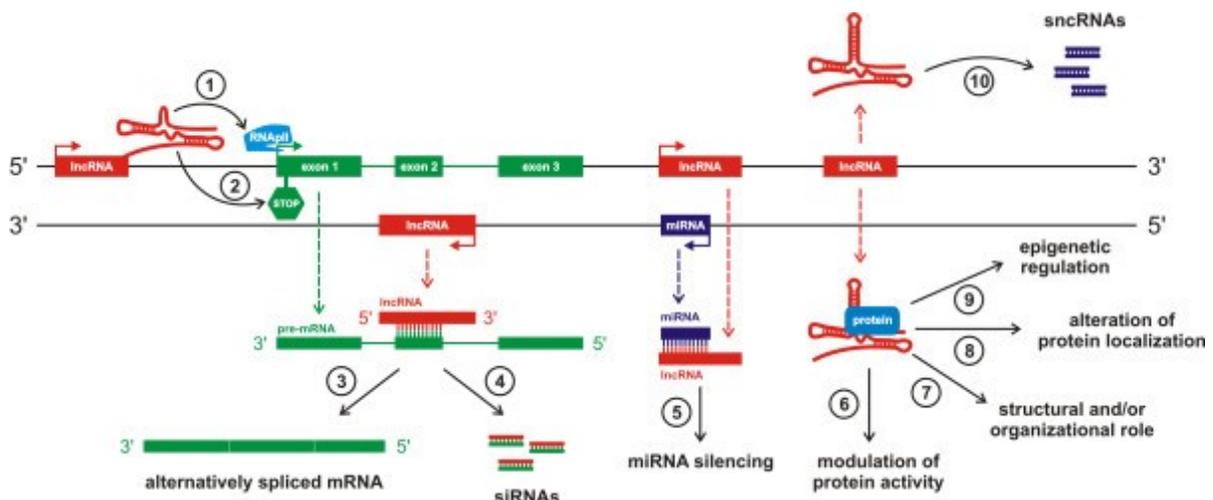


Figura 4. Rappresentazione schematica dei meccanismi di azione dei lncRNA: i lncRNA trascritti da un promotore a monte possono (1) influenzare negativamente o (2) positivamente l'espressione del gene a valle inibendo il reclutamento della RNA polimerasi II e/o inducendo il rimodellamento della cromatina, rispettivamente. I lncRNA sono in grado di ibridizzare con i pre-mRNA ed impedire il riconoscimento dei siti di splice da parte dello spliceosoma, dando quindi un trascritto alternativo (3). Alternativamente, l'ibridazione tra i trascritti senso e antisenso può consentire all'enzima DICER di generare i siRNA (4). Il legame del lncRNA al miRNA determina il silenziamento funzionale del miRNA (5). Il complesso formato da lncRNA e specifiche proteine può modulare l'attività delle proteine (6), è coinvolto nella struttura e nell'organizzazione delle cellule (7), altera la localizzazione delle proteine nella cellula (8), ed influenza i processi epigenetici (9). Infine, i lncRNA possono essere processati in piccoli RNA (10) (da Sana et al., J Transl Med. 2012).

I lncRNA possono esplicare le loro funzioni attraverso 4 differenti modalità di azione (Figura 5):

- **lncRNA come molecole segnale**

La funzione proposta per alcune molecole di lncRNA come segnali molecolari deriva dalla constatazione che la loro trascrizione potrebbe servire ad integrare segnali di sviluppo, a “interpretare” il contesto cellulare e rispondere a diversi *stimuli*. In alcuni casi i lncRNA controllano la trascrizione legando in maniera reversibile diversi componenti del macchinario trascrizionale, come la RNA Polimerasi II²⁴.

- **lncRNA come “esche” molecolari**

Le “esche” molecolari (*e.g.*, Gas5, PANDA) sono polinucleotidi che regolano negativamente un effettore precludendo l’accesso al DNA a proteine regolatrici e impedendo, dunque la trascrizione di specifici geni^{25, 26}. Gas5 è un lncRNA con una struttura a forcina (che ricorda la struttura degli elementi responsivi ai glucocorticoidi del DNA). In condizioni di *starvation* da fattori di crescita, Gas5 si lega al dominio di legame al DNA dei recettori per i glucocorticoidi (GR) agendo da “*Glucocorticoid Response Element (GRE)*” e competendo, quindi, con i GRE del DNA per il legame ai GR. Gas5 reprime, quindi, l’attivazione trascrizionale dei geni metabolici indotta dai GR, inibendo il legame di questi ultimi ai geni *target* GRE²⁵.

- **lncRNA come “guide molecolari”**

I lncRNA sono richiesti per la corretta localizzazione di specifiche proteine, inclusi i complessi ribonucleoproteici. *Homebox Antisense Intergenic RNA (HOTAIR)* è un

esempio di lncRNA guida, che dirige la localizzazione del *Polycomb Repressor Complex 2* (PRC2), modulando il silenziamento epigenetico di geni *target* nei processi associati al cancro²⁷.

▪ **lncRNA come “scaffold”**

I lncRNA possono agire da *scaffold* per l'assemblaggio di due o più proteine in un complesso funzionale, presentando quindi un ruolo strutturale. I lncRNA, chiamati *Telomere-associated ncRNA* (*Telomeric Repeat-Containing RNA*: TERRA), regolano negativamente la lunghezza dei telomeri probabilmente inibendo l'attività della telomerasi²⁸.

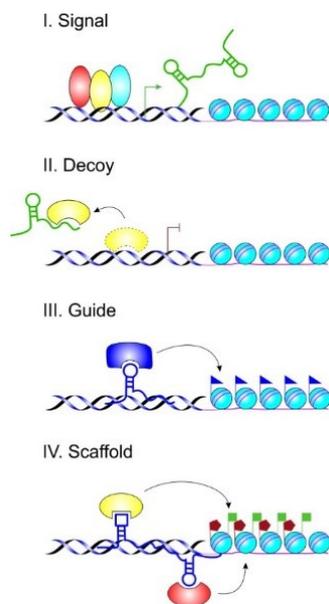


Figura 5. Rappresentazione schematica dei possibili ruoli regolativi dei lncRNA: (I) i lncRNA possono agire da segnali e regolare l'espressione genica; (II) i lncRNA possono allontanare i fattori di trascrizione o altre proteine dalla cromatina oppure agire da “esche” per i miRNA; (III) i lncRNA possono reclutare gli enzimi di rimodellamento della cromatina sui geni target e quindi, funzionare da guide; (IV) i lncRNA possono agire da scaffold per l'assemblaggio di più proteine in complessi ribonucleoproteici (da Sana et al., J Transl Med. 2012).

Oltre le funzioni sopra menzionate, è stato riportato che i lncRNA possono essere funzionali in alcune substrutture del cervello dei topi e hanno un qualche ruolo

associato ai fattori di trascrizione coinvolti nella pluripotenza²⁹. I *long intergenic non-coding RNA* (lincRNA) hanno un importante ruolo nel mantenimento dello stato pluripotente delle cellule. Studi sulle cellule staminali embrionali di topo hanno rivelato che il *knockdown* dei lincRNA influenza la regolazione genica³⁰.

2.2 I “microRNA” (miRNA)

I *microRNA* (miRNA) sono una classe di piccoli (19-24 nucleotidi di lunghezza) RNA endogeni, evolutivamente conservati, che fungono da regolatori post-trascrizionali dell'espressione genica. Essi agiscono principalmente legando la sequenza complementare presente nell'RNA messaggero (mRNA) *target*, interferendo col macchinario traduzionale, inibendo o alterando la produzione della proteina. Recenti studi hanno rivelato che, oltre ad inibire la traduzione, il legame del miRNA al proprio mRNA *target* può indurre il reclutamento e l'associazione di fattori che degradano l'mRNA, portando alla sua destabilizzazione con conseguente riduzione dei livelli di espressione. I miRNA sono stati scoperti nel 1993 da Lee e colleghi nel nematode *C. elegans*. In questo organismo, la *down-regolazione* della proteina LIN-14 è essenziale per la progressione dallo stadio larvare L1 allo stadio L2. Inoltre, è stato dimostrato che la *down-regolazione* di LIN-14 dipende dalla trascrizione di un secondo gene, chiamato *lin-4*. Sorprendentemente, il trascritto di *lin-4* non viene tradotto in una proteina biologicamente attiva. Esso dà, invece, origine a due piccoli RNA lunghi approssimativamente 21 e 61 nucleotidi. La sequenza più lunga forma una struttura a *stem-loop* e serve da precursore dell'RNA più corto. Più tardi, questo

gruppo insieme a Wightman *et al.* scoprì che l'RNA più corto aveva complementarità antisenso con diversi siti nel 3' UTR dell'mRNA di *lin-14*. Il legame tra queste regioni complementari causava una riduzione dei livelli della proteina LIN-14 senza determinare alcun cambiamento significativo dei livelli di espressione del suo mRNA. Questi due studi hanno portato alla luce un modello in cui l'appaiamento di basi si verifica tra diversi piccoli RNA *lin-4* e i siti complementari sul 3' UTR dell'mRNA di *lin-14*, inibendo quindi la traduzione di *lin-14* con conseguente progressione dallo stadio L1 allo stadio L2 durante lo sviluppo di *C. elegans*. Inizialmente si ritenne che questo nuovo modello di regolazione dell'espressione genica fosse un fenomeno esclusivo di *C. elegans*. Nel 2000, due gruppi separati scoprirono che un piccolo RNA, *let-7*, era essenziale per lo sviluppo dell'ultimo stadio larvare nella forma adulta, in *C. elegans*. Cosa più importante, omologhi di questo gene sono stati successivamente scoperti in molti altri organismi, incluso l'uomo. Nel periodo che seguì diversi laboratori clonarono numerosi *small RNA* da uomo, mosca e vermi. Questi RNA erano non codificanti, lunghi circa 19-24 nt, e derivavano da un precursore più lungo con una struttura a *stem-loop*. Molti furono trovati evolutivamente conservati tra le specie e mostravano specificità cellula-tipo. Il riconoscimento e la conferma dell'esistenza di questi piccoli RNA, oggi chiamati microRNA (miRNA), ha incoraggiato un'intensa ricerca allo scopo di identificare nuovi membri di questa famiglia³¹.

2.2.1 Biogenesi dei miRNA

I geni per i miRNA possono avere varie localizzazioni sul genoma. È stato stimato che approssimativamente il 50% dei miRNA sono espressi da trascritti non codificanti

proteine. Il resto è localizzato principalmente negli introni di geni codificanti ed è generalmente cotrascritto con i geni parentali e processato separatamente³².

I passaggi coinvolti nella biogenesi dei miRNA sono schematicamente rappresentati nella Figura 6. I trascritti codificanti per i miRNA sono inizialmente trascritti dalla RNA polimerasi II come lunghi miRNA primari (lunghe diverse centinaia di nucleotidi), e presentano il *cap* all'estremità 5' e la coda Poly(A) all'estremità 3'. Questi pri-miRNA possono essere sia non codificanti che codificanti (presenti all'interno dell'introne di un gene codificante). Il miRNA primario viene, in seguito, processato in un precursore, il pre-miRNA, lungo 70-120 nt, da un complesso multiproteico chiamato *Microprocessor* (Figura 6A). Questo complesso comprende una RNasi III nucleare di 160-kDa chiamata Droscha. Questo enzima è altamente conservato negli animali. Droscha dimerizza con un'altra *double-stranded RNA(dsRNA) binding protein*, chiamata *DiGeorge syndrome critical region gene 8* (DGCR8) o Pasha, per formare il complesso *Microprocessor* funzionale (Figura 6A). Il pre-miRNA con il tipico gruppo fosfato al 5' e circa 2 nt sporgenti al 3' è, in seguito, esportato nel citoplasma dall'esportina 5 (exp-5) (Figura 6B). Nel citoplasma, il pre-miRNA è, infine, processato in un miRNA *duplex* maturo di 18-23 nt da un'altra RNasi III, Dicer-1, con l'aiuto di *dsRNA-binding proteins* e di altre proteine (Figura 6 C, D). I due filamenti del miRNA vengono in seguito separati; questo processo dipende da vari fattori come l'asimmetria termodinamica del *duplex* e la stabilità dell'appaiamento all'estremità 5'. Un filamento, chiamato filamento guida, insieme alle proteine sopra menzionate e ad altre *Rna Binding Proteins* tra cui 6A TNRC6A (*trinucleotide repeat-containing gene 6*), si

associa con le proteine Argonata (AGO), formando un complesso proteico microribonucleare (miRNP) chiamato “RNA-Induced Silencing Complex” (RISC) (Figura 6E). Il filamento del miRNA con l’appaiamento più instabile al 5’ agisce solitamente da filamento guida, mentre il filamento con un appaiamento stabile all’estremità 5’ (noto come filamento *passenger* o miR*) è generalmente degradato o, in rari casi, si associa con le proteine AGO, consentendo ad entrambi i filamenti di servire da miRNA funzionali. Il filamento guida dirige il complesso sull’*mRNA target* attraverso la complementarità di sequenza e causa la sua repressione traduzionale³¹ (Figura 6F).

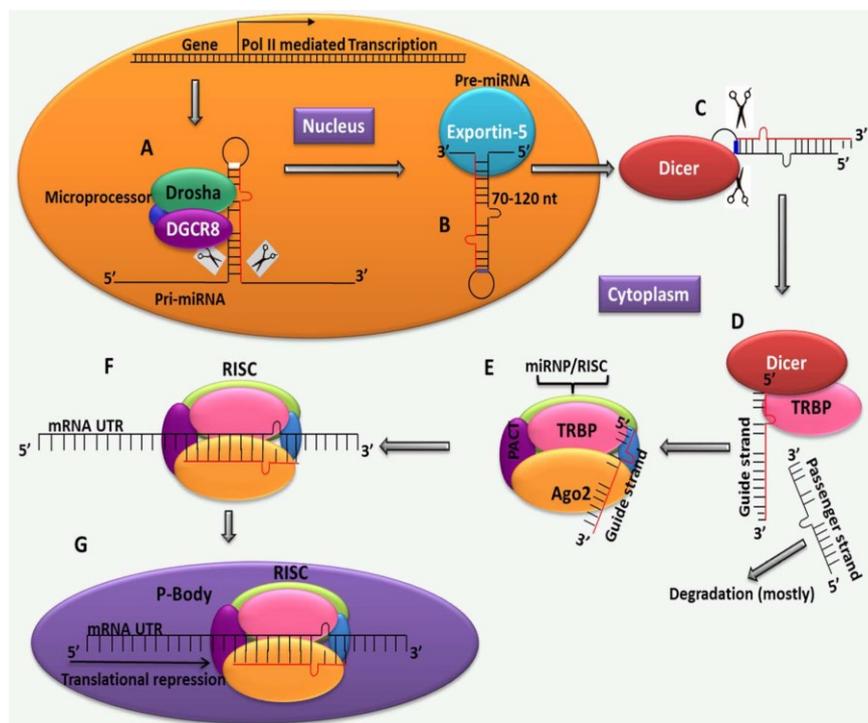


Figura 6. Biogenesi dei miRNA: i miRNA sono inizialmente trascritti dalla RNA polimerasi II come lunghi miRNA primari (pri-miRNA) (A). Nel nucleo, questi pri-miRNA sono nuovamente processati dal Microprocessor, un grande complesso che include l’RNasi II Drosha e l’RNA binding protein DGCR8, in pre-miRNA lunghi ~70-120-nucleotidi, con struttura ad hairpin (A). I pre-miRNA sono, in seguito, esportati nel citoplasma dall’esportina 5 (B), dove il loop del pre-miRNA viene tagliato dall’RNasi III Dicer per dare il miRNA maturo di ~18-23-nucleotidi (C). Una proteina cellulare chiamata transactivation response RNA binding protein (TRBP) (D) facilita l’associazione del complesso Dicer-miRNA col complesso RISC, che contiene le proteine Argonate 2 (Ago2), la protein kinase RNA activator (PACT), il trinucleotide repeat-containing gene 6A (TNRC6A), ed altre RNA binding proteins (E). Dopo l’incorporazione del miRNA maturo nel complesso RISC, il filamento “passeggero” è degradato, e il filamento guida è diretto al 3’UTR del’*mRNA target* per essere degradato (nel caso di perfetta complementarità tra coppie di basi) o per inibirne la traduzione (nel caso di complementarità incompleta) (F) (da Bhaskaran and Mohan. Vet Pathol. 2014).

2.2.2 Funzioni dei miRNA

I miRNA possono presentare diversi *modus operandi* attraverso i quali reprimono la traduzione dell'mRNA *target*. La repressione può essere il risultato di una ridotta efficienza di traduzione oppure può essere dovuta ad una effettiva riduzione dei livelli di mRNA³³. Si ritiene che il grado di complementarità tra il miRNA e le sequenze *target* nell'mRNA determini la scelta del meccanismo post-traduzionale: una volta incorporato nel complesso RISC, il miRNA degraderà in maniera specifica l'mRNA *target* se quest'ultimo presenta perfetta complementarità di sequenza, oppure ne inibirà la traduzione se l'mRNA non ha sufficiente complementarità per essere degradato ma presenta comunque un insieme adeguato di siti complementari al miRNA³³ (Figura 7A/B). Ciò nonostante, esistono delle eccezioni a questa regola generale in cui anche una complementarità quasi perfetta porta all'inibizione della traduzione e non alla degradazione³⁴. È stato, inoltre, proposto un meccanismo di azione alternativo in cui il legame dei miRNA porta ad una più rapida deadenilazione degli mRNA, diminuendone quindi la stabilità e accelerando la loro degradazione³⁵. Studi più recenti hanno inoltre dimostrato che i miRNA possono agire da "esche" ed interferire con il funzionamento di proteine regolatrici legandosi all'mRNA senza necessitare della regione *seed*. L'idea che il 3'UTR sia il solo *target* nell'interazione con i miRNA sta anch'essa cambiando dal momento che alcuni recenti studi hanno dimostrato che i siti di legame per i miRNA possono essere presenti anche nel 5' UTR e all'interno delle sequenze codificanti³⁶.

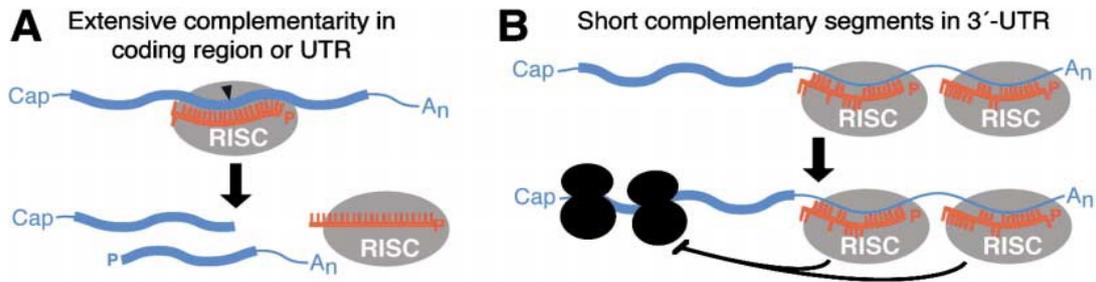


Figura 7. Meccanismi di azione dei miRNA. (A) Degradazione dell'mRNA nel caso di perfetta complementarità di sequenza tra il miRNA e il sito target al 3' UTR dell'mRNA. (B) Inibizione della traduzione nel caso di non sufficiente complementarità di sequenza tra il miRNA e l'mRNA target (da Bartel. Cell.2004).

Una importante questione, ancora oggi oggetto di dibattito è quella relativa alla specificità dell'interazione tra i miRNA e i propri *target*. Nel caso degli studi animali, metodi computazionali usano la conservazione evolutiva per distinguere i siti *target* dei miRNA da una moltitudine di segmenti 3'UTR che altrimenti avrebbero uno *score* parimenti buono per la qualità e la stabilità dell'appaiamento³⁷. La cellula, d'altro canto, non può utilizzare il filtro della conservazione evolutiva per scegliere tra le varie possibilità. Questo significherebbe che molti di questi altri mRNA potrebbero, difatti, essere *target* se espressi nelle stesse cellule dei miRNA affini. In realtà, probabilmente, l'appaiamento tra basi non è l'unico e principale determinante di specificità. Proteine o strutture dell'mRNA potrebbero limitare l'accessibilità del miRNP (*miRNA Ribonucleoprotein Complex*) alle UTR. Le suddette proteine o strutture potrebbero, anche, facilitare il riconoscimento dei veri mRNA *target* tramite elementi negli mRNA fino a quel momento sfuggiti al rilevamento³⁸. D'altra parte, deve essere considerata anche l'ipotesi alternativa che la qualità e la stabilità dell'appaiamento sia il principale determinante di specificità. Questi requisiti di complementarità

includono un *match* perfetto o quasi perfetto di 7 nt vicino all'estremità 5' del miRNA, che di per sé rappresenterebbe un grado di specificità comparabile a quello dei siti di DNA riconosciuti da molti fattori di trascrizione³⁹. Un appaiamento al di fuori del *core* dei 7 nt potrebbe fornire un ulteriore elemento di specificità. Proprio come la struttura della cromatina limita le possibilità di legame ai fattori di trascrizione, il *set* ridotto di geni trascritti in ciascuna cellula limita quali geni del genoma andranno sotto il controllo dei miRNA in quella cellula. E, come l'azione cooperativa di molteplici fattori di trascrizione aumenta la specificità del loro controllo, allo stesso modo l'azione cooperativa delle interazioni omotipiche ed eterotipiche miRNA: UTR potrebbe fornire un ulteriore meccanismo per aumentare la specificità del controllo da parte del miRNA³⁷.

2.3 I "Circular RNA" (circRNA)

I "circular RNA" (circRNA) sono una classe di RNA endogeni non-codificanti di recente scoperta. A differenza degli RNA lineari che terminano con il *cap* all'estremità 5' e la coda di Poly(A) all'estremità 3', i circRNA formano delle strutture circolari covalentemente chiuse senza polarità 5'-3' né code poliadenilate. I circRNA furono scoperti nei virus a RNA già negli anni 70. Tuttavia, si riteneva che tali trascritti circolari fossero un accidentale sottoprodotto derivante da rari errori durante lo *splicing* canonico. Con lo sviluppo delle tecnologie di sequenziamento *highthroughput* e delle analisi computazionali, sono stati scoperti migliaia di circRNA dagli Archea all'uomo. Recenti lavori hanno

dimostrato che molti di questi circRNA sono endogeni, abbondanti, conservati e stabili nelle cellule di mammifero⁴⁰.

2.3.1 Proprietà dei circRNA

I circRNA presentano diverse proprietà interessanti. Anzitutto, questi RNA circolari hanno una struttura covalentemente chiusa senza polarità 3'-5' né coda poliadenilata, caratteristiche che rendono queste molecole più stabili rispetto agli RNA lineari e non suscettibili alla degradazione da parte delle RNA esonucleasi o dell'RNAsi R⁴¹. Secondo, esiste una grande eterogeneità tra i circRNA. In alcuni casi, le molecole circolari possono essere fino a 10 volte più abbondanti dei corrispondenti mRNA lineari⁴². Terzo, i circRNA spesso esibiscono espressione specifica per tessuto e/o stadio di sviluppo. Infine, i circRNA sono evolutivamente conservati tra le diverse specie⁴⁰.

2.3.2 Biogenesi dei circRNA

Recenti studi hanno dimostrato che i circRNA vengono generati tramite un processo di *backsplicing*, che differisce dallo *splicing* canonico degli RNA lineari. Jeck *et al.* hanno avanzato due modelli per la formazione dei circRNA⁴². Il modello 1 è chiamato "*lariat-driven circularization*" o "*exon-skipping*" (Figura 8A), mentre il modello 2 è noto come "*intron-pairing-driven circularization*" o "*direct backsplicing*" (Figura 8B). Poco dopo, Zhang e collaboratori scoprirono un nuovo tipo di circRNA nelle cellule umane che deriva dagli introni e denominato "*circular intronic RNA*" (ciRNA). La biogenesi dei

ciRNA dipende da un motivo consenso contenente una sequenza di 7 nt arricchita in GU vicino al sito di *splicing* 5' e una sequenza di 11 nt arricchita in C vicino il punto di diramazione⁴³ (Figura 8C). Molto di recente, Li *et al.* hanno anche trovato esoni che vengono circolarizzati con gli introni ritenuti tra gli esoni. Questi autori hanno chiamato tali molecole "*exon-intron circRNA*" o "*EIciRNA*"⁴⁴. Infine, Yuan Gao *et al.* hanno scoperto una nuova categoria di circRNA, chiamati "*intergenic circRNA*", tramite l'utilizzo di CIRI (*circRNA Identifier*), un nuovo *tool* di identificazione di circRNA in grado di rilevare circRNA nell'intero genoma. Queste molecole contengono due frammenti di circRNA intronici fiancheggiati da segnali di *splicing* GT-AC che agiscono da siti donatore e accettore della giunzione circolare formando un circRNA integrato⁴⁵.

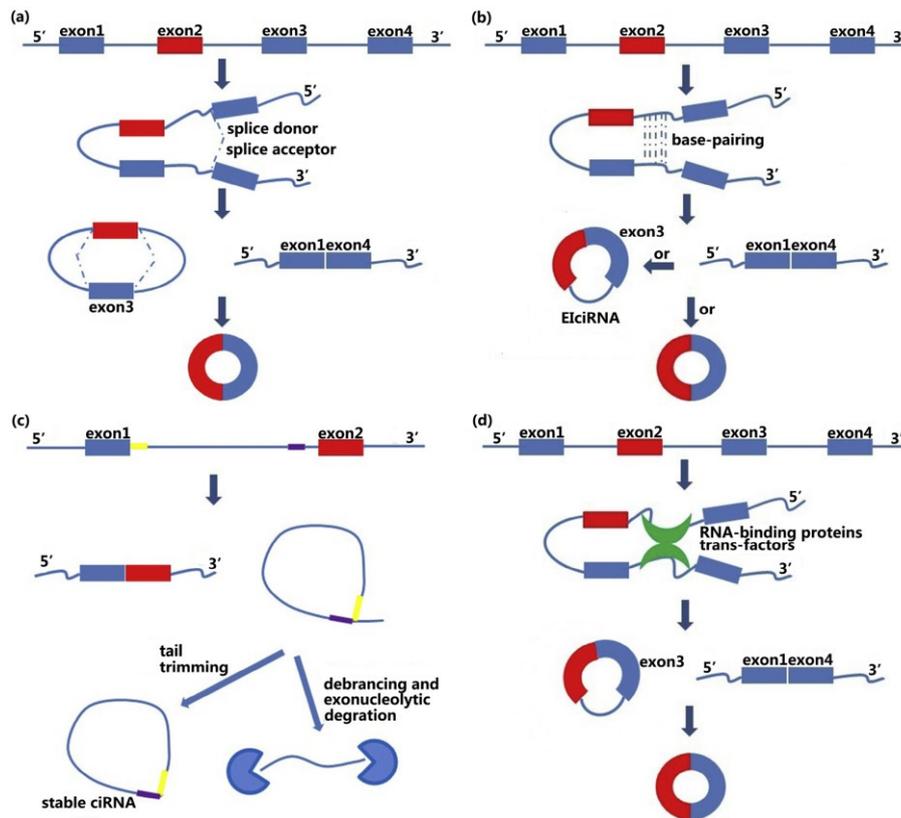


Figura 8. Modelli di biogenesi dei circRNA. (A) Lariat-driven circularization. Il sito di splice donatore all'estremità 3' dell'esone 1 covalentemente unito al sito accettore all'estremità 5' dell'esone 4 forma un lariat tramite skipping esonico. In seguito, gli introni vengono rimossi mediante lo spliceosoma. Infine, si forma il circRNA. (B) Intron-pairing-driven circularization. L'introne 1 e l'introne 3 formano una struttura circolare tramite l'appaiamento tra basi complementari. In seguito gli introni vengono rimossi o sono ritenuti per formare circRNA o EicRNA, rispettivamente. (C) Circular-intronic RNA. Il lariat intronico è generato dalla reazione di splicing. Le sequenze ricche in GU vicino al sito di splice 5' (rettangolo giallo) e le sequenze ricche in C vicino al punto di ramificazione (rettangolo viola) sono minimamente sufficienti perché l'introne sfugga alla degradazione. (D) Circularizzazione guidata da RBPs (RNA binding proteins) (da Qu et al., Cancer Lett. 2015).

2.3.3 Funzioni dei circRNA

- **I circRNA come "Competitive endogenous RNA" o "miRNA sponges"**

I ceRNA (Competitive Endogenous RNA) contengono MRE (miRNA Responsive Elements), proprio come gli mRNA, gli pseudogeni e i lncRNA, e possono competere per il legame ai miRNA. Di conseguenza, la presenza o l'assenza dei ceRNA influenza l'attività dei miRNA riguardante la regolazione dell'espressione genica. Recenti

evidenze hanno indicato che i circRNA possono agire da spugne molecolari per i miRNA o come potenti ceRNA⁴⁰ (Figura 9). Ad esempio, la prima osservazione di un circRNA che agisce da spugna è CDR1as, che possiede più di 70 siti di legame conservati per il miRNA-7⁴⁶. Ancora, circ-SRY è un RNA circolare espresso specificamente nei testicoli di topo e presenta 16 siti di legame per il miRNA-138⁴⁷. È stato dimostrato CDR1as e circ-Sry legano i miRNA senza essere degradati, il che li rende eccellenti candidati quali ceRNA⁴⁰.

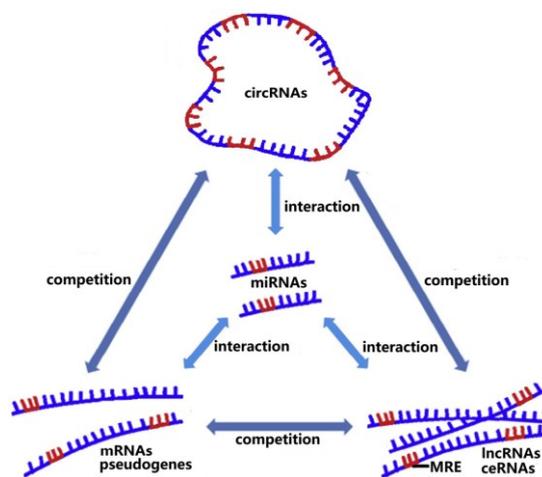


Figura 9. Network di ceRNA. CircRNA, mRNA, pseudogeni e lncRNA agiscono da ceRNA per regolare i loro rispettivi ruoli. Essi contengono MRE e competono per il legame ai miRNA (da Qu et al., Cancer Lett. 2015).

▪ Interazione con le RBP (RNA binding proteins)

Oltre alla funzione di “miRNA sponges”, i circRNA possono interagire anche con le RNA Binding Proteins, come circ-Foxo3, circ-MBL (*muscleblind*) e alcuni ciRNA. Circ-Foxo3 è in grado di legare molti tipi di proteine. Esso può reprimere il ciclo cellulare e bloccare la transizione dalla fase G1 alla fase S interagendo con CDK2 e p21. Inoltre, è stato riscontrato che i *lariat* intronici possono accumularsi nel citoplasma e legare la

proteina TDP43, sopprimendo la tossicità di TDP43 nell'ALS (*Amyotrophic Lateral Sclerosis*)⁴⁵.

- **I circRNA modulano la stabilità degli mRNA**

Alcuni circRNA possono modulare la stabilità degli mRNA. Ad esempio, l'RNA circolare antisenso CDR1as può formare un *duplex* con l'mRNA, stabilizzandolo⁴⁸.

- **I circRNA regolano la trascrizione genica**

I circRNA possono anche regolare la trascrizione genica. I meccanismi possono essere vari. Questi circRNA sono sempre presenti nel nucleo. Ad esempio, entrambi gli *ElciRNA* circ-EIF3J e circ-PAIP2, possono combinarsi con U1 snRNP per interagire ulteriormente con la RNA polimerasi II ed aumentare l'espressione dei loro geni parentali nelle cellule HeLa e nelle HEK293⁴⁴. Anche i ciRNA possono regolare la trascrizione genica. È stato dimostrato che ci-ankrd52 e ci-sirt7 possono agire da regolatori positivi della trascrizione dei loro geni parentali interagendo con la RNA Polimerasi II⁴³ (Figura 10).

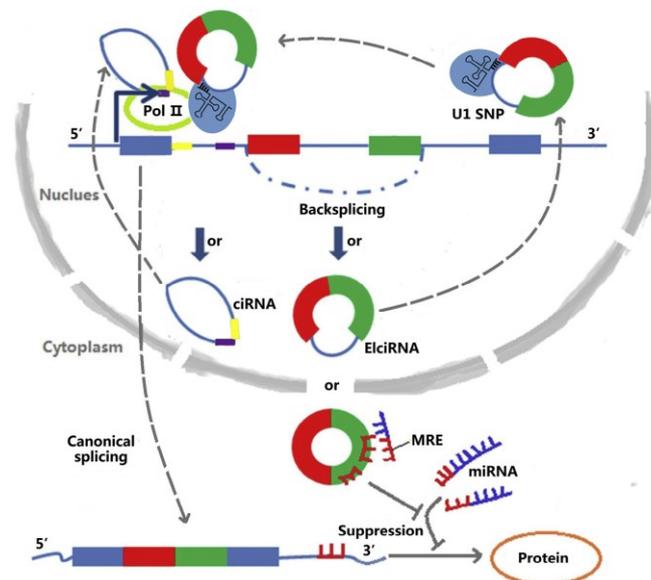


Figura 10. circRNA: regolazione dell'espressione dei geni parentali (da Qu et al., *Cancer Lett.* 2015).

3. Meccanismi di interazione tra ncRNA e loro influenza sulla biologia del cancro

L'interazione tra molecole di RNA non codificanti si verifica a causa della complementarità di sequenza tra esse: ad esempio, i ncRNA possono presentare dei "miRNA response elements" (MRE) al pari degli mRNA e quindi essere soggetti allo stesso meccanismo di *targeting*. Il legame dei miRNA ad altri ncRNA (*i.e.*, lncRNA e circRNA) potrebbe avere un duplice effetto: da una parte i miRNA potrebbero essere sequestrati, impedendo così la loro azione sugli mRNA, dall'altra il legame del miRNA ai lncRNA ed ai circRNA potrebbe promuovere la loro degradazione, similmente a quanto accade per gli mRNA⁴⁹.

3.1 I miRNA inducono la degradazione dei lncRNA

Diversi lavori hanno riportato che i miRNA possono legare i lncRNA e promuoverne la degradazione, contribuendo ai processi tumorali. I lncRNA sono strutturalmente simili agli mRNA: presentano, infatti, il 5'-*cap* e la coda 3'-Poly(A); di conseguenza, le proteine coinvolte nella regolazione del *decapping*, della deadenilazione e della degradazione degli mRNA possono controllare anche il *turnover* dei lncRNA legando specifici miRNA.

UCA1 (*Urothelial Cancer Associated 1*), un lncRNA *up*-regolato in molti tumori (*e.g.*, tumore alla cistifellea, tumore al seno, tumore ovarico)^{50, 51} presenta due siti di legame per il miR-1, un miRNA oncosoppressore ben noto. Il legame del miR-1 ad UCA1 è stato confermato

dal saggio della luciferasi nel cancro alla cistifellea e, come atteso, l'*up*-regolazione del miR-1 induceva la *down*-regolazione di UCA1 e causava una riduzione della crescita cellulare e della migrazione ed anche un aumento dell'apoptosi. Tali effetti funzionali venivano revertiti in seguito alla sovra-espressione di UCA1 e al silenziamento di Ago2, suggerendo che il miR-1 era in grado di *down*-regolare l'espressione di UCA1 tramite l'azione di Ago2.

MALAT1 (*Metastasis Associated Lung Adenocarcinoma Transcript 1*) è uno dei lncRNA più studiati ed abbondanti; la sua espressione fu inizialmente associata alle metastasi nel carcinoma ai polmoni non a piccole cellule⁵² (NSCLC), ma risulta deregolato in numerosi altri tipi di tumore. L'estremità 3' di MALAT1 è degradata dall'RNAsi P e dall'RNAsi Z, producendo un ncRNA simile ad un t-RNA, chiamato mascRNA (*MALAT1-associated small cytoplasmic RNA*), che sarà esportato nel citoplasma, mentre la maggior parte delle molecole di MALAT1 sono localizzate negli *speckles* nucleari dove regolano lo *splicing* alternativo di specifici pre-mRNA. Inoltre, MALAT1 può legare CBX4 (*chromobox 4*), un componente del *Polycomb Repressive Complex 1* (PRC1), e modulare la sua localizzazione nei granuli di intercromatina, portando all'attivazione o all'inibizione dell'espressione genica⁵³. Attraverso questi meccanismi molecolari, MALAT1 controlla l'espressione di diversi geni correlati al ciclo cellulare e al processo metastatico, influenzando quindi la proliferazione cellulare, la migrazione e l'invasione. Recenti lavori hanno riportato che MALAT1 è un *target* di numerosi miRNA oncosoppressori, che ne indurrebbero la degradazione reprimendo i suoi effetti oncogenici. MALAT1 viene regolato a livello post-trascrizionale dal miR-101 e dal miR-217 nel carcinoma a cellule squamose dell'esofago. Il

miR-101 ed il miR-217 sono coinvolti in diversi tipi di tumore come oncosoppressori e mostrano una correlazione negativa significativa con MALAT1 nei tessuti ESCC e nei tessuti adiacenti sani. La sovra-espressione del miR-101 e del miR-217 reprime significativamente l'espressione di MALAT1 portando all'inibizione della crescita cellulare, dell'invasione e delle metastasi nelle cellule ESC. Nel cancro alla vescica, MALAT1 presenta espressione inversa al miR-125b. Questo miRNA è parzialmente complementare a MALAT1 e si lega ad esso in modelli *in vitro*. Il miR-125b è *down-regolato* nel cancro alla vescica e la sua sovra-espressione riduce l'espressione di MALAT1, con conseguente inibizione della proliferazione cellulare e attivazione dell'apoptosi⁴⁹. MiR-125b è stato, inoltre, identificato quale regolatore post-trascrizionale di HOTTIP (*HOXA Distal Transcript Antisense RNA*) nel carcinoma epatocellulare⁵⁴ (HCC). HOTTIP è un lncRNA *up-regolato* in HCC, anche negli stadi più precoci dall'insorgenza del tumore, ed è posizionato in direzione antisenso all'estremità distale del *cluster* dei geni HOX. HOTTIP promuove la crescita del tumore e le metastasi *in vitro* ed *in vivo* tramite la regolazione dell'espressione dei vicini geni HOXA (*e.g.*, HOXA10, HOXA11, HOXA13). Il miR-125b è stato trovato spesso *down-regolato* nei tumori HCC e in alcuni tumori esiste una correlazione negativa tra l'espressione del miR-125b e HOTTIP. L'interazione tra miR-125b e HOTTIP è stata validata tramite il saggio della luciferasi e confermata tramite l'espressione ectopica del miR-125b che induceva la *down-regolazione* di HOTTIP⁵⁴.

HOTAIR (*HOX Antisense Intergenic RNA*) è uno dei lncRNA maggiormente studiati in quanto è stato frequentemente associato a diverse neoplasie. HOTAIR esercita le sue funzioni oncogeniche agendo da *scaffold* per l'assemblaggio del *Polycomb Repressive*

Complex 2 (PRC2) sul *cluster* dei geni HOXD e inducendo il silenziamento trascrizionale di numerosi geni soppressori di metastasi²⁷ (e.g., la famiglia delle protocaderine). HOTAIR è destabilizzato a livello post-trascrizionale da diversi miRNA oncosoppressori in vari tipi di cancro. Chiyomaru *et al.* hanno riportato l'esistenza di un legame funzionale tra il miR-34a e HOTAIR nelle linee cellulari di cancro alla prostata trattate con genisteina, un isoflavone con attività antitumorale: il miR-34a si lega direttamente a 2 MRE presenti nella molecola di RNA di HOTAIR e riduce i suoi livelli⁵⁵. In un altro lavoro Chiyomaru *et al.* hanno riportato che l'espressione di HOTAIR è correlata negativamente a quella del miR-141 nelle cellule del carcinoma al rene (RCC). Il miR-141 appartiene alla famiglia del miRNA-200, che è noto inibire la transizione epitelio-mesenchima (EMT) tramite inibizione di ZEB1 (*Zinc Finger E-box Binding Homeobox 1*) e *up*-regolazione della E-caderina. MiR-141 è in grado di legare e degradare HOTAIR in maniera Ago2-dipendente, e tale azione riduce l'espressione di ZEB2 (*Zinc Finger E-box Binding Homeobox 2*) indotta da HOTAIR⁵⁶.

L'espressione del miR-141 è stata, inoltre, trovata negativamente correlata a quella del lncRNA H19 (*H19, imprinted maternally expressed transcript*) nel cancro allo stomaco. H19, un lncRNA oncofetale, è altamente espresso durante l'embriogenesi ed è *up*-regolato in molti tumori, incluso il cancro allo stomaco. H19 agisce come precursore primario del miR-675, che a sua volta lega e reprime l'mRNA di RB1⁵⁷ (*RB transcriptional corepressor 1*). La sovra-espressione di H19 aumenta la crescita tumorale e induce la EMT; inoltre, H19 modula il processamento dei miRNA tramite l'interazione con le proteine coinvolte in

questo processo molecolare (*i.e.*, Drosha, Dicer). È stato dimostrato che il miR-141 lega H19 nel cancro allo stomaco, e reprime l'espressione di H19 e le sue funzioni tumorigeniche⁵⁸.

La prima evidenza sperimentale che i lncRNA possono essere *target* dei miRNA è relativa al trascritto antisenso di CDR1 (*Cerebellar Degeneration-Related protein*, noto anche come CiRS-7 o CDR1AS), un RNA circolare generato da un processo di *backsplicing*. MiR-671, un miRNA nucleare, induce il clivaggio di CDR1AS in maniera Ago2 dipendente. La repressione del miR-671 promuove l'*up*-regolazione sia di CDR1AS che di CDR1, suggerendo che CDR1AS sia in grado di stabilizzare il trascritto senso CDR1. Ad oggi, questo rappresenta l'unico caso riferito di un circRNA legato e degradato da un miRNA.

L'interazione tra il miR-671 e CDR1AS potrebbe influenzare l'assetto biomolecolare nel Glioblastoma Multiforme (GBM), il tumore più aggressivo e prevalente del sistema nervoso centrale⁴⁹. Barbagallo *et al.* hanno dimostrato che il miR-671-5p è significativamente *up*-regolato nel GBM; la sua sovra-espressione forzata determinava un aumento della migrazione e una riduzione del tasso di proliferazione delle linee cellulari di GBM, suggerendo un suo potenziale ruolo quale nuovo onco-miRNA nel GBM⁵⁹.

L'espressione del miR-671 era inversamente correlata a quella di CDR1AS e CDR1 nelle biopsie di GBM e l'espressione di CDR1AS e CDR1 diminuiva utilizzando il miR-671 *mimic*, suggerendo che l'interazione di queste molecole potrebbe essere funzionalmente alterata in un modello di GBM.

3.2 LncRNA come “esche” per i miRNA

Il meccanismo di interazione tra lncRNA e miRNA maggiormente esplorato si basa sull'esistenza della stessa sequenza *target* per i miRNA presente nelle molecole di lncRNA e nelle molecole di mRNA. In questo modo i lncRNA sono in grado di sequestrare i miRNA agendo da “*miRNA sponges*”. Attraverso questo meccanismo di interazione competitiva, i lncRNA riducono la quantità di miRNA disponibili e aumentano, di conseguenza, la traduzione dei loro mRNA *target*. LncRNA che agiscono da “*competitive endogenous RNA*” (ceRNA), sono stati ampiamente descritti nei circuiti molecolari coinvolti nel cancro⁴⁹.

EWSAT1 (*Ewing Sarcoma Associated Transcript 1*) è un lncRNA che agisce da oncogene nel sarcoma di Ewing e nel carcinoma naso-faringeo (NPC). EWSAT1 presenta due MRE per il *cluster* miR-326/330-5p e promuove lo sviluppo e la progressione dei tumori agendo da ceRNA per questi miRNA, inducendo l'espressione della ciclina D1, *target* dei miRNA appartenenti al *cluster* miR-326/330-5p⁶⁰.

Xia *et al.* hanno dimostrato che il lncRNA FER1L4 (*FER-1-Like Family Member 4, pseudogene*) e l'mRNA di PTEN (*Phosphatase and Tensin homolog*) presentano entrambi siti di legame per l'onco-miR miR-106a-5p e sono *down-regolati* nel cancro allo stomaco. Dal momento che FER1L4 agisce da ceRNA per il miR-106a-5p, la *down-regolazione* di FER1L4 determina il rilascio del miR-106a-5p che ha come *target* l'mRNA di PTEN, riducendo la sua espressione. La deregolazione dell'asse FER1L4 - miR-106a-5p - PTEN aumenta la proliferazione cellulare promuovendo la transizione dalla fase G0/G1 alla fase S del ciclo cellulare⁶¹.

L'elevata espressione del lncRNA H19 nelle cellule staminali di cancro al seno (BCSC) è critica per il mantenimento della staminalità. In queste cellule H19 agisce da spugna molecolare per let-7a/b, con conseguente *up*-regolazione del fattore della pluripotenza LIN28, un *target* di let-7a/b molto abbondante nelle BCSC. Curiosamente, H19 è reciprocamente represso dai suoi *target* let-7a/b, ma questo *loop* a *feedback* negativo può essere alterato da LIN28 a causa della sua capacità di inibire l'espressione di let-7a/b. L'espressione di let-7b è tamponata anche dal lncRNA HOST2 (*Human Ovarian Cancer-Specific Transcript 2*) nelle cellule di tumore ovarico. Legandosi a let-7b, HOST2 regola negativamente la sua disponibilità ed induce l'espressione dei suoi *target* oncogenici con conseguente aumento della crescita cellulare e della motilità nel cancro alle ovaie⁴⁹.

Un lavoro di Deng *et al.* riporta un altro caso in cui let-7 viene sequestrato da un lncRNA. L'*up*-regolazione del lncRNA CCAT1 (*Colon Cancer Associated Transcript 1*) nei tessuti HCC è stata associata ad un'aumentata proliferazione e migrazione cellulare⁶²; questa azione oncogenica è mediata dalla sua funzione di spugna molecolare per let-7: l'inibizione di let-7 porta alla *up*-regolazione dell'espressione dei *target* di let-7, HMGA2 (*high mobility group AT-hook 2*) e MYC (*MYC proto-oncogene, bHLH transcription factor*). È interessante notare che altri studi riportano che MYC, legandosi al promotore di CCAT1, induce la trascrizione di CCAT1 nel cancro al colon e allo stomaco, suggerendo l'esistenza di un *loop* a *feedback* positivo tra CCAT1 e MYC mediato da let-7⁴⁹.

Recenti lavori hanno riferito l'effetto inibitorio di HOTAIR sulle funzioni dei miRNA in diverse neoplasie. Su *et al.* hanno trovato che HOTAIR era altamente espresso nei tessuti HCC e promuoveva la proliferazione e la progressione dei tumori xenografi. Questi effetti

oncogenici erano in parte dovuti alla sua capacità di reprimere l'espressione del miR-1. Inoltre, anche il miR-1 era in grado di regolare negativamente l'espressione di HOTAIR, generando dunque un *loop* reciproco di repressione tra questi due ncRNA⁶³. Altre evidenze sperimentali mostrano che HOTAIR è in grado di legare e *down*-regolare il miR-152 nel cancro allo stomaco. La sovra-espressione di HOTAIR nei tessuti tumorali porta ad una diminuita espressione del miR-152 e alla *up*-regolazione del suo *target*, HLA-G (*human leukocyte antigen G*), che a sua volta facilita i meccanismi di propagazione del tumore⁶⁴.

Jin *et al.* hanno riportato un'associazione tra l'*up*-regolazione di MALAT1 e le metastasi nel cancro al seno triplo negativo (TNBC). Queste proprietà oncogeniche di MALAT1 sono mediate dalla sua capacità di sequestrare il miR-1 e, di conseguenza, aumentare l'espressione del suo *target*, SNAI2 (*Snail Family Transcriptional Repressor 2*), anche noto come Slug, un oncogene coinvolto nella regolazione dei meccanismi di invasione cellulare. Inoltre, la sovra-espressione del miR-1 è in grado di ridurre l'espressione di MALAT1, dimostrando l'esistenza di un reciproco *loop a feedback* negativo tra lncRNA e miRNA⁶⁵.

TUG1 (*Taurine UpRegulated 1*), un lncRNA ben noto e frequentemente *up*-regolato nel cancro, agisce da spugna molecolare per il miR-145. Nel cancro alla vescica, TUG1 riduce l'espressione del miR-145 e causa l'*up*-regolazione di ZEB2, *target* del miR-145, promuovendo la transizione epitelio-mesenchima e la propensione alla formazione di metastasi delle cellule tumorali⁶⁶. Il ruolo di TUG1 quale ceRNA è stato dimostrato anche in altri tumori. La sovra-espressione di TUG1 è coinvolta nell'angiogenesi nel glioblastoma tramite modulazione della proliferazione delle cellule endoteliali, della migrazione e della formazione dei vasi. Questi processi cellulari sono mediati dall'interazione di TUG1 con il

miR-299, *down*-regolato nel glioblastoma. Infatti, il *knockdown* di TUG1 induce l'*up*-regolazione del miR-299 e la concomitante riduzione del VEGFA (*Vascular Endothelial Growth Factor A*), *target* del miR-299. Questi eventi molecolari esitano in una diminuzione della densità dei microvasi in modelli xenografi di glioblastoma. Ma *et al.* hanno mostrato che l'*up*-regolazione di TUG1 nel cancro alla cistifellea (GBC) è correlata alla proliferazione cellulare e alle metastasi nel GBC e tali effetti oncogenici erano, almeno in parte, dovuti alla funzione di "*miRNA sponge*" di TUG1, che lega il miR-300 e regola negativamente la sua espressione⁶⁷. Nell'osteosarcoma, TUG1 agisce da ceRNA sequestrando il miR-9-5p, inducendo l'*up*-regolazione del fattore di trascrizione POU2F1 (*POU class 2 homeobox 1*). POU2F1 è spesso *up*-regolato nell'osteosarcoma ed è coinvolto nella proliferazione cellulare, nel differenziamento e nei processi infiammatori. Poiché POU2F1 è un *target* del miR-9-5p, il silenziamento di TUG1 inibisce la proliferazione cellulare e la formazione di colonie, mentre induce l'arresto del ciclo cellulare in fase G0/G1 e l'apoptosi. Questi processi cellulari sono mediati dall'*up*-regolazione del miR-9-5p e dalla repressione dell'espressione di POU2F1⁶⁸.

TUSC7 (*Tumor Suppressor Candidate 7*; anche chiamato LOC285194) è un lncRNA oncosoppressore trascrizionalmente indotto da TP53 (*Tumor Protein 53*); è stato inizialmente scoperto in quanto deletato nell'osteosarcoma, portando ad una incontrollata proliferazione degli osteoblasti. È, inoltre, associato alla ridotta sopravvivenza nei pazienti affetti da osteosarcoma. Diversi lavori riportano l'esistenza di un meccanismo di legame competitivo tra TUSC7 e diversi onco-miRNA associato ai processi relativi al cancro. Wang *et al.* hanno studiato il significato biopatologico della forte *down*-regolazione di

TUSC7 nei tumori HCC. Essi provarono che l'espressione ectopica di TUSC7 inibiva le metastasi, l'invasione e la EMT, agendo da spugna per il miR-10a. Questo miRNA era in grado di promuovere la EMT nelle HCC legando e reprimendo EphA4⁶⁹ (*EPH Tyrosine Kinase Receptor A4*). Inoltre, l'esone 4 di TUSC7 presenta due siti di legame per il miR-211. Nel cancro al colon, il miR-211 aumenta la proliferazione cellulare, ma questo effetto può essere revertito inducendo l'espressione forzata di TUSC7, che tampona l'attività del miR-211⁷⁰. Il ruolo di TUSC7 quale oncosoppressore è stato dimostrato anche nel cancro allo stomaco. TUSC7, *down-regolato* in questo tipo di tumore, è un *marker* prognostico indipendente di sopravvivenza libera da malattia e la sua espressione ectopica inibisce la crescita tumorale nei modelli *in vitro* e *in vivo*, in parte tramite regolazione negativa dell'espressione del miR-23⁷¹.

Senza dubbio, uno dei più iconici lncRNA con funzione di "*miRNA sponge*" è UCA1, che lega e reprime diversi miRNA in molti tumori. Il legame di UCA1 al miR-143 è stato dimostrato nel cancro al seno, dove UCA1 è in grado di modulare la crescita cellulare e l'apoptosi *down-regolando* il miR-143: questo, a sua volta, porta alla *up-regolazione* di BCL2 (regolatore dell'apoptosi) ed ERBB3⁷² (*erb-b2 receptor tyrosine kinase 3*). Il ruolo di UCA1 nel cancro alla vescica è stato associato al metabolismo dei ROS. Il silenziamento di UCA1 riduceva la produzione dei ROS, e promuoveva la glutaminolisi mitocondriale nelle cellule tumorali. In queste cellule, UCA1 agisce da ceRNA sequestrando e *down-regolando* il miR-16. Ciò induce l'*up-regolazione* di GLS2 (*Glutaminase 2*), *target* del miR-16, che aumenta l'*uptake* di glutamina e il tasso di glutaminolisi, processi che sono amplificati nelle cellule tumorali. L'espressione di GLS2 indotta da UCA1 mantiene l'equilibrio redox

e protegge le cellule tumorali riducendo l'eccessiva produzione di ROS⁷³. L'attività oncogenica di UCA1 nel CRC è il risultato della sua funzione di "miRNA sponge" per il miR-204-5p, un miRNA oncosoppressore critico. UCA1, *up*-regolato nel CRC, inibisce l'attività del miR-204-5p promuovendo, quindi, l'*up*-regolazione dei *target* di questo miRNA, CREB1, BCL2 e RAB22A (membro della famiglia dell'oncogene RAS) e regolando la proliferazione cellulare e l'apoptosi. L'*up*-regolazione di UCA1 nelle HCC è stata associata alla crescita cellulare e alle metastasi; questi processi sono indotti dal legame di UCA1 al miR-216b e portano alla *down*-regolazione del miR-216b. Ridotti livelli del miR-216b portano alla derepressione del suo *target* FGFR1 (*Fibroblast Growth Factor Receptor 1*) e all'attivazione della *pathway* di ERK⁷⁴. Un'associazione tra UCA1 e il processo metastatico è stata riportata anche per il cancro ovarico epiteliale. UCA1, infatti, promuove l'espressione di MMP14 (*Matrix Metallo Peptidase 14*), una proteina chiave coinvolta nell'invasione cellulare, agendo da "miRNA sponge" per il miR-485-5p, un miRNA che ha come *target* MMP14⁷⁵. Infine, un ruolo di ceRNA nel cancro è stato riportato anche per XIST (*X Inactivated-Specific Transcript*). XIST è stato il primo lncRNA ad essere caratterizzato funzionalmente ed è considerato il principale effettore dell'inattivazione del cromosoma X nelle femmine di mammifero durante lo sviluppo embrionale. La sua deregolazione è stata trovata in diversi tumori (*e.g.*, cancro al seno, glioblastoma, carcinoma epatocellulare), suggerendo che XIST potrebbe avere un potenziale potere diagnostico nel cancro. La *down*-regolazione o la *up*-regolazione *in vitro* di XIST è associata ad un'alterata proliferazione cellulare, alle metastasi, e all'apoptosi in diversi modelli tumorali. Song *et al.* hanno scoperto che la sovra-espressione di XIST è correlata alle

metastasi e alla prognosi negativa nei pazienti NPC⁷⁶ (*Nasopharynx Cancer*). XIST induce l'*up*-regolazione di E2F3 (*E2F Transcription Factor 3*), una proteina critica per la proliferazione delle cellule tumorali. Gli autori hanno dimostrato che l'attivazione di E2F3 promossa da XIST è causata dalla sua funzione di spugna molecolare per il miR-34a-5p (un miRNA oncosoppressore ben noto), che ha come *target* E2F3⁷⁶. Al contrario, Chang *et al.* hanno dimostrato che XIST agisce da oncosoppressore e inibisce la formazione di metastasi e la progressione delle HCC: legando il miR-181a, e riducendo la sua disponibilità, XIST induce l'*up*-regolazione di PTEN, riducendo quindi la proliferazione cellulare, l'invasione e la migrazione⁷⁷.

3.3 *CircRNA come "miRNA sponges"*

I circRNA sono considerati nuovi potenziali ceRNA: possono contenere MRE e competere con gli mRNA per il legame ai miRNA. Di conseguenza, i circRNA reprimono competitivamente l'attività dei miRNA legandoli e sequestrandoli. Poiché i miRNA sono fortemente coinvolti nella quasi totalità degli aspetti della fisiologia cellulare e giocano un ruolo chiave nell'insorgenza e nella progressione del cancro, i circRNA possono essere ragionevolmente considerati come una nuova classe di molecole di RNA strettamente associate alla regolazione della proliferazione, del differenziamento e dei processi metastatici⁷⁸.

Zheng *et al.* hanno riportato che il circ-TTBK2 (*Tau Tubulin Kinase 2*) è significativamente *up*-regolato nei tessuti e nelle cellule di glioma, a differenza delle controparti lineari. La sovra-espressione del circ-TTBK2 è associata ad un aumentato tasso della proliferazione

cellulare e dell'invasione e ad una riduzione dell'apoptosi. Circ-TTBK2 contiene MRE per il miR-217, che svolge un ruolo di oncosoppressore nelle cellule di glioma. Infatti, il circ-TTBK2 e il miR-217 interagiscono l'uno con l'altro in maniera AGO2-dipendente e l'*up*-regolazione di circ-TTBK2 promuove la trasformazione maligna delle cellule di glioma tramite la *down*-regolazione del miR-217. Di conseguenza, HNF1 β , un diretto *target* del miR-217, viene derepresso e lega il promotore di Derlin-1 aumentando la sua espressione. Infine, Derlin-1 è in grado di promuovere la proliferazione cellulare, la migrazione e l'invasione, e inibire l'apoptosi delle cellule di glioma attivando le *pathway* PI3K/AKT e ERK. Inoltre, ripristinando l'espressione del miR-217 si ottiene la reversione della progressione tumorale indotta da circ-TTBK2, suggerendo l'esistenza di un reciproco *loop* a *feedback* negativo tra circ-TTBK2 e miR-217⁷⁹.

MiR-145 è un miRNA oncosoppressore ben noto nel CRC, ed esplica la sua funzione legando e degradando gli mRNA degli oncogeni ERK5 ed IRS1; è, inoltre, ben noto il suo ruolo di *marker* predittivo di sopravvivenza nei pazienti CRC. In uno studio di Xie *et al.* è stato dimostrato che la *down*-regolazione del miR-145 nel CRC è dovuta al ruolo del circ_001569, che agisce da "*miRNA sponge*" inibendo direttamente l'azione del miR-145. Circ_001569 è stato trovato *up*-regolato nei tessuti CRC e correlato con la progressione e l'aggressività della malattia.

Nello specifico, circ_001569 non influenza direttamente l'espressione del miR-145, ma tramite un meccanismo di "*sponging*" inibisce la sua attività post-trascrizionale: di conseguenza, questo determina l'*up*-regolazione dei *target* del miR-145, E2F5 (E2F

Transcription Factor 5), BAG4 (*BCL2-Associated Athanogene 4*) e FMNL2 (*Formin Like 2*), che sono responsabili della proliferazione cellulare e dell'invasione indotta dal circ_001569⁸⁰.

Ulteriori lavori sul CRC, studiando il ruolo del cir-ITCH nella biopatologia di questo tumore, hanno trovato una potenziale interazione tra il cir-ITCH ed i miRNA miR-7 e miR-20a. Cir-ITCH è *down*-regolato nei tessuti CRC e la sua espressione ectopica determina una riduzione della proliferazione cellulare. Questo effetto è dovuto all'attività di spugna molecolare di cir-ITCH per il miR-7 e il miR-20a: entrambi possono legare il 3'-UTR di ITCH (*Itchy E3 Ubiquitin Protein Ligase*), l'isoforma lineare di cir-ITCH. L'*up*-regolazione di ITCH indotta da cir-ITCH promuove l'ubiquitinazione e la degradazione della forma fosforilata di DVL2 (*Dishevelled Segment Polarity Protein 2*) e, di conseguenza, inibisce la *pathway Wnt/β-catenin*, reprimendo l'espressione di MYC e di CCND1⁸¹ (*cyclin D1*). È interessante notare che altri autori hanno trovato risultati molto simili nelle ESCC: cir-ITCH agisce da spugna per i miRNA miR-7, miR-17, e miR-214, aumenta l'espressione di ITCH, e promuove la degradazione ubiquitina-dipendente di DVL2, inibendo quindi la via canonica di Wnt⁸².

Oltre al meccanismo di "*sponging*" del miR-7 da parte di cir-ITCH appena descritto, il sequestro del miR-7 da parte di CDR1AS è stato uno dei primi meccanismi di ceRNA studiati nella biologia dei ncRNA, anche in relazione al cancro. L'espressione di CDR1AS è stata trovata aumentata nei tessuti HCC ed inversamente correlata all'espressione del miR-7, che era poco espresso nei medesimi campioni. Nonostante il ruolo oncogenico del miR-7 (precedentemente riferito per il CRC e le ESCC), questo miRNA esibiva proprietà di oncosoppressore nelle HCC. CDR1AS presenta 63 MRE per il

miR-7 e reprime fortemente la sua attività. Il *knockdown* di CDR1AS promuove l'espressione del miR-7 e reprime i suoi *target*, CCNE1 (*cyclin E1*) e PIK3CD (*phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit delta*): questa cascata molecolare risulta in una riduzione della proliferazione cellulare e dell'invasione nelle HCC⁸³.

Tramite un profilo di espressione nelle OSCC, Chen *et al.* hanno identificato l'*up*-regolazione di un circRNA chiamato circRNA_100290, che era funzionalmente correlato ad un alterato controllo del ciclo cellulare e alla proliferazione cellulare nelle cellule OSCC. CircRNA_100290 funge da "*miRNA sponge*" per diversi membri della famiglia del miR-29, riducendo la quantità disponibile di miR-29 e, di conseguenza, promuovendo la traduzione di uno dei suoi *target*, CDK6 (*Cyclin Dependent Kinase 6*), che a sua volta induce la transizione dalla fase G1 alla fase S nel tumore⁸⁴. Il primo trascritto circolare ad essere identificato fu il circRNA Sry: esso codifica un gene che mappa sulla regione "*sex-determining*" del cromosoma Y umano ed è altamente espresso nei testicoli dei topi adulti. Inizialmente, il circRNA Sry fu considerato un artefatto derivato da un aberrante meccanismo di *splicing* e non vi fu attribuita alcuna funzione. Si è iniziato ad indagare il ruolo del circRNA Sry solo di recente. Il circRNA Sry contiene 16 putativi siti *target* per il miR-138 e la sua funzione di "*miRNA sponge*" per il miR-138 è stata dimostrata da Hansen *et al*⁸⁵. Ad oggi, non è riportata alcuna evidenza sperimentale della deregolazione dell'asse Sry circRNA-miR-138 nel cancro; tuttavia, come descritto da Zhao e Shen, il miR-138 potrebbe legare diversi trascritti correlati al cancro⁸⁶. Ad esempio, la *down*-regolazione del miR-138 promuove la progressione maligna nel colangiocarcinoma tramite l'azione del

suo *target* RhoC⁸⁷ (*Ras homolog gene family, member C*). Queste osservazioni potrebbero suggerire che il ruolo del legame endogeno competitivo tra il circRNA Sry e il miR-138 meriterebbe un'analisi approfondita nell'ambito dei fenotipi del cancro.

4. Il Carcinoma del Colon-Retto (CRC)

Il carcinoma del colon-retto (CRC) è una delle patologie più comuni nelle società industrializzate. Nel 2017, sono state stimate 135.430 diagnosi di CRC, e 50.260 decessi causati da questa malattia (<http://seer.cancer.gov/statfacts/html/colorect.html>) (<https://www.cancer.org/research/cancer-facts-statistics/colorectal-cancer-facts-figures.html>). Nelle società occidentali il CRC rappresenta la seconda causa di morte per cancro ed è il terzo tipo di tumore più frequente⁸⁸.

Il carcinoma del colon-retto è in assoluto il tumore a maggiore tasso di insorgenza nella popolazione italiana, con quasi 52.000 diagnosi stimate per il 2014. Tra gli uomini si trova al terzo posto, preceduto da prostata e polmone, costituendo il 14% di tutti i nuovi casi di tumore; tra le donne è al secondo posto, preceduto dalla mammella, con il 13%. Nella classifica dei tumori più frequenti per gruppi di età il carcinoma del colon-retto occupa sempre posizioni elevate, variando nelle diverse età tra l'8% e il 14% negli uomini e tra il 5% e il 17% nelle donne. La diffusione dei fattori di rischio, l'anticipazione diagnostica e l'aumento dell'età media della popolazione sono alla base della progressiva crescita dell'incidenza di questo tumore negli ultimi decenni. Il *trend* temporale per gli uomini è passato da un andamento in lieve crescita (+1%/anno) tra la metà degli anni '90 e l'inizio degli anni Duemila, ad un marcato incremento tra il 2004 e il 2007, con successiva riduzione, effetti potenzialmente associati all'attivazione dei programmi di *screening* organizzati. Tra le donne l'aumento (+0,6%/anno) appare omogeneo per tutto il periodo esaminato. Nel 2011 sono stati osservati 19.077 decessi per carcinoma del colon-retto

(ISTAT) (di cui il 54% negli uomini), neoplasia al secondo posto nella mortalità per tumore (10% nei maschi, 12% nelle femmine), e tra il secondo e terzo posto nelle varie fasce di età. Il carcinoma del colon-retto presenta una prognosi sostanzialmente favorevole e progressivamente in miglioramento: dal 50% a 5 anni dei primi anni '90 al 64% del 2005-2007 nei maschi e dal 51% al 63% nelle femmine (<http://www.registri-tumori.it>). Nel corso della vita, il rischio di sviluppare il CRC è circa il 6%. L'età media di insorgenza è 62 anni. Il rischio aumenta dopo i 40 anni sia per gli uomini sia per le donne⁸⁸.

La malattia si presenta all'inizio come un polipo adenomatoso benigno, che evolve poi in adenoma avanzato con alto grado di displasia e infine progredisce in cancro invasivo. I tumori invasivi confinati nella parete del colon (stadio I e II del sistema TNM) sono curabili, ma se non trattati diffondono ai linfonodi regionali (stadio III) e in seguito determinano la formazione di metastasi in altri siti distali (stadio IV). I tumori di stadio I e II sono curabili tramite escissione chirurgica, e fino al 73% dei tumori di stadio III sono curabili tramite la chirurgia coadiuvata dalla chemioterapia. I recenti progressi nell'ambito della chemioterapia hanno incrementato la sopravvivenza, ma solitamente lo stadio IV della malattia è incurabile⁸⁹.

La prognosi del CRC si basa, in gran parte, sullo stadio patologico al momento della diagnosi ed è influenzata da numerosi fattori. Il più importante fattore prognostico è lo stadio del tumore. Il grado d'invasione della parete intestinale, le metastasi linfonodali e le metastasi a distanza influiscono negativamente sulla prognosi.

Tre differenti sistemi sono utilizzati per la stadiazione del CRC:

1. Sistema di Dukes

2. Sistema di Astler-Coller

3. Sistema TNM

La classificazione di Dukes, proposta dal dottor Cuthbert E. Dukes nel 1932, pone l'accento sulla infiltrazione tissutale, sul coinvolgimento linfonodale e sulla presenza di metastasi a distanza. Il sistema di Astler-Coller è una classificazione anatomo-chirurgica dei tumori dell'intestino lievemente modificata rispetto a quella proposta da Dukes.

La "American Joint Committee on Cancer" (AJCC) e la "International Association of Cancer" (UICC) hanno formulato il sistema di classificazione TNM tramite il raggruppamento di tre componenti: tumore, linfonodi e metastasi⁸⁸ (Tabella 1).

La classificazione TNM, più attuale, pone maggiore attenzione al grado di infiltrazione tissutale, discriminando tra l'invasione della sola mucosa, dello strato muscolare o della sierosa. Benché i reperti strumentali siano molto importanti per diagnosticare le lesioni neoplastiche a distanza ed il coinvolgimento linfonodale, per una corretta valutazione della stadiazione, e quindi della prognosi, si rende necessaria la valutazione istologica post-operatoria.

Attualmente, le scelte terapeutiche si basano su questa classificazione.

Dukes	Astler-Coller	Classificazione TNM	Classificazione Numerica	Descrizione anatomopatologica	Sopravvivenza a 5 anni
A	A	T1 N0 M0	I	Interessa la mucosa e la sottomucosa	>90%
A	B1	T2 N0 M0	I	Si estende allo strato muscolare	>90%
B	B2	T3 N0 M0	IIA	Si estende alla tonaca muscolare	70-85%
B	B3	T4 N0 M0	IIB	Si estende alla sierosa	55-65%
C	C1	T1-2 N1 M0	M0 IIIA	Metastasi a 1-3 linfonodi regionali	45-55%
C	C2	T3-4 N1 M0	M0 IIIB		45-55%
C	C3	Tx N2 M0	M0 IIIc	Metastasi a 4 o più linfonodi regionali	20-30%
D	D	Tx Nx M1	IV	Metastasi a distanza	5%

Tabella 1. Classificazione TNM.

4.1 Patogenesi del CRC

Tradizionalmente, si distinguono due tipologie di carcinoma del colon-retto: forme sporadiche e forme familiari (ereditarie). Il 75-80% dei tumori del colon-retto ha origine sporadica. Di tutti i pazienti, una percentuale elevata ha parenti da primo fino al terzo grado affetti da carcinoma del colon-retto. È abbastanza chiaro, quindi, che anche nelle forme sporadiche i familiari presentano un rischio maggiore di sviluppare il CRC.

I tumori sporadici possono essere classificati in due sottotipi chiaramente distinti, sulla base degli eventi molecolari che danno origine al tumore. Si ritiene esistano due principali *pathway* implicate nella carcinogenesi del CRC. La prima di queste, definita canonica o "*suppressor*" *pathway*, coinvolge l'instabilità cromosomica (CIN). Essa è caratterizzata da perdite alleliche ai cromosomi 5q (APC), 17p (p53), e 18q (DCC/SMAD4). La seconda *pathway* coinvolge l'instabilità dei microsatelliti (MSI), ed è chiamata "*mutator*" *pathway*. La *pathway* MSI è presente in circa il 15-20% dei tumori del colon-retto sporadici.

A parte le differenze a livello molecolare, queste due *pathway* si manifestano con distinte caratteristiche cliniche e istopatologiche⁸⁸.

▪ *Pathway* canonica

La *pathway* canonica è presente nell'80-85% dei carcinomi del colon-retto e si presume evolva secondo la sequenza di eventi molecolari descritti nel modello di Fearon e Vogelstein. È ormai accertato che nella maggior parte dei casi, i carcinomi insorgono da adenomi preesistenti. Fearon e Vogelstein proposero un modello per la carcinogenesi del CRC che correlava specifici eventi genetici con la morfologia del

tessuto. Ogni passaggio, dalla mucosa normale verso il carcinoma, coinvolge specifiche e definite alterazioni genetiche (Figura 11). Questo modello lineare si è evoluto in un approccio più complesso ed esaustivo. Mutazioni in geni oncosoppressori, quali APC, p53, e DCC, e oncogeni, come K-ras, sono caratteristiche di questo modello e della *pathway* canonica. I tumori con CIN (instabilità cromosomica) sono anche caratterizzati da squilibrio allelico (che coinvolge comunemente i cromosomi 5q, 8p, 17p e 18q), duplicazioni cromosomiche e traslocazioni. Secondo questa *pathway*, le mutazioni del gene oncosoppressore APC si verificano nella prima fase di adenoma, seguite poi da mutazioni di K-ras e p53 e infine da delezioni del cromosoma 18q⁹⁰.

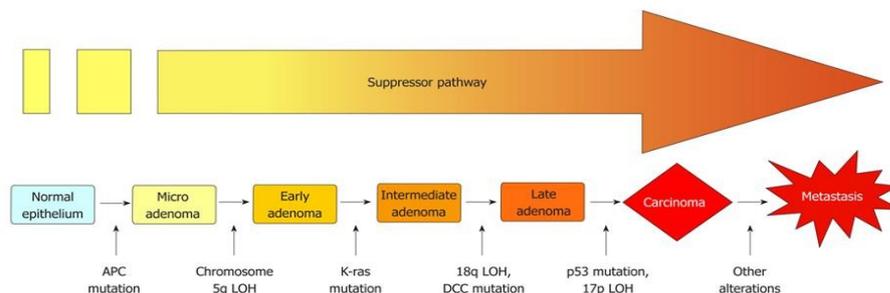


Figura 11. Alterazioni molecolari nella "suppressor" pathway (da Moràn et al., World J Gastrointest Oncol, 2010).

▪ MSI pathway

La *pathway* MSI o "mutator" pathway, è coinvolta in circa il 15-20% delle forme sporadiche di CRC. I tumori che presentano MSI sono caratterizzati da accumulo di mutazioni nelle sequenze microsatellite (*High Microsatellite Instability*, MSI-H). I microsatelliti sono corte sequenze di DNA non codificante ripetute *in tandem* lungo il

genoma. Questo accumulo di mutazioni *frameshift* è causato da alterazioni nei geni codificanti il sistema di riparazione dei *mismatch* (MMR). I geni coinvolti nel sistema MMR sono almeno sette: hMLH1, hMLH3, hMSH2, hMSH3, hMSH6, hPMS1 e hPMS2. Quando le proteine del complesso MMR sono funzionali, gli errori introdotti dalla DNA polimerasi nelle sequenze microsatellite durante la replicazione sono riparati. L'acquisizione di migliaia di mutazioni, caratteristica del fenotipo MSI-H, richiede l'inattivazione dei geni MMR. Mutazioni germinali o cambiamenti epigenetici, in hMLH1 (principalmente silenziato tramite metilazione) e in hMSH2 rappresentano la causa più comune di MSI-H nei tumori sporadici (e in HNPCC, *Hereditary Non Polyposis Colorectal Cancer*). Mutazioni in hMSH6 sono meno frequenti e le alterazioni degli altri geni MMR sono molto rare. Questi dati rafforzano l'idea che la perdita di hMLH1 e hMLH2 sia associata con la completa inattivazione del sistema MMR, mentre l'alterazione delle altre proteine causa solamente una parziale inattivazione del complesso. I tumori sporadici che presentano MSI-H generalmente non mostrano grandi anomalie citogenetiche né aneuploidia. Tali tumori sono caratterizzati da bassa frequenza, o assenza, di mutazioni o perdite alleliche nei geni APC, K-ras e p53, solitamente alterati nella "*suppressor*" pathway, e perdita di eterozigosi ai cromosomi 5q, 17p, e 18q. Queste mutazioni sono invece descritte nelle sequenze microsatellite presenti nei geni coinvolti nella carcinogenesi, quali TGF β RII, IGF2R, BAX, MSH3, MSH6, caspasi 5, APC, β -catenina, Tcf4, axina, MMP-3, E2F-4, BCL-10, cdx-2 e hRAD50. Inoltre, un certo numero di geni normalmente funzionali è silenziato tramite

metilazione. La maggior parte dei tumori sporadici con MSI-H mostra un *CpG island methylator phenotype* (CIMP), caratterizzato da metilazione aberrante del DNA⁹⁰.

Il carcinoma del colon-retto è il risultato di numerose alterazioni genetiche, tuttavia alcune vie di segnalazione cellulare sono state identificate quali fattori chiave nella patogenesi del CRC (Figura 12).

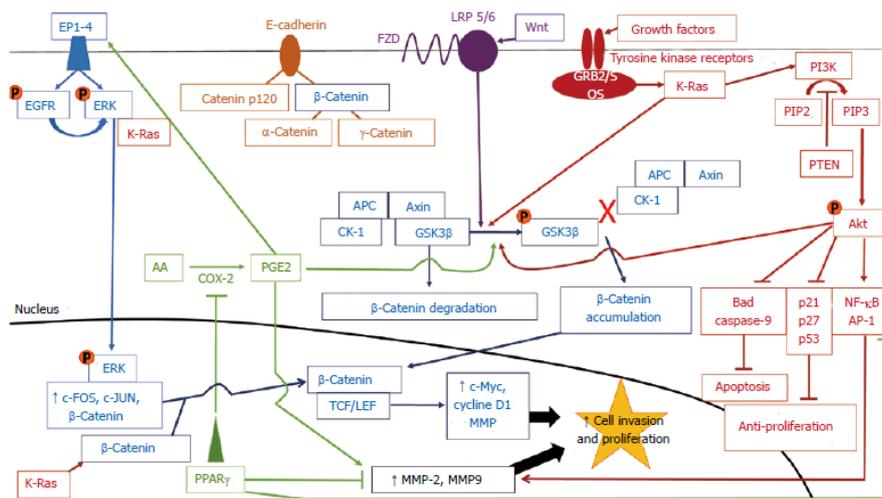


Figura 12. Crosstalk tra le pathway coinvolte nella patogenesi del CRC (da Applegate and Lane. World J Gastrointest Oncol 2015).

- APC e la via di segnalazione WNT/ β -catenina

L'evento iniziale nella patogenesi del CRC è rappresentato dall'attivazione della via di segnalazione Wnt. Tale via regola la proliferazione, il differenziamento e la motilità delle cellule dell'epitelio intestinale. In condizioni fisiologiche, la proteina APC (*Adenomatous Polyposis Coli*) forma un complesso citoplasmatico con l'Axina presente nel citosol. Entrambe le proteine possiedono siti di legame per i membri del loro complesso funzionale. Il complesso APC-Axina recluta altri elementi, la

serina/treonina chinasi GSK3 β (*Glicogen synthase kinase 3 β*) e la caseina chinasi 1 (CK-1). Insieme, queste proteine costituiscono il complesso noto come *β -catenina destruction complex*. La β -catenina nel citosol è legata e fosforilata da queste chinasi e quindi marcata per la degradazione attraverso un meccanismo ubiquitina-proteasoma dipendente. La β -catenina svolge una duplice funzione nelle cellule: essa, infatti, agisce da fattore di trascrizione nel nucleo e stabilizza l'adesione cellulare. Quando presente nel citosol la β -catenina si lega alla E-caderina, proteina transmembrana responsabile della formazione e del mantenimento delle giunzioni aderenti tra le cellule epiteliali. Si ritiene che la perdita di funzione della E-caderina si verifichi nella fase finale della carcinogenesi e conduca alla distruzione del complesso ECCU (*E-Caderin-Catenin Unit*), evento che causa la perdita delle giunzioni aderenti con conseguente incremento della motilità e della migrazione cellulare. E-caderina e APC competono per il legame alla β -catenina e cooperano al fine di mantenere l'equilibrio della concentrazione della β -catenina nel citosol. La perdita di funzione di APC risulta nella saturazione della E-caderina e nel conseguente accumulo della β -catenina nel citosol; da qui la β -catenina viene traslocata nel nucleo dove agisce da co-fattore trascrizionale aumentando la trascrizione di geni che regolano proliferazione, differenziamento e migrazione cellulare, quali c-Myc, la metalloproteinasi-7 (MMP-7) e la ciclina D1. Tramite fosforilazione della β -catenina e sua successiva degradazione via ubiquitina-proteasoma, APC e le altre proteine del complesso, agiscono da regolatori negativi della via di segnalazione Wnt/ β -catenina e mantengono l'omeostasi delle cellule della cripta e delle cellule staminali. Si ritiene che, a causa del

ruolo chiave nella regolazione della *pathway* Wnt, le mutazioni con perdita di funzione di APC costituiscono l'evento iniziale nella patogenesi del CRC. Difatti, APC è mutato in circa l'80% dei carcinomi del colon-retto, mentre mutazioni della β -catenina sono state trovate in circa il 5% dei CRC. Le mutazioni di APC possono essere di origine ereditaria, come nel caso di FAP (*Familial Adenomatous Polyposis*), o dovute a ipermetilazione o deregolazione del gene APC. Nelle mutazioni con perdita di funzione di APC è persa la capacità di degradare la β -catenina, permettendo l'attivazione costitutiva della *pathway* Wnt/ β -catenina e la trascrizione di oncogeni importanti per la proliferazione cellulare e la formazione di metastasi⁹¹.

- K-ras, BRAF e la via di segnalazione ERK/MAPK

La via di segnalazione ERK/MAP chinasi gioca un ruolo critico nella patogenesi del CRC. Infatti, il 30-35% dei pazienti con tumori sporadici del colon-retto presenta mutazioni di K-ras, tipicamente associate all'insorgenza del tumore tramite la *pathway* CIN, seguendo la classica sequenza adenoma-carcinoma proposta da Vogelstein. D'altra parte, mutazioni di BRAF, una molecola a valle di K-ras, si osservano solo nel 9-13% dei pazienti affetti dalla forma sporadica della malattia, ma sono frequentemente osservate nei tumori con instabilità dei microsatelliti (MSI) e prendono origine dalla "*mutator*" *pathway*⁹².

K-ras è responsabile della trasduzione di segnali mitogeni dai recettori dei fattori di crescita sulla superficie cellulare al nucleo. K-ras agisce da interruttore molecolare, regolando la chinasi ERK (*Extracellular signal-Regulated Kinase*) e la via di segnalazione

PI3K/Akt. Durante l'attivazione di K-ras, il legame dei fattori di crescita ai recettori tirosina chinasi determina il reclutamento del complesso GRB2/SOS (*Growth factor Receptor/Bound protein2/Son Of Sevenless*) sul lato interno della membrana citoplasmatica. Questo complesso attiva la proteina G Ras (*Rat sarcoma*). Quando K-ras lega il GTP diviene attiva grazie ai cambiamenti conformazionali di domini che controllano la specificità di legame per molecole effettrici, quali le chinasi Raf e PI3K, avviando una cascata di eventi di fosforilazione che si conclude con la fosforilazione e successiva traslocazione nucleare di ERK (Figura 13). Nel nucleo, ERK fosforila e attiva specifici fattori di trascrizione, aumentando l'espressione di geni *target*, quali c-FOS e c-JUN, che regolano proliferazione, differenziamento e apoptosi. Inoltre, l'attivazione di K-ras risulta in una incrementata espressione della β -catenina, che si accumula nel citosol. L'attività di K-ras termina con l'idrolisi del GTP. Mutazioni di K-ras distruggono la sua attività GTPasica e mantengono K-ras in uno stato perennemente attivo, aumentando la segnalazione di ERK. Mutazioni di K-ras interagiscono con la via di segnalazione Wnt/ β -catenina determinando la fosforilazione di GSK3 β attraverso l'attivazione di PI3K. Come già discusso, l'inattivazione di GSK3 β porterà alla traslocazione nucleare della β -catenina. Pertanto, in condizioni fisiologiche, GSK3 β regola negativamente entrambe le vie di segnalazione, K-ras e Wnt/ β -catenina, tramite fosforilazione di K-ras, contribuendo così alla sua degradazione. GSK3 β gioca un ruolo fondamentale nella regolazione delle due *pathway* inducendo la degradazione di intermedi chiave di ciascuna via e prevenendo la trascrizione di geni importanti nella patogenesi del CRC. Nelle cellule tumorali e in modelli murini di CRC,

mutazioni con guadagno di funzione di K-ras accoppiate a mutazioni con perdita di funzione di APC sono state associate all'aumento dei livelli nucleari di β -catenina e all'aumento della dimensione, del numero e dell'incidenza dei tumori rispetto alle cellule o ai topi con mutazioni di K-ras o APC solamente. I tumori risultanti mostravano un tasso di migrazione più alto ed una capacità invasiva maggiore dovuta all'incrementata attività della ciclina D1, che promuove la progressione del ciclo cellulare⁹¹.

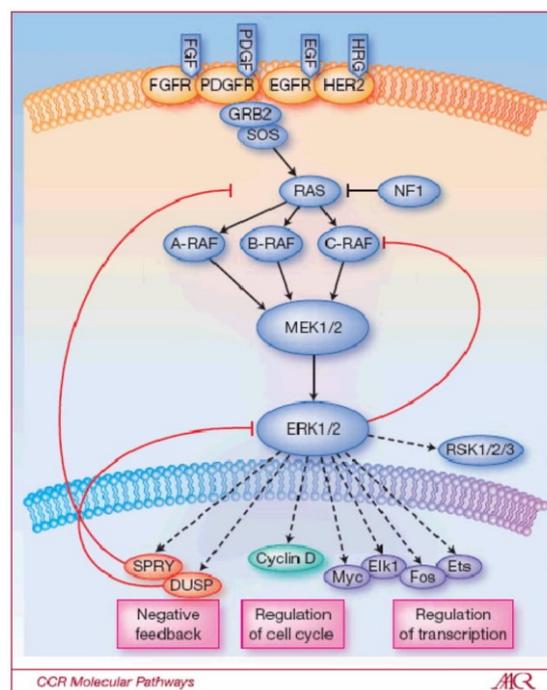


Figura 13. Rappresentazione schematica della via di segnalazione Ras/ERK/MAP CHINASI (da Pratilas and Solit, Clin Cancer Res. 2010).

- La via di segnalazione PI3K/AKT

La via di segnalazione PI3K/Akt costituisce un'altra importante *pathway*, la cui attivazione induce trasformazione cellulare, proliferazione, migrazione, e sopravvivenza, processi che lavorano insieme per promuovere la progressione

tumorale. Nel 30-35% di tutti i carcinomi del colon-retto sono coinvolte mutazioni che determinano un'attivazione aberrante di questa *pathway*⁹¹. PI3K (Fosfatidilinositolo 3-chinasi) appartiene ad una famiglia di chinasi dei lipidi, e riveste un ruolo chiave nella fisiologia cellulare e nello sviluppo tumorale. Sono state identificate tre classi di PI3K, ciascuna delle quali caratterizzata da distinte proprietà strutturali e funzionali. La classe principalmente coinvolta nella patogenesi tumorale è la Classe I, costituita da una subunità regolatoria (p85) e una subunità catalitica (p110). Tre geni, denominati PIK3R1, PIK3R2 e PIK3R3, codificano tre diverse isoforme della subunità regolatoria p85; le subunità catalitiche p110 α , p110 β e p110 γ sono codificate, rispettivamente, da PIK3CA, PIK3CB e PIK3CD. Le mutazioni somatiche note per la capacità di indurre l'attivazione di PI3K nelle cellule tumorali si trovano comunemente nei geni PIK3CA e PIK3CB. L'attivazione di PI3K può realizzarsi anche in risposta a stimoli esterni, tramite l'azione dei recettori tirosina chinasi (RTK) o l'attivazione di K-ras. A seguito del legame della subunità p85 ai residui di fosfotirosina localizzati nella porzione intracellulare di RTK, viene rilasciata la subunità catalitica p110 e PI3K diviene attiva. PI3K fosforila il fosfatidilinositolo 4,5 bifosfato (PIP2) a fosfatidilinositolo 3,4,5 trifosfato (PIP3) e, successivamente, l'accumulo di PIP3 sulla membrana cellulare porta al reclutamento di AKT e alla sua fosforilazione tramite PDK1 e mTORC2 (Figura 14). Alcune delle vie attraverso cui Akt attivata promuove la crescita e la sopravvivenza cellulare includono: 1) inibizione di proteine proapoptotiche della famiglia Bcl-2; 2) aumentata degradazione di p53, tramite una maggiore disponibilità citoplasmatica di MDM2; 3) induzione della trascrizione di geni antiapoptotici tramite

il fattore di trascrizione NF- κ B. Un altro importante effetto di Akt è l'attivazione del gruppo di proteine mTOR (*Mammalian Target Of Rapamycin*), che a sua volta aumenta la sintesi proteica tramite mTORC1 e mTORC2. La via di segnalazione PI3K è regolata negativamente dalla proteina fosfatasi PTEN (*Phosphatase and Tensin homologue protein*), una molecola oncosoppressore che defosforila il PIP3 in PIP2, inibendo l'attivazione a valle di Akt. L'inattivazione di PTEN è generalmente associata alla presenza di una forma tronca, non funzionale della proteina - risultato di mutazioni somatiche - o al silenziamento epigenetico, spesso tramite ipermetilazione del promotore⁹³. A causa del suo ruolo di regolatore negativo della via di segnalazione PI3K/Akt, PTEN è associato a inibizione della progressione del ciclo cellulare, induzione dei processi di morte cellulare, e stimolazione dell'angiogenesi. In molti carcinomi del colon-retto sono state riscontrate mutazioni e perdita di espressione di PTEN, correlata con l'aggressività e l'invasività del tumore.

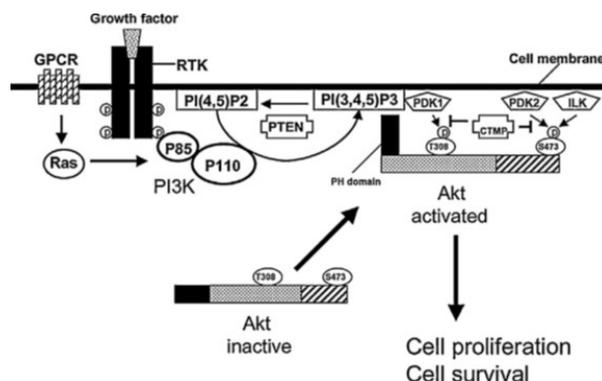


Figura 14. Rappresentazione schematica della via di segnalazione PI3K/Akt (da Dionysios Papadatos-Pastosa et al. *Critical Reviews in Oncology/Hematology* 2015).

- La via di segnalazione di p53

Le mutazioni nel gene oncosoppressore p53 sono le più comuni riscontrate nei tumori umani. La mutazione o l'assenza di p53 è osservata nel 50% dei carcinomi del colon-retto. Quale gene oncosoppressore, p53 è attivato da stimoli genotossici, ai quali risponde arrestando la progressione del ciclo cellulare e inducendo l'apoptosi. Nelle cellule sane, la soppressione di p53 è necessaria per consentire la normale crescita cellulare, pertanto la sua concentrazione è mantenuta bassa e la sua espressione è regolata tramite un meccanismo di degradazione ubiquitina-proteasoma dipendente, mediato dalla ubiquitina-ligasi MDM2. MDM2 è essa stessa un *target* di p53, creando così un *loop a feedback* negativo che controlla l'espressione e l'attività di p53. In risposta a un danno citotossico, p53 viene attivata da proteine chinasi, che la fosforilano a livello del suo dominio di legame per MDM2, stabilizzandola e permettendone il legame al DNA per indurre la trascrizione di geni quali l'inibitore delle chinasi ciclina-dipendenti p21 e il fattore proapoptotico BAX⁹¹ (*Bcl-2 Associated X protein*). P53 induce l'arresto del ciclo cellulare in fase G1 per consentire la riparazione del DNA durante la replicazione delle cellule esposte a stress genotossico. Se il danno al DNA è troppo esteso perché possa essere efficacemente riparato, p53 può indurre apoptosi, e questa è considerata la principale *pathway* attraverso cui p53 esplica la sua funzione di oncosoppressore⁹⁰. Le mutazioni missenso ricorrono nell'80% di tutte le mutazioni di p53, risultando in una proteina stabile che si accumula nel nucleo ma, tuttavia, manca di attività trascrizionale. L'accumulo di p53 mutata nei soggetti affetti da CRC è fortemente correlato con l'aumento delle metastasi e la prognosi sfavorevole. La

maggior parte delle mutazioni di p53 si verifica negli ultimi stadi dell'evoluzione da adenoma a carcinoma, e la frequenza delle mutazioni aumenta al progredire della lesione⁹¹.

4.2 Meccanismi epigenetici nel CRC: metilazione e modificazioni istoniche

I cambiamenti epigenetici giocano un ruolo fondamentale nello sviluppo del carcinoma del colon-retto e includono metilazione del DNA, modificazioni degli istoni e rimodellamento della cromatina. A differenza delle mutazioni genetiche, le alterazioni epigenetiche influenzano l'espressione genica senza cambiare la sequenza del DNA e sono potenzialmente reversibili. La metilazione del DNA è mediata da enzimi chiamati DNA metiltrasferasi e si riscontra tipicamente nelle isole CpG (CGI). Le CGI sono regioni genomiche caratterizzate da un insolito arricchimento nei dinucleotidi CpG e si trovano comunemente nei promotori di quasi il 60% dei geni umani. La metilazione delle CGI associate ai promotori dei geni è tipicamente associata al silenziamento genico. Durante lo sviluppo tumorale, la demetilazione dell'intero genoma e l'ipermetilazione dei promotori avvengono per lo più simultaneamente. Tra tutti i meccanismi epigenetici, la metilazione del DNA ha acquisito particolare interesse quando è stato scoperto che l'ipermetilazione del promotore associata alla repressione trascrizionale costituisce un evento precoce del processo di carcinogenesi. Nel genoma del CRC centinaia di geni mostrano una metilazione aberrante del DNA, inclusi i geni coinvolti nella via di segnalazione Wnt/ β -catenina, quali APC e AXIN2, i geni coinvolti nella riparazione del DNA MLH1 e MLH2, i

geni associati alla regolazione del ciclo cellulare come CDKN2A, ed i geni della via di segnalazione RAS quali RASSF1A e RASSF1B.

Un altro meccanismo epigenetico cruciale nella patogenesi del CRC è rappresentato dalle modificazioni covalenti delle code istoniche, tra le quali particolare rilevanza assume l'acetilazione/deacetilazione degli istoni. *Loci* ipoacetilati normalmente inibiscono l'espressione genica tramite l'azione delle istone deacetilasi (HDAC), mentre istoni iperacetilati sono coinvolti nell'attivazione dell'espressione genica tramite le istone acetiltrasferasi (HAT). Nei tumori gastrointestinali è stata osservata una ipoacetilazione globale delle lisine degli istoni H3 e H4, spesso associata alla sotto-espressione dei geni oncosoppressori. Nel CRC è stata riscontrata *up*-regolazione di diverse HDAC, quali HDAC1-3, HDAC5 e HDAC7, associata alla *down*-regolazione dei geni coinvolti nella *pathway* Wnt. Inoltre, HDAC2 è *up*-regolata nelle fasi iniziali della carcinogenesi. La sovra-espressione di HDAC2 è accompagnata dalla ipoacetilazione degli istoni H4K12 e H3K18 durante la progressione da adenoma a CRC, il che suggerisce che l'aumentata espressione di HDAC2 e la conseguente perdita dell'acetilazione sono significativamente associate alla progressione del CRC⁹⁴.

4.3 Diagnosi e terapia

Lo *screening* per il carcinoma del colon-retto è eseguito sui pazienti per ricercare la neoplasia prima che questa si manifesti. I pazienti positivi al *test* di *screening*, quale la ricerca di sangue occulto nelle feci, vengono sottoposti ad ulteriori indagini diagnostiche. Il carcinoma del colon-retto è una patologia comune e grave, caratterizzata da una lenta

evoluzione. Una volta sviluppatosi, infatti, il tumore avanza lentamente attraverso diversi stadi. Per questo motivo, lo *screening* rappresenta uno strumento importante per la diagnosi precoce e per organizzare una migliore strategia terapeutica ⁹⁵.

Attualmente, le principali strategie di *screening* impiegate sono due: la ricerca di sangue occulto nelle feci (*Fecal Occult Blood Testing*, FOBT) e l'endoscopia (sigmoidoscopia flessibile e colonscopia)⁹⁶.

Il carcinoma del colon-retto, se precocemente diagnosticato, può essere trattato e spesso curato con le terapie *standard*. La resezione chirurgica del tratto di colon coinvolto rappresenta il trattamento iniziale per molti tumori coloretali. La probabilità di una recidiva è correlata al grado di penetrazione del tumore attraverso la parete intestinale ed al coinvolgimento o meno dei linfonodi. Queste caratteristiche sono alla base del sistema di classificazione TNM e guidano le raccomandazioni per una terapia addizionale⁹⁷. Quando, nel 1994, il ruolo della terapia sistemica per il carcinoma del colon-retto è stato per la prima volta descritto, il 5-fluorouracile era l'unico agente chemioterapico efficace per il trattamento di questo tumore. La molecola, una pirimidina contenente fluoro, agisce tramite l'inibizione della timidilato sintasi, l'enzima che catalizza la tappa limitante nella biosintesi delle pirimidine. Il 5-fluorouracile è generalmente somministrato con la leucovorina, un folato ridotto che stabilizza il legame del 5-fluorouracile alla timidilato sintasi, incrementando quindi l'inibizione della sintesi del DNA. Nei pazienti con carcinoma allo stadio avanzato, il trattamento con 5-fluorouracile e leucovorina riduce la dimensione del tumore del 50% in circa il 20% dei pazienti e prolunga la sopravvivenza media da 6 mesi (senza trattamento) a 11 mesi. Gli effetti collaterali associati al 5-

fluorouracile dipendono dalla via di somministrazione. Quando somministrato tramite bolo, i principali effetti collaterali sono neutropenia, stomatiti e diarrea. La somministrazione endovenosa è associata ad una minore tossicità ematologica e gastrointestinale, ma l'eritrodisestesia palmo-plantare è più comune⁹⁸.

L'irinotecano (Camptosar, anche noto come CPT-11) è un derivato semisintetico della camptotecina, che esercita un effetto citotossico tramite l'interazione con l'enzima topoisomerasi I. Questo enzima svolge la doppia elica del DNA per permetterne la replicazione e la trascrizione, determinando rotture a singolo filamento. Questi danni, reversibili e normalmente riparati, sono invece stabilizzati dalla camptotecina, portando alla frammentazione del DNA e alla morte cellulare. L'irinotecano è un profarmaco che viene idrolizzato nel suo metabolita attivo, SN-38, dalle carbossilesterasi epatiche. I principali effetti tossici dell'irinotecano includono diarrea, soppressione del midollo osseo, nausea, vomito e alopecia.

L'oxaliplatino è un derivato del platino di terza generazione che forma addotti voluminosi di DNA e induce morte cellulare tramite apoptosi.

Studi di laboratorio hanno identificato siti molecolari nei tessuti tumorali che possono servire da specifici *target* per il trattamento. L'obiettivo di una tale strategia terapeutica è l'interruzione delle *pathway* cellulari essenziali per la crescita tumorale, la sopravvivenza e la formazione di metastasi e, potenzialmente, la riduzione degli effetti collaterali associati a chemioterapie meno specifiche. Attualmente, due promettenti classi di composti sono stati introdotti nel trattamento clinico del carcinoma del colon-retto: gli antagonisti dei recettori dei fattori di crescita e gli inibitori dell'angiogenesi⁹⁸.

- Cetuximab

Il recettore del fattore di crescita epidermico (EGFR), è una glicoproteina transmembrana coinvolta nelle vie di trasduzione del segnale che regolano crescita cellulare, differenziamento, proliferazione e morte cellulare programmata. Il recettore è localizzato sulla superficie dell'epitelio ed è sovraespresso in alcuni tumori. Il recettore può essere inattivato tramite l'impiego di anticorpi diretti contro il dominio extracellulare (in cui si trova il sito di legame per il ligando), di ligandi solubili che inibiscono la dimerizzazione del recettore, o di piccole molecole che impediscono la fosforilazione del recettore da parte del suo dominio intracellulare dotato di attività tirosina chinasi. Il cetuximab è un anticorpo monoclonale diretto contro il dominio di legame extracellulare del recettore e recentemente è diventato il primo di questi inibitori ad essere approvato negli Stati Uniti per il trattamento dei carcinomi del colon-retto metastatici. Gli effetti collaterali del cetuximab sono abbastanza modesti, tra i più comuni *rash* simile ad acne e secchezza cutanea⁹⁸.

- Bevacizumab

La considerazione che i tumori inducono la formazione di nuovi vasi sanguigni per sostenere il proprio sviluppo, ha promosso la ricerca di inibitori dell'angiogenesi quali strumenti di controllo della crescita e della diffusione delle cellule tumorali. L'angiogenesi, cioè la formazione di nuovi vasi sanguigni da altri preesistenti, è un processo cruciale nello sviluppo e nella progressione del tumore. Il VEGF (*Vascular Endothelial Growth Factor*) è un potente fattore angiogenico, stimola la proliferazione cellulare, la sopravvivenza e la maturazione dei vasi. Diversi lavori hanno riportato che l'impiego di

un regime chemioterapico combinato a un anticorpo monoclonale anti-VEGF può aumentare la sopravvivenza del paziente⁹⁹. Il bevacizumab (Avastin), un anticorpo diretto contro il VEGF, è stato esaminato in combinazione con agenti chemioterapici in diversi *trials* clinici in pazienti con carcinoma del colon-retto di stadio avanzato. L'aggiunta del bevacizumab ha portato a un notevole incremento del tasso di risposta e prolungato la sopravvivenza complessiva di 4-7 mesi⁹⁸.

4.4 Ruolo dei miRNA nella patogenesi e nell'evoluzione del CRC

Sin dalla prima scoperta relativa al coinvolgimento dei miRNA nella leucemia cronica linfocitica¹⁰⁰ è stato ampiamente dimostrato che queste molecole giocano un ruolo critico nel cancro. L'attività oncogenica dei miRNA può essere tessuto-specifica. L'alterata espressione dei miRNA gioca un ruolo eziologico nell'insorgenza e nella progressione del cancro al colon: il *pattern* di espressione dei miRNA è in grado di discriminare più efficacemente tra tessuti sani e tessuti tumorali CRC rispetto al *pattern* di espressione degli mRNA. Nel 2003, Michael *et al.* pubblicarono il primo studio sui miRNA nel CRC: essi identificarono il **miR-143** ed il **miR-145** come nuovi miRNA deregolati nel cancro al colon¹⁰¹. Da allora, la letteratura sui miRNA nel CRC è cresciuta considerevolmente. I primi miRNA scoperti correlati al CRC, miR-143 e miR-145, agiscono da oncosoppressori e sono *down*-regolati nel CRC rispetto ai colonociti normali. Il miR-143 ha come *target* KRAS (*Kirsten Rat Sarcoma Viral Oncogene Homolog*) e MACC1 (*Metastasis-Associated in Colon Cancer-1*), giocando quindi un importante ruolo nella regolazione del *signalling* di EGFR (*Epidermal Growth Factor Receptor*) e HGFR (*Hepatocyte Growth Factor Receptor*). Il miR-145

ha come bersaglio anche l'oncogene MYC (*v-Myc Avian Myelocytomatosis Viral Oncogene Homolog*), riducendone l'espressione: questo spiega l'inibizione miR-145 dipendente della proliferazione cellulare *in vitro* e *in vivo*¹². Il **miR-21** è uno dei più comuni miRNA *up*-regolati nel cancro, incluso il CRC. Ci sono notevoli effetti a valle di elevati livelli del miR-21: il suo *locus* genico sul 17q23 è amplificato in molti tumori solidi^{102, 103}; la sua espressione è indotta da vari processi correlati al cancro, quali l'infiammazione e l'ipossia. Il miR-21 ha come *target* vari geni oncosoppressori, come PDCD4 (*Programmed cell death 4*), CCL20 [*Chemokine (C-C motif) ligand 20*], CDC25A (*Cell division cycle 25 homolog A*), PTEN (*phosphatase and tensin homolog*), promuovendo quindi la proliferazione cellulare, l'invasione e le metastasi nel CRC¹⁰⁴⁻¹⁰⁷.

Oltre al miR-21, molti altri miRNA possono essere indotti nelle cellule tumorali sotto condizioni ipossiche. In questo gruppo, dei cosiddetti "*hypoxamir*", si trova il **miR-210**, che può modulare la formazione di metastasi indotta da ipossia nelle cellule di CRC. Il miR-210 è spesso *up*-regolato nei tessuti CRC. La sua espressione forzata nelle cellule di CRC promuove la migrazione e l'invasione attraverso la repressione del suo *target* VMP1¹⁰⁸ (*Vacuole Membrane Protein 1*).

Uno dei più noti *cluster* di miRNA oncogenici nel CRC è il *cluster* dei **miR-17-92**. Esso comprende 6 membri (**miR-17**, **miR-18a**, **miR-19a**, **miR-19b**, **miR-20a**, e **miR-92a**) con funzioni oncogeniche, tutti *up*-regolati nel CRC¹⁰⁹.

La famiglia del miR-200 è una delle famiglie di miRNA più conosciute nei mammiferi. Essa comprende cinque membri, localizzati in due differenti *loci* sui cromosomi 1 (**miR-**

200a, miR-200b, miR-429) e 12 (miR-200c e miR-141)¹¹⁰. Si tratta di miRNA oncosoppressori significativamente coinvolti nella EMT, nella repressione della autorigenerazione e del differenziamento delle cellule staminali tumorali, nella modulazione della divisione cellulare e dell'apoptosi. Il miR-200c è particolarmente interessante nel CRC, a causa della sua capacità di regolare la proliferazione cellulare, l'invasione, la migrazione, la EMT e le metastasi. L'espressione del miR-200c è statisticamente più bassa nei campioni clinici di CRC e nelle linee cellulari di CRC altamente metastatiche¹¹¹. Un'altra famiglia di miRNA oncosoppressori coinvolti nella EMT nel CRC è la famiglia del miR-34. Essa include tre miRNA oncosoppressori (miR-34a, miR-34b e miR-34c), che sono regolati da p53; essi sono coinvolti nella repressione delle metastasi promuovendo la Transizione Mesenchima-Epitelio (MET) tramite l'inibizione del fattore di trascrizione SNAI1¹¹² (*Snail family zinc finger-1*).

4.5 Impatto dei lncRNA sulla patobiologia del CRC

Negli ultimi anni, diversi lavori hanno riferito il coinvolgimento dei lncRNA nella genesi e nell'evoluzione del CRC tramite una varietà di meccanismi molecolari. I lncRNA hanno un impatto critico sulle vie di segnalazione del CRC agendo sia da oncogeni sia da oncosoppressori tramite l'interazione con altre molecole regolatorie, come DNA, RNA, e proteine. Sebbene per molti lncRNA sia stata riportata una deregolazione nel CRC, che suggerisce un loro potenziale potere diagnostico/prognostico, i meccanismi molecolari della loro azione nella biologia del CRC sono stati chiariti solo per pochi di essi. Uno dei

più noti lncRNA associati al cancro è **MALAT1** (*Metastasis-Associated Lung Adenocarcinoma Transcript 1*), localizzato sul cromosoma 11q13.1 e lungo 8000 nt. MALAT1 è sovra-espresso nelle metastasi di vari tumori, come il carcinoma polmonare non a piccole cellule, il carcinoma epatocellulare e il sarcoma dello stroma endometriale^{52, 113, 114}. Sono state proposte diverse funzioni intracellulari di MALAT1. Esso potrebbe svolgere un importante ruolo nel metabolismo dei pre-mRNA: è, infatti, associato ai domini SC35 all'interno del nucleo. Si concentra nel nucleolo agendo da "riboregolatore", controllando l'espressione dei suoi geni *target*¹¹⁵. Inoltre, MALAT1 è coinvolto nella regolazione di proteine oncosoppressorie [e.g., PSF (*PTB-associated Splicing Factor*)], e regola anche l'attività di E2F1, fattore di trascrizione essenziale per la progressione del ciclo cellulare¹¹⁶. Il silenziamento *in vitro* di MALAT1 ha mostrato influenzare la migrazione cellulare nel tumore alla vescica. MALAT1 agisce da regolatore negativo dei fattori associati alla transizione epitelio-mesenchima (EMT) ZEB1, ZEB2 e SNAI2, e da regolatore positivo delle E-caderine¹¹⁷. La deplezione di MALAT1 determina mitosi aberranti, con una larga frazione di cellule che si bloccano al confine G2/M, ed una incrementata morte cellulare. È stato recentemente identificato un meccanismo di processamento al 3' per MALAT1: il prodotto primario della trascrizione del *locus* MALAT1 è un lncRNA di 6.7 kb ritenuto nel nucleo ed un ncRNA citoplasmatico di 61 nucleotidi simile a un tRNA, noto come mascRNA¹¹⁸ (*MALAT1-Associated Small Cytoplasmatic RNA*) (Figura 15). La sovra-espressione del segmento di RNA di MALAT1 contenente mascRNA in cellule di CRC ha indotto proliferazione ed invasione cellulare. Mutazioni puntiformi di MALAT1 sono state individuate nelle linee cellulari di CRC e nei tessuti tumorali¹¹⁹.

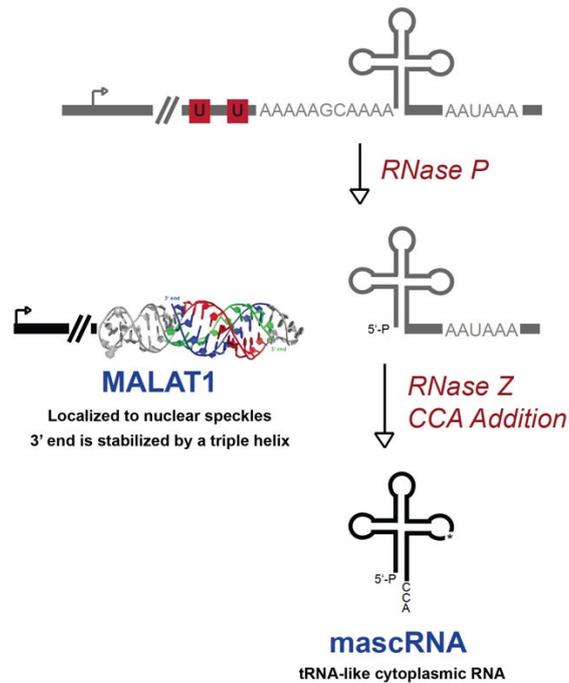


Figura 15. Meccanismo di processingo di MALAT1 (da Wilusz, *Genes Dev.* 2012).

BANCR (*BRAF-Activated Non-protein Coding RNA*) sembra essere strettamente associato con V600E, una delle mutazioni del gene BRAF più frequenti in diversi tumori, incluso il CRC. BANCR è frequentemente sovra-espresso nei tessuti di CRC: tale sovra-espressione è significativamente correlata alle metastasi ai linfonodi ed allo stadio tumorale¹²⁰.

L'aumentata espressione di BANCR incrementa la migrazione delle linee cellulari di CRC, mentre il suo *knockdown* la inibisce. BANCR induce la EMT influenzando l'espressione dei marcatori epiteliali e mesenchimali attraverso un meccanismo dipendente da MEK (*Mitogen-activated protein kinase kinase 7*) /ERK, contribuendo quindi alla progressione del CRC. Studi di associazione *genome-wide* (GWAS) hanno identificato una serie di *loci* di suscettibilità, che sono legati alla sensibilità a diverse malattie (incluso il CRC), sul cromosoma 8q24¹²¹. È interessante che tale segmento (di circa 2Mb) si trovi all'interno di

un'ampia regione genica nella quale non sono presenti geni codificanti per proteine (deserto genico), ma diversi geni non codificanti mappano su di essa (*i.e.*, CCAT1, CCAT2, PCAT2 e PRNCR1). Il polimorfismo a singolo nucleotide (SNP) rs6983267, che mappa sulla regione cromosomica 8q24.21, è stato strettamente associato ad un maggiore rischio di sviluppare il CRC¹²². È stato riportato che la regione genomica contenente rs6983267 comprende elementi *enhancer*, e le varianti alleliche conferiscono una diversa affinità di legame a TCF7L2 [*Transcription factor 7-like 2 (T-cell specific, HMG-box)*], un fattore di trascrizione che ha un ruolo centrale nell'attivazione trascrizionale dei geni *target* di Wnt^{123, 124}.

Lo stato polimorfico del lncRNA **CCAT2** (*Colon Cancer Associated Transcript 2*), che contiene rs6983267, influenza l'espressione di CCAT2: l'allele G produce più trascritti CCAT2 nel CRC. CCAT2 interagisce con TCF7L2 e porta alla *up*-regolazione di MYC, del miR-17-5p e del miR-20; esso inoltre attiva la via di segnalazione Wnt¹²⁵. Inoltre, CCAT2 è esso stesso un *target* a valle di Wnt, il che suggerisce l'esistenza di un *loop* a *feedback* positivo. L'isoforma lunga di CCAT1 (CCAT1-L) è sovra-espressa e positivamente correlata allo stadio tumorale e alla progressione del CRC. CCAT1-L è localizzato nel nucleo, mentre l'isoforma corta (CCAT1-S) è stata trovata nel citoplasma. CCAT1-L può interagire con un *enhancer* trascrizionale di MYC (MYC-335) tramite la formazione di un *loop*; l'*enhancer* a sua volta interagisce con il promotore di MYC. Il *knockdown* di CCAT1-L porta ad una diminuzione del livello di mRNA di MYC, suggerendo fortemente che questo lncRNA possa regolare l'espressione di MYC in *cis*¹²⁶.

Un altro lncRNA che mappa sul cromosoma 8q24 è **PVT1** (*Plasmacytoma Variant Translocation 1*): PVT1 è localizzato a valle di MYC sul cromosoma 8q24 e produce un'ampia varietà di RNA "spliced", come un *cluster* di sei miRNA (*i.e.*, miR-1204, miR-1205, miR-1206, miR1207-5p, miR1207-3p, e miR-1208)¹²⁷. PVT1 è *up*-regolato nel CRC a causa di amplificazione genica sul cromosoma 8q24. Il *knockdown* di PVT1 nelle cellule di CRC porta a una significativa riduzione della proliferazione e dell'invasione attraverso l'attivazione di TGF- β e della via apoptotica. L'aumentata espressione di PVT1 è necessaria per avere elevati livelli di proteina MYC nelle cellule di CRC aventi la regione 8q24 amplificata. PVT1 e l'espressione della proteina MYC correlano nei tumori primari, mentre l'ablazione di PVT1 dalla linea cellulare HCT-116 guidata da MYC riduce il suo potenziale tumorigenico¹²⁸. Sorprendentemente, PVT1 è anche un gene *target* di p53: p53 lega e attiva un elemento di risposta nelle vicinanze del miR-1204, e induce i trascritti di PVT1 con conseguente *up*-regolazione del miR-1204¹²⁹. L'espressione ectopica del miR-1204 porta a un incremento dei livelli di p53 e causa morte cellulare tramite un meccanismo in parte dipendente da p53.

Il lncRNA maggiormente studiato in diverse malattie neoplastiche è **HOTAIR** (*Hox Transcript Antisense Intergenic RNA*). Esso è localizzato all'interno del *cluster* genico HOXC (Homeobox C) sul cromosoma 12 ed è coespresso con i geni HOXC¹³⁰. HOTAIR interagisce con PRC2; esso funziona da *scaffold* per l'assemblaggio di PRC2 sul *cluster* genico HOXD, inducendo la trimetilazione della lisina-27 dell'istone H3 (H3K27me3) del *locus* HOXD¹³¹ (Figura 16). Tramite il silenziamento di geni soppressori di metastasi multiple (come HOXD10, PGR, e i geni della famiglia delle protocaderine), HOTAIR regola

epigeneticamente l'espressione di HOXD e promuove la formazione di metastasi nel cancro al seno²⁷. Studi di espressione sul CRC rivelano una stretta correlazione tra l'espressione di HOTAIR ed i membri del complesso PRC2 (*i.e.*, SUZ1, EZH2, e H3K27me3): questo suggerisce che l'espressione di HOTAIR sia associata con la funzione epigenetica di PRC2, il quale è coinvolto nel mantenimento dello stato mesenchimale e indifferenziato delle cellule di CRC.

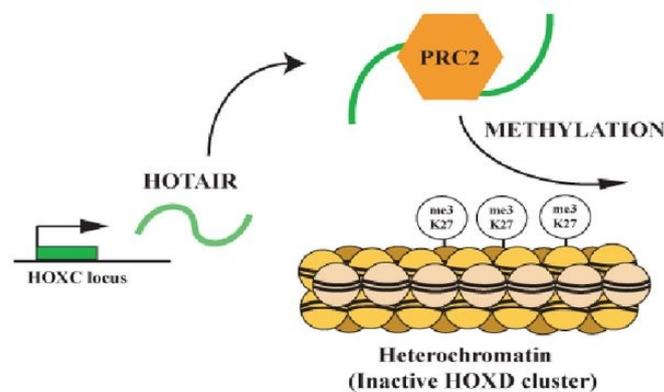


Figura 16. Meccanismo di interazione tra HOTAIR e PRC2 (da Calet et al., *ScientificWorldJournal*. 2010).

Allo stesso modo di HOTAIR, anche il lncRNA **CRNDE** (*Colorectal Neoplasia Differentially Expressed*) sembra essere fisicamente e funzionalmente associato a PRC2. È stata evidenziata una sovrapposizione tra le liste di geni influenzati dal *knockdown* di CRNDE e PRC2. Questi dati potrebbero suggerire un coinvolgimento di CRNDE nel rimodellamento epigenetico della cromatina, e specificamente nella *down*-regolazione dell'espressione genica tramite metilazione istonica mediata dal complesso PRC2¹³². Il promotore di CRNDE è legato da diversi fattori di trascrizione associati alla pluripotenza (*i.e.*, MYC, MYCN) e il *knockdown* di CRNDE diminuisce l'espressione di diversi marcatori della

pluripotenza (*i.e.*, SOX2, KLF4, NANOG, e OCT4). L'espressione di CRNDE è massima nei primi stadi dello sviluppo dei mammiferi per poi decrescere progressivamente. Essa è richiesta per il mantenimento della pluripotenza in cellule staminali embrionali di topo ed è potenzialmente coinvolta nella tumorigenesi¹³³.

TUSC7 (*Tumor Suppressor Candidate 7*) è un oncosoppressore regolato da p53, che riduce la crescita cellulare nel CRC sia *in vitro* sia *in vivo*. Più specificamente, il suo quarto esone (contenente due siti di legame per il miR-211) è responsabile dell'inibizione della crescita cellulare nel tumore. Studi *in vivo* hanno confermato che TUSC7 è in grado di legare il miR-211 inducendo la sua *down*-regolazione. È stato dimostrato che il miR-211 promuove la crescita cellulare nelle linee di CRC. Di conseguenza, TUSC7 agisce da spugna per il miRNA. TUSC7 è, a sua volta, un *target* del miR-211, indicando l'esistenza di un reciproco *loop a feedback* negativo tra TUSC7 ed il miR-211⁷⁰.

H19 è un oncogene fetale soggetto a *imprinting* paterno, espresso sul cromosoma materno. È altamente espresso dalle prime fasi dell'embriogenesi fino allo stadio fetale in molti organi, ma è quasi completamente *down*-regolato dopo la nascita¹³⁴. H19 è *up*-regolato in molti tumori, incluso il CRC, e agisce da precursore primario del miR-675 nei mammiferi. L'mRNA di RB è un *target* diretto del miR-675: infatti, il *knockdown* del mir-675 aumenta l'espressione di RB e allo stesso tempo riduce la crescita cellulare nel CRC. Al contrario, mutazioni con guadagno di funzione nel miR-675 causano una *down*-regolazione di RB e aumentano la crescita delle cellule tumorali. Ipossia e TGF- β inducono, in concomitanza, H19 e miR-675, insieme all'induzione dei marcatori della EMT e alla repressione della E-caderina⁵⁷. È interessante il fatto che H19 possieda siti di legame, sia canonici sia non

canonici, per la famiglia dei miRNA let-7. H19 è in grado di modulare negativamente la disponibilità di let-7 agendo da spugna molecolare¹³⁵.

Recentemente, il ruolo di un altro lncRNA, funzionalmente associato alla EMT, è stato caratterizzato nel CRC. **N-BLR** è un lncRNA coinvolto nella via apoptotica: la sua inibizione porta a una *down*-regolazione di XIAP (*X-linked Inhibitor of Apoptotic Protein*) e al conseguente incremento della morte cellulare. N-BLR, inoltre, promuove l'invasione e la migrazione modulando l'espressione della vimentina e della E-caderina. N-BLR sembra sia regolato dai membri della famiglia del miR-200 (*i.e.*, miR-141, e miR-200c). La famiglia del miR-200 è stata strettamente associata alla EMT: infatti, i suoi componenti hanno come *target* i fattori di trascrizione ZEB1/ZEB2, che sono repressori dell'espressione della E-caderina. L'aumentata espressione di N-BLR in campioni CRC attirerebbe i miR-141/miR-200c endogeni disponibili sostituendosi al loro *target* ZEB1, con successivo incremento del livello di quest'ultimo. Si osserva, di conseguenza, una diminuzione della E-caderina; inoltre, questo conferirebbe un fenotipo mesenchimale alle cellule, correlato alla maggiore invasività delle cellule di CRC¹³⁶.

Urothelial Cancer Associated 1 (**UCA1**), un oncogene fetale coinvolto nello sviluppo embrionale e nella carcinogenesi¹³⁷ è stato per la prima volta scoperto nel cancro alla vescica, e il suo trascritto più lungo è anche chiamato *Cancer Upregulated Drug-Resistant transcript* (CUDR). Sono state descritte tre diverse isoforme di UCA1 lunghe 1.4 kb, 2.2 kb, e 2.7 kb ma, in genere, si studia l'isoforma più abbondante di 1.4 kb¹³⁸.

Il gene di UCA1 codifica 3 esoni localizzati sul cromosoma 19 ed è altamente espresso nelle cellule tumorali. Infatti, la sua trascrizione è *up*-regolata da diverse *pathway*

oncogeniche. È stato dimostrato che ETS-2 (*Ras-Responsive Transcription Factor*) regola l'espressione di UCA1 sia nel cancro alla vescica che nel cancro al colon-retto^{139, 140}.

UCA1 è *up*-regolato dal principale induttore della transizione epitelio-mesenchima (EMT), TGF, nelle cellule del tumore allo stomaco e al seno^{141, 142}.

BMP9 (*Bone Morphogenetic Protein-9*) ha un ruolo ambiguo nella progressione tumorale, ma è stato recentemente scoperto che esso induce l'espressione di UCA1 nelle cellule del tumore alla vescica¹⁴³. È interessante notare come in queste cellule l'espressione di UCA1 venga stimolata anche durante l'ipossia tramite HIF1 (*Hypoxia-Inducible Factor-1*) e che la secrezione di UCA1 negli esosomi aumenta in queste condizioni^{144, 145}.

Le analisi di espressione di UCA1 e i dati relativi alla sopravvivenza dei pazienti estrapolati dal *dataset* TCGA mostrano che la sua espressione è correlata ad un aumentato tasso di rischio in diversi tipi di tumore, in particolare nell'adenocarcinoma pancreatico. UCA1, come altri lncRNA, svolge diverse funzioni e può influenzare sia la regolazione epigenetica sia la trascrizione genica, così come la regolazione post-trascrizionale agendo come ceRNA per diversi miRNA. UCA1 è stato studiato in molti tipi di cellule tumorali, incluso il cancro al colon retto¹³⁸.

I livelli di UCA1 sono marcatamente aumentati nei tessuti e nelle cellule di CRC rispetto ai controlli, e questa elevata espressione è significativamente correlata ad una maggiore dimensione del tumore, al minore differenziamento istologico, e ad una maggiore infiltrazione del tumore. Inoltre, i pazienti con un'elevata espressione di UCA1 presentano una prognosi più sfavorevole rispetto ai pazienti con espressione di UCA1 più bassa. UCA1 influenza, altresì, la proliferazione, l'apoptosi e la progressione attraverso il ciclo

cellulare delle cellule di CRC. Infatti, i dati di RT-qPCR indicano che i livelli di espressione di UCA1 sono significativamente più alti nelle linee cellulari LoVo, CaCO2, SW480 e HCT116 rispetto alle cellule CCC-HIE-2 (linee cellulari della mucosa intestinale normale). L'espressione ectopica di UCA1 nelle cellule LOVO promuove la proliferazione e la progressione attraverso il ciclo cellulare e inibisce l'apoptosi nelle cellule LOVO, mentre il silenziamento genico di UCA1 nelle SW480 inibisce la proliferazione e la progressione attraverso il ciclo cellulare e promuove l'apoptosi. Questi dati suggeriscono che UCA1 potrebbe svolgere un ruolo importante nella regolazione dell'insorgenza e dello sviluppo del CRC¹³⁷.

Brabagallo *et al.*, analizzando l'espressione di 17 lncRNA e 31 circRNA nelle biopsie e negli esosomi del siero di pazienti affetti da CRC, hanno rilevato l'*up*-regolazione di UCA1 nelle biopsie di CRC e la sua *down*-regolazione negli esosomi del siero. Le analisi *in vitro* (i.e., RNA silencing e inibizione delle *mitogen-activated protein kinase* [MAPK]) e *in silico* (i.e., correlazione di espressione e predizione dell'interazione RNA-RNA), hanno mostrato che (1) UCA1 potrebbe essere controllato dalle MAPK tramite CEBP; (2) UCA1 sequestra i miRNA miR-135a, miR-143, miR-214, e miR-1271, proteggendo ANLN, BIRC5, IPO7, KIF2A, e KIF23 dalla degradazione indotta dai miRNA; e (3) interagisce con le 3'UTR degli mRNA, impedendo il legame dei miRNA. Inoltre, UCA1 e il suo antisenso LINC01764 potrebbero interagire e mascherare reciprocamente i loro siti di legame per i miRNA¹⁴⁶.

Molti studi riportano che UCA1 potrebbe agire come miRNA *sponge* in diversi modelli *in vitro*. È stato riportato che l'espressione forzata di UCA1 risulta in una ridotta espressione

dei suoi miRNA *target*, o, viceversa, la ridotta espressione di UCA1 determina aumentati livelli dei miRNA^{72, 147-150}.

5. *Il processo metastatico*

Le metastasi sono la prima causa di morte nei pazienti con cancro. Il cancro metastatico è responsabile di circa il 90% delle morti e, approssimativamente, circa il 50% di tutti i pazienti presenta metastasi rilevabili al momento delle diagnosi, mentre nella maggior parte dei pazienti le micrometastasi non possono essere identificate tramite le tecniche diagnostiche attualmente utilizzate¹⁵¹. Le metastasi rappresentano il prodotto finale di un processo biologico multi-stadio noto come cascata metastatica, per mezzo della quale le cellule epiteliali nel tumore primitivo migrano localmente attraverso la matrice extracellulare (ECM) circostante e lo strato di cellule stromali (Figura 17a), entrano nel lume dei vasi sanguigni (Figura 17b), sopravvivono al trasporto attraverso i vasi (Figura 17c), si fermano in un nuovo sito distale (Figura 17d), fuoriescono dai vasi e invadono il parenchima del tessuto distale (Figura 17e), inizialmente sopravvivono in questo nuovo ambiente formando micrometastasi (Figura 17f) e in seguito, riprendono a proliferare nel sito metastatico, generando quindi macrometastasi clinicamente rilevabili¹⁵² (Figura 17g). Diversi geni e i loro prodotti sono coinvolti nella cascata metastatica, mediando fasi cruciali come l'inizio e la progressione, così come funzioni organo-specifiche delle metastasi (virulenza). Questi prodotti genici includono proteasi, chemochine, citochine e i loro recettori, fattori angiogenici, chinasi intracellulari e transmembrana, molecole di adesione e componenti della matrice extracellulare (ECM). Più di recente, è emerso un importante impatto delle molecole di *non-coding RNA* nel processo metastatico. I miRNA e

i lncRNA modulano le metastasi tramite *network* regolatorie, e possono promuovere o inibire le metastasi in diversi tipi di tumore¹⁵³.

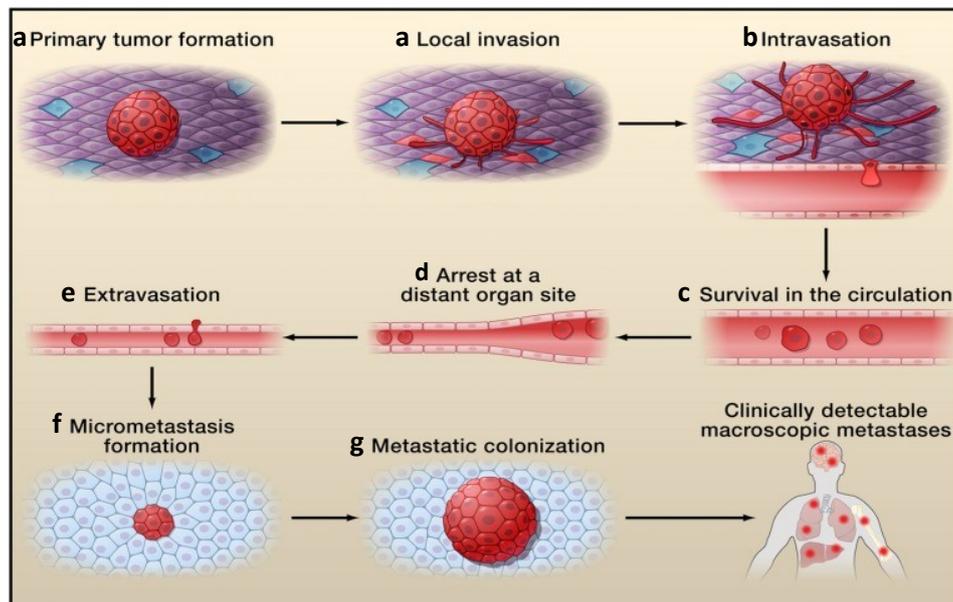


Figura 17. Rappresentazione della cascata metastatica: a) le cellule nel tumore primario migrano localmente attraverso la matrice extracellulare (ECM) circostante e lo strato di cellule stromali, b) entrano nel lume dei vasi sanguigni, c) sopravvivono al trasporto attraverso i vasi, d) si fermano in un nuovo sito distale, e) fuoriescono dai vasi e invadono il parenchima del tessuto distale, f) inizialmente sopravvivono in questo nuovo ambiente formando micrometastasi e g) in seguito, riprendono a proliferare nel sito metastatico, generando quindi macrometastasi clinicamente rilevabili (da Valastyan and Weinberg. *Cell*. 2011).

5.1 Ruolo della Transizione Epitelio-Mesenchima (EMT) nel processo metastatico

La transizione epitelio-mesenchima (EMT) è un programma di sviluppo che permette alle cellule epiteliali immobili di acquisire la capacità di migrare ed invadere come cellule singole. La EMT è stata inizialmente descritta come parte integrante della morfogenesi durante lo sviluppo embrionale, e solo in seguito è stata osservata in diversi processi patologici, inclusi la guarigione delle ferite, la fibrosi e le metastasi del cancro. Le cellule tumorali riattivano la EMT per acquisire alterazioni molecolari che consentono la perdita

delle caratteristiche epiteliali e l'acquisizione di un fenotipo mesenchimale. Nei tessuti distali, le cellule tumorali devono adattarsi al nuovo micro-ambiente e creare una nicchia metastatica di supporto, passando da una modalità migratoria ad una modalità proliferativa per generare le macrometastasi. Un evento chiave nel promuovere tali proprietà invasive delle cellule tumorali è la EMT. I punti chiave di questo processo sono la perdita delle giunzioni aderenti e della polarità apico-basale e l'acquisizione di un fenotipo mesenchimale con incremento della motilità ed invasività¹⁵⁴. Le cellule epiteliali sono caratterizzate da interazioni intatte cellula-cellula mediate da molecole di adesione come le E-caderine e la citocheratina all'interno delle *tight junctions*, delle giunzioni aderenti, dei desmosomi, e delle *gap junctions*. La E-caderina è il prototipo delle caderine di tipo I. Queste caderine mediano interazioni omofile formando legami tra uno o più domini Ig presenti nella loro regione extracellulare e connettendosi indirettamente ai microfilamenti di actina tramite l'alfa e la beta catenina nel citoplasma. Nelle cellule epiteliali i primi contatti sono mediati dalle E-caderine che clusterizzano in piccoli complessi giunzionali che, in seguito, si espandono formando giunzioni aderenti stabili e promuovendo la formazione dei desmosomi. L'E-caderina è richiesta per il mantenimento di giunzioni stabili: difatti, anticorpi anti E-caderina possono distruggere tali contatti e indurre il fenotipo mesenchimale. Durante la progressione tumorale si osserva la *down-regolazione* delle E-caderine. Nella maggior parte dei casi, tale *down-regolazione* è dovuta a cambiamenti epigenetici, quali la repressione trascrizionale e l'ipermetilazione del promotore. La comprensione della regolazione della produzione della E-caderina è aumentata in seguito all'identificazione di specifici repressori trascrizionali, il primo dei

quali è stato SNAIL. In risposta a vari segnali cellula e tessuto-specifici le cellule epiteliali sovra-esprimono un gruppo di fattori di trascrizione che inducono la EMT (EMT-TF) al fine di realizzare tutti i cambiamenti morfologici, cellulari e molecolari durante la EMT. I principali EMT-TF includono i membri della famiglia dei fattori di trascrizione *zinc finger* SNAIL (SNAIL1 e SNAIL2), e un membro strettamente correlato appartenente alla stessa famiglia (SLUG); i fattori di trascrizione *basic helix-loop-helix* della famiglia TWIST (TWIST1 e TWIST2), e le proteine *homeobox* ZEB1 e ZEB2 (*zinc finger e-box binding*). I membri della superfamiglia dei TF SNAIL e ZEB sono in grado di legare le sequenze E-box sul promotore della E-caderina reprimendone la trascrizione. È stato dimostrato che SIP1, un'altra proteina *zinc-finger E-box-binding*, agisce anch'essa come forte repressore della E-caderina¹⁵¹. Infine, varie *pathway* di sviluppo, come quella del TGF- β , di WNT, NOTCH, e i recettori dei fattori di crescita tirosina chinasi sono coinvolti nell'indurre la EMT in determinate condizioni fisiologiche. In risposta ai vari stimoli che inducono l'EMT e alle loro *pathway* di segnalazione a valle, i fattori di trascrizione della EMT (EMT-TF) sono attivati per dare avvio al programma (Figura 18). L'attivazione dei EMT-TF porta alla *down*-regolazione dei geni codificanti le proteine delle giunzioni epiteliali, quindi al disassemblaggio delle giunzioni aderenti, dei desmosomi e delle *tight junctions* e alla perdita della polarità apico-basale. Poiché tali giunzioni non fungono semplicemente da supporto strutturale all'epitelio ma regolano anche diverse vie di segnalazione tramite le proteine ad esse associate, i cambiamenti nelle giunzioni epiteliali influenzano un gran numero di *pathway* a valle che possono ulteriormente promuovere la EMT e l'invasione. Ad esempio, la β -catenina rilasciata dalle giunzioni aderenti può potenzialmente entrare

nel nucleo per guidare l'espressione dei geni *target* WNT e potenziare la EMT. Allo stesso tempo l'espressione dei geni mesenchimali come la N-caderina, la fibronectina, la vimentina, è *up*-regolata. In particolare, è aumentata l'espressione di una serie di caderine non epiteliali e di proteine della superficie cellulare critiche per la migrazione cellulare. Mentre la rete di citocheratina ancorata ai desmosomi nelle cellule epiteliali viene distrutta, l'*up*-regolazione della vimentina porta alla riorganizzazione dell'actina del citoscheletro in una rete citoplasmatica di filamenti intermedi. Questa transizione verso lo stato mesenchimale facilita la motilità cellulare e la formazione di nuove protrusioni di membrana, come gli invadopodi, per la degradazione della matrice¹⁵⁴. Le fibre di actina avviano e aumentano anche la contrattilità cellulare. Infine, le protrusioni cellulari e l'espressione delle metalloproteinasi si concludono con la degradazione della ECM, la migrazione cellulare, e l'acquisizione della capacità invasiva¹⁵⁵.

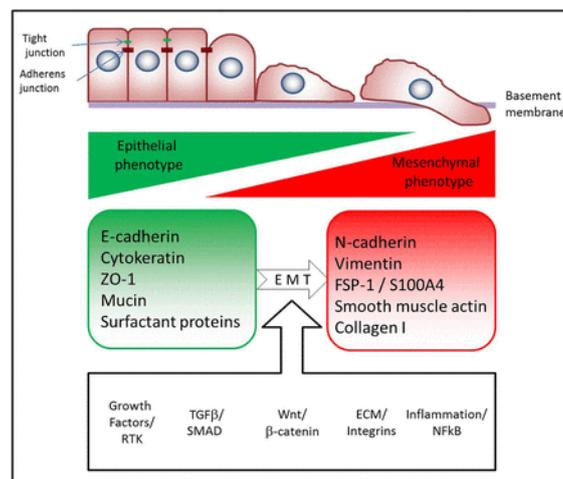


Figura 18. Principali caratteristiche della Transizione Epitelio-Mesenchima: in risposta ai vari stimoli che inducono la EMT vengono attivate varie pathway di segnalazione a valle (e.g., TGF-β, WNT) con conseguente attivazione dei fattori di trascrizione della EMT (EMT-TF) che danno avvio al programma (da Bartis et al., Thorax. 2014).

5.2 Coinvolgimento dei ncRNA nel processo metastatico

Numerose evidenze sperimentali suggeriscono che i miRNA giocano un ruolo cruciale nella regolazione degli eventi molecolari che conducono alle metastasi. Essi controllano l'espressione di geni pro-metastatici in diversi tipi di cancro e possono promuovere o reprimere le metastasi attraverso vari meccanismi¹⁵⁶. Sono stati identificati numerosi miRNA coinvolti nella regolazione della EMT (Tabella 2) che agiscono attraverso la *network* trascrizionale che regola la EMT o regolando le *pathway* molecolari coinvolte nella EMT. I fattori di trascrizione ZEB1/ZEB2 promuovono la EMT principalmente attraverso la repressione della E-caderina e l'attivazione della N-caderina e della vimentina. La E-caderina può essere sia un *target* diretto dei miRNA, oppure può essere regolata in maniera indiretta tramite la regolazione dei fattori di trascrizione. I miRNA miR-200c, miR-10b, e miR-9 inattivano direttamente l'mRNA della E-caderina¹⁵⁷⁻¹⁵⁹. Molti miRNA sono noti regolare l'espressione di ZEB1/2 nelle cellule tumorali. La famiglia del miR-200, un noto miRNA oncosoppressore, è coinvolta nella inibizione della EMT e della staminalità¹¹⁰. La ridotta espressione dei membri della famiglia del miR-200 è stata documentata in vari tumori inclusi il cancro al seno, il CRC, il cancro alla prostata e ai reni¹⁶⁰. La famiglia del miR-200 regola l'espressione di ZEB1/2 a livello post-trascrizionale inibendo i loro messaggeri e portando alla repressione della EMT nelle cellule tumorali. Numerosi miRNA regolano l'espressione dei fattori di trascrizione SNAI1 e SNAI2 (SLUG). Il miR-30a è significativamente *down*-regolato nei tessuti NSCLC (carcinoma polmonare non a piccole cellule). Il ripristino del miR-30a inibisce la EMT reprimendo l'espressione di SNAIL1 nelle linee cellulari NSCLC^{161, 162}.

miRNA	miRNA level status and role		Target
	Expression	Function	
miR-30a	Downregulated	EMT-Inhibitor	SNAI1
miR-34a	Downregulated	EMT-Inhibitor	SNAI1
miR-101	Downregulated	EMT-Inhibitor	ZEB1/ZEB2
miR-132	Downregulated	EMT-Inhibitor	ZEB2
miR-138	Downregulated	EMT-Inhibitor	ZEB2, vimentin, EZH2
miR-139-5p	Downregulated	EMT-Inhibitor	ZEB1/ZEB2
miR-186	Downregulated	EMT-Inhibitor	TWIST
miR-200	Downregulated	EMT-Inhibitor	ZEB1/ZEB2
miR-203	Downregulated	EMT-Inhibitor	SNAI2
miR-205	Downregulated	EMT-Inhibitor	ZEB1/ZEB2
miR-503	Downregulated	EMT-Inhibitor	c-myc
miR-655	Downregulated	EMT-Inhibitor	ZEB1, TGFB2
miR-93	Upregulated	EMT-Promoter	FOXA1
miR-181b-3p	Upregulated	EMT-Promoter	YWHAG
miR-374a	Upregulated	EMT-Promoter	WIF1, PTEN, WNT5A
miR-503-3p	Upregulated	EMT-Promoter	SMAD2
miR-1269a	Upregulated	EMT-Promoter	HOXD10

Tabella 2. miRNA coinvolti nella Transizione Epitelio-Mesenchima (da Jafri et al., *Semin Cancer Biol.* 2017).

Numerose ricerche hanno dimostrato che i lncRNA svolgono un importante ruolo nell'inizio e nella progressione di vari tumori. Evidenze sempre maggiori suggeriscono che i lncRNA siano regolatori cruciali della EMT, influenzando così il processo metastatico. LINC01186 è *down*-regolato nei tessuti di tumore ai polmoni rispetto ai tessuti adiacenti sani e analisi bioinformatiche suggeriscono sia associato al TGF- β e partecipi alla EMT. Inoltre, il *knockdown* di LINC01186 promuove la migrazione cellulare e l'invasione, mentre la sua sovra-espressione previene le metastasi. LINC01186 è regolato da SMAD3. È stato dimostrato che aumentati livelli di SMAD2 e SMAD3 promuovono la crescita e la formazione di metastasi nel tumore ai polmoni. Il *knockdown* di LINC01186 influenza i livelli di espressione di diversi marcatori molecolari della EMT. È stata ipotizzata l'esistenza di un aberrante asse di segnalazione TGF- β /SMAD3/LINC01186 che conduce ad una maggiore aggressività del carcinoma polmonare attraverso l'induzione della EMT¹⁶³.

Il lncRNA NEAT è altamente sovra-espresso nel tumore al seno e la sua espressione è strettamente correlata alla dimensione del tumore e alle metastasi ai linfonodi. Inoltre, la

repressione di NEAT1 porta ad una riduzione della proliferazione e delle metastasi. L'analisi tramite *Western blot* ha rivelato che i livelli di β -catenina ed N-caderina diminuiscono, mentre i livelli di E-caderina aumentano dopo il *knockdown* di NEAT1. Ciò dimostra che NEAT1 può agire da oncogene nel tumore al seno, dove può promuovere la proliferazione e le metastasi¹⁶⁴.

5.3 miRNA e lncRNA come potenziali candidati per la terapia anti-metastatica

Il tumore metastatico è una condizione difficile da gestire principalmente a causa della sua natura sistemica e ai frequenti meccanismi di chemio resistenza ai regimi terapeutici, così come la variabilità delle risposte ai farmaci antitumorali. I miRNA stanno emergendo come promettenti candidati per la terapia anti-metastatica. Generalmente, vengono utilizzati due approcci basati sui miRNA. La terapia anti-oncomiR è usata per silenziare la funzione dei miRNA oncogenici sovra-espressi, mentre i *miRNA-mimics* sono impiegati per ripristinare la funzione dei miRNA *down*-regolati (Figura 19). L'attività degli oncomiR può essere inibita tramite la somministrazione di oligonucleotidi sintetici anti-miRNA (AMO), di "*Locked Nucleic Acid Anti-miRNA*" (LNA) e "*miRNA sponges*". Questi anti-miR, sequestrano o degradano i miRNA maturi o competono con il messaggero *target* per il legame al miRNA. Sono stati identificati numerosi potenziali miRNA. Ad esempio, il miR-96 e il miR-182 sono sovra-espressi nella maggior parte dei tumori e promuovono la

migrazione e l'invasività delle cellule tumorali, potendo essere, quindi, utilizzati come potenziali *target* per inibire la formazione di metastasi¹⁵⁶.

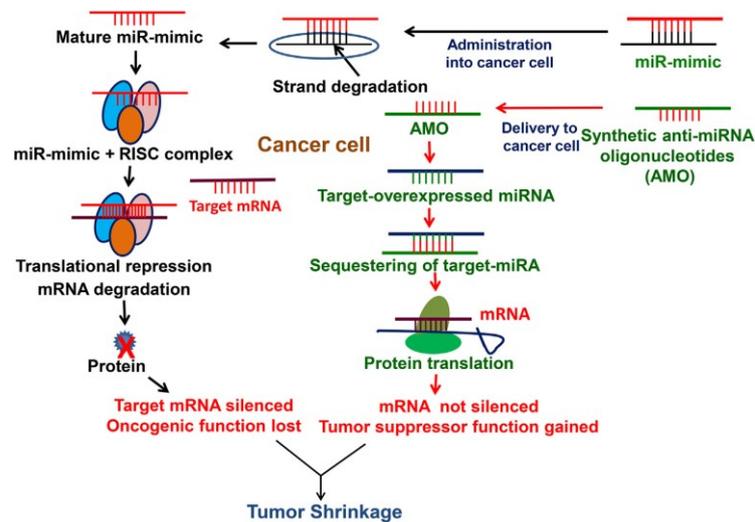


Figura 19. Approcci terapeutici basati sui miRNA: a) gli anti-miRNA oligonucleotides (AMOs) possono essere introdotti nelle cellule tumorali per silenziare la funzione del miRNA sovra-espresso. Gli AMOs inibiscono il miRNA target in modo tale che possa riprendere la traduzione dell'mRNA oncosoppressore; b) utilizzo di miRNA-mimics per ripristinare la funzione del miRNA down-regolato (da Jafri et al., *Semin Cancer Biol.* 2017).

Lo stesso tipo di approccio potrebbe essere applicato ai lncRNA. L'impiego dei lncRNA come *target* farmacologici presenta diversi vantaggi rispetto ad altre strategie terapeutiche. I lncRNA sono molecole lineari, di piccole dimensioni e risultano, pertanto, più "druggable" rispetto alle proteine; il loro *targeting*, infatti, è principalmente basato sulla complementarità di sequenza tra acidi nucleici. La loro progettazione sarebbe, quindi, più semplice, e la loro sintesi meno dispendiosa. La deplezione dei lncRNA oncogenici potrebbe essere ottenuta tramite l'impiego di siRNA o shRNA (*small hairpin RNA*): tali molecole mostrano una grande selettività nei confronti dell'RNA e un'elevata efficienza di *knockdown*¹⁶⁵. Tuttavia, negli ultimi anni sono stati sviluppati nuovi metodi che stanno

sostituendosi all'approccio basato sull'RNA *interference* (RNAi) (*i.e.*, ASOs e aptameri). Gli ASOs (*antisense oligonucleotides*) sono molecole di DNA *single strand* o RNA (8-50 nt) con sequenza complementare ai loro RNA *target*. Quando legano gli RNA *target*, l'RNasi H1 riconosce l'*eteroduplex* DNA: RNA/RNA: RNA e catalizza il clivaggio delle molecole di RNA. Gli aptameri sono piccoli oligonucleotidi di DNA o RNA o peptidi con una struttura tridimensionale stabile che si lega specificamente ai propri *target* (*i.e.*, RNA, proteine, e piccole molecole), adattandosi alla loro forma tridimensionale; il loro meccanismo di *targeting* è, pertanto, indipendente dalla complementarità di sequenza¹⁶⁶ (Figura 20). Una conoscenza più approfondita sui *target* molecolari dei lncRNA insieme allo sviluppo di nuovi metodi selettivi per la "*delivery*" degli RNA nelle cellule tumorali, potrebbe portare a grandi benefici per il trattamento dei tumori metastatici. L'elevato numero di lncRNA putativamente "*druggable*" e le strategie terapeutiche basate sui lncRNA si potrebbero rivelare, pertanto, strumenti molto promettenti per integrare le strategie terapeutiche attualmente in uso.

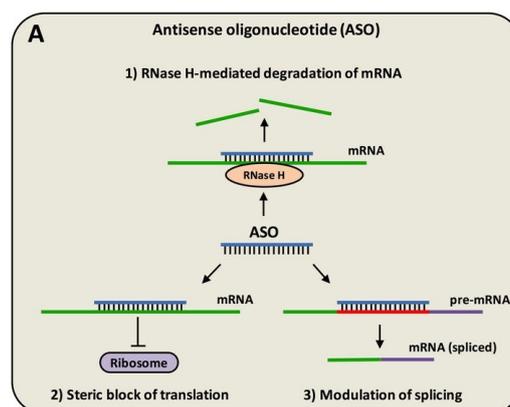


Figura 20. Antisense oligonucleotides (ASOs) (da Rossor et al., Pract Neurol. 2018).

In questo lavoro abbiamo esplorato l'esistenza di lncRNA la cui alterata espressione fosse associata alle caratteristiche più aggressive del CRC. Per fare questo, abbiamo utilizzato un approccio combinato di biologia computazionale e sperimentale che ci ha consentito di individuare un lncRNA (LINC00483) costantemente *down-regolato* in tutti i *dataset* analizzati e nelle biopsie di CRC rispetto alla parte adiacente sana, la cui funzione sembra essere associata all'inibizione della proliferazione e della transizione epitelio-mesenchima.

6. Materiali e Metodi

6.1 Analisi Computazionale

Il lavoro è iniziato tramite la consultazione di “R2 Genomics” (<https://hgserver1.amc.nl/cgi-bin/r2/main.cgi?&species=hs>) per la ricerca dei *dataset* contenenti dati di espressione relativi al carcinoma del colon-retto (Tabella 3), al fine di identificare quei geni che fossero differenzialmente espressi in associazione alle caratteristiche più gravi del fenotipo tumorale [e.g., presenza di metastasi, instabilità dei microsatelliti, stadio TNM avanzato, mutazioni di *k-ras*, *CpG island methylator phenotype (CIMP) status*]. I risultati restituiti dall’analisi sono stati filtrati per *p-value* (≤ 0.01), impostando come *test* statistico l’ANOVA *test*, mentre la correzione per *test* multipli è stata effettuata tramite il “*False Discovery Rate*”.

DATASET SELEZIONATI
GSE33114 (108)
GSE2109 (315)
GSE21510 (148)
GSE14333 (290)
GSE41258(390)
GSE39582 (566)
GSE2109 (39)
TCGA COAD (286)

Tabella 3. Lista dei *dataset* selezionati per lo studio. Il numero dei campioni è indicato tra parentesi.

Da ogni *dataset* sono stati estrapolati i geni che fossero *down*-regolati e *up*-regolati nel CRC in relazione alle caratteristiche clinico-patologiche sopra menzionate.

Sono state incrociate le liste dei geni *down*-regolati e di quelli *up*-regolati, estrapolate da ciascun *dataset*, ottenendo una lista unica di geni che mostrassero lo stesso tipo di disregolazione (*i.e.*, sovra- o sotto-espressione) in almeno 3 dei *dataset* consultati.

I geni deregolati nelle condizioni prese in esame sono stati utilizzati come “*bait genes*” (geni esca) (BG) per individuare i lncRNA che fossero correlati positivamente o negativamente con tali geni, tramite calcolo del coefficiente di *Pearson* ($p\text{-value} < 0.01$) tra i valori di espressione dei geni deregolati e i valori di espressione di tutti i lncRNA presenti in tutti i *dataset* consultati. I lncRNA sono stati individuati utilizzando le liste degli *Official Gene Symbol* di *LNCipedia* (<https://lncipedia.org/>). Nel caso di geni rappresentati da più sonde, l'analisi è stata ripetuta in ogni *dataset* per ciascuna delle sonde rappresentanti il gene. Dei lncRNA restituiti dall'analisi di correlazione sono stati selezionati quelli correlati ad almeno metà delle sonde rappresentanti ciascun BG, generando così, per ogni *dataset*, un'unica lista di geni correlati positivamente ed un'unica lista di geni correlati negativamente a ciascuno dei BG. Sono state incrociate le liste dei lncRNA associati a ciascun BG per ciascun *dataset* e selezionati solo quei lncRNA presenti in almeno 3 diversi *dataset*.

Sono stati, quindi, incrociati i lncRNA correlati positivamente/negativamente con tutti i geni *down*-regolati e, allo stesso modo, i lncRNA correlati positivamente/negativamente

con tutti i geni *up*-regolati in associazione alle caratteristiche clinico-patologiche più gravi del CRC.

Infine, sono state generate 4 liste di lncRNA correlati ai BG associati alle caratteristiche clinico-patologiche più gravi del CRC:

1. lncRNA correlati positivamente ai BG *down*-regolati;
2. lncRNA correlati negativamente ai BG *down*-regolati;
3. lncRNA correlati positivamente ai BG *up*-regolati;
4. lncRNA correlati negativamente ai BG *up*-regolati;

Sono state sovrapposte le liste 1 e 4 e le liste 2 e 3, poiché ci si aspetta, in linea del tutto teorica, che i lncRNA correlati positivamente ai geni *down*-regolati fossero presenti anche nella lista dei lncRNA correlati negativamente ai geni *up*-regolati, e che i lncRNA correlati negativamente ai geni *down*-regolati fossero presenti anche nella lista dei lncRNA correlati positivamente ai geni *up*-regolati. Sono stati, quindi, selezionati i lncRNA comuni alle due liste confrontate.

6.1.1 Identificazione dei target dei lncRNA

Una delle funzioni più comuni e caratterizzate dei lncRNA è quella di “spugne molecolari” per i miRNA: essi legano e sequestrano i miRNA, impedendo a questi ultimi di degradare i loro mRNA *target* e inibirne l’espressione. Tra lncRNA, miRNA e mRNA si viene, dunque, a creare un circuito di interazione in cui alla *down*-regolazione del lncRNA corrisponde la *up*-regolazione del miRNA *target* (e

viceversa) e, di conseguenza, la *down*-regolazione degli mRNA *target* del miRNA, poiché vengono degradati da questi ultimi. In questa *network* di interazione a tre nodi le molecole alle estremità del circuito, *i.e.*, lncRNA e mRNA, mostrerebbero lo stesso *trend* di espressione, mentre la molecola che fa da *tres d'union* tra le due - il miRNA - presenterebbe espressione opposta (Figura 21).

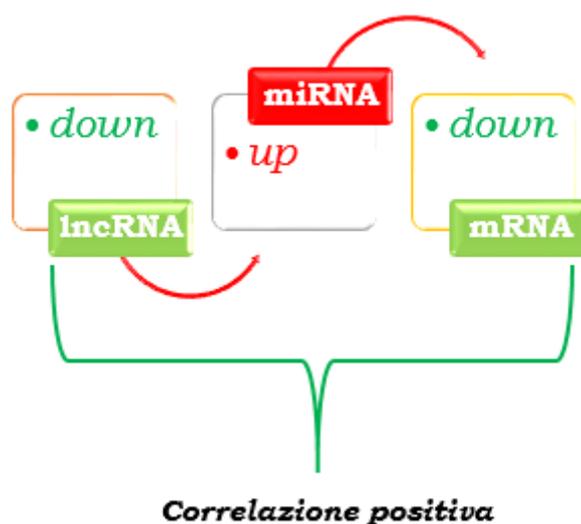


Figura 21. Rappresentazione schematica del circuito di interazione lncRNA-miRNA-mRNA.

Sulla base della suddetta funzione di “*miRNA sponges*” dei lncRNA, sono stati identificati quei miRNA che potessero legare contemporaneamente i lncRNA selezionati e i geni correlati positivamente con essi.

A tal fine, sono stati ricercati i miRNA correlati negativamente ai lncRNA in studio estrapolando - tramite UCSC XENA (<https://xenabrowser.net/datapages/>) - i valori di espressione dei miRNA dalla coorte TCGA Colon Cancer (COAD) (28 *dataset*) e - tramite “*R2 Genomics*” - i valori di espressione dei lncRNA dallo stesso

dataset. È stato, quindi, calcolato il coefficiente di correlazione di *Pearson* tra i valori di espressione dei miRNA e dei lncRNA, selezionando quei miRNA che fossero correlati negativamente ai lncRNA, e il cui valore di *p-value* associato fosse ≤ 0.001 .

Sono stati, quindi, ricercati i miRNA aventi siti di legame per i lncRNA in studio tramite i tool **LncBase** (http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=lncbasev2%2Findex-experimental) e **miRcode** (<http://www.mircode.org/>) e confrontati con i miRNA correlati negativamente ai lncRNA. I miRNA risultati dal confronto incrociato presentano, dunque, siti di legame per i lncRNA e sono negativamente correlati con essi. Questa analisi è stata rifinita tramite il tool RNA22 v.2 (<https://cm.jefferson.edu/rna22/Interactive/>). Più specificamente, sono stati selezionati i miRNA che presentavano almeno 2 siti di legame per i lncRNA ed un valore di *Folding Energy* (energia libera) ≤ -12 (tanto più negativo è il valore di *Folding Energy* tanto più forte è l'entità dell'interazione lncRNA-miRNA).

Sono stati identificati gli mRNA *target* di tali miRNA tramite **TarBase v.8** (http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=tarbasev8%2Findex). Per un miRNA (miR-544a) la ricerca dei *target* è stata implementata tramite consultazione del database **StarBase** (<http://starbase.sysu.edu.cn/>).

Per completare l'asse molecolare dell'interazione lncRNA-miRNA-mRNA sono stati consultati nuovamente i *dataset* selezionati nella fase iniziale dello studio al

fine di identificare quei geni che mostrassero la correlazione positiva più forte con i lncRNA in studio. Tali liste di geni (per ciascun *dataset*) sono state confrontate con le liste degli mRNA *target* dei miRNA correlati negativamente ai lncRNA e aventi siti di legame per questi ultimi. Sono state unificate le liste dei geni *target* di ciascun miRNA di tutti i *dataset* consultati, ottenendo così, per ciascun miRNA, un'unica lista di geni *target* legati dai miRNA associati ai lncRNA e, al tempo stesso, correlati positivamente ai lncRNA, in accordo con la supposta funzione di “*miRNA sponges*” dei lncRNA. Da questa lista finale sono stati selezionati quei geni *target* che presentavano i coefficienti di correlazione più elevati.

Sono stati consultati i *database mirCancer* (<http://mircancer.ecu.edu/>) e *mir2Disease* (<http://www.mir2disease.org/>) per verificare che tali miRNA associati sia ai lncRNA che ai geni positivamente correlati con questi ultimi fossero coinvolti nel CRC e quale fosse la loro espressione.

6.2 Linee cellulari

Per il presente studio sono state utilizzate le linee cellulari HCT-116, derivate da colture primarie del tumore al colon e mutate nel gene K-ras (resistenti al trattamento con Cetuximab). Le HCT-116 sono state ottenute dall'*Interlab Cell Line Collection (ICLC)*, una “*International Repository Authority*” all'interno dell'IRCCS Azienda Ospedaliera Universitaria San Martino-IST Istituto Nazionale per la Ricerca sul Cancro (Genova, Italia). Le HCT-116 sono state coltivate in *Roswell Park Memorial Institute-1640 (RPMI-1640)*

medium (Gibco), con aggiunta di siero bovino fetale (FBS) al 10% e di 2 mM L-glutammina (Lonza) e penicillina/streptomicina (10.000 U/ml) all'1%. Le cellule sono state mantenute a 37°C con una percentuale di CO₂ pari al 5%.

6.3 Estrazione di RNA da cellule

L'RNA totale delle cellule è stato estratto utilizzando il TriZol® (Invitrogen, Carlsbad, CA, USA) seguendo il protocollo fornito dal *kit*. L'RNA estratto è stato quantificato utilizzando lo spettrofotometro GenQuant pro (Biochrom).

Dopo la quantificazione, l'RNA è stato trattato con DNasi, seguendo il protocollo fornito dal Kit DNase I, Amp Grade (Invitrogen). Il trattamento si è reso necessario poiché alcuni geni studiati (RUNX1-IT1 e MCF2L-AS1) sono monoesonici.

6.4 Estrazione di RNA da tessuti FFPE

Per il presente studio, sono stati reclutati 20 pazienti CRC tramite l'Azienda Ospedaliero-Universitaria Policlinico Vittorio Emanuele (Catania, Italia). Sono stati, quindi, esaminati 20 campioni di tessuto tumorale, con le rispettive parti sane prese dallo stesso paziente. Le fettine di tessuto fissato in formalina ed incluso in paraffina (FFPE) sono state fornite dalla Sezione di Anatomia Patologica, Dipartimento G.F. Ingrassia, dell'Università di Catania (Catania, Italia). L'RNA è stato estratto dai tessuti FFPE tramite il *PureLink FFPE RNA Isolation Kit* (Ambion), secondo il protocollo fornito dal *kit*, e quantificato tramite lo spettrofotometro GenQuant pro (Biochrom) e Qubit (Invitrogen).

6.5 Estrazione di RNA da tessuti congelati inclusi in OCT e da biopsie

fresche

Sono stati esaminati 25 campioni di tessuto tumorale primitivo e 11 campioni di tessuto tumorale primitivo metastatico inclusi in OCT e 3 biopsie tumorali fresche. Tutti i campioni sono stati forniti dall'Istituto Oncologico del Mediterraneo (IOM) di Viagrande (Catania, Italia). Le fettine di tessuto tumorale incluse in OCT sono state tagliate tramite criostato e fornite dalla Sezione di Anatomia Patologica, Dipartimento G.F. Ingrassia, dell'Università di Catania (Catania, Italia). Prima di procedere all'estrazione dell'RNA, si è reso necessario sottoporre i campioni al seguente trattamento per eliminare i residui di OCT:

1. 3 lavaggi con 500 µl di DEPC H₂O; centrifugare a 4°C x 5 min a 11000 g (Eppendorf Centrifuge 5424 R) e rimuovere il surnatante dopo ogni *step*;
2. Un lavaggio con 500 µl di ETOH 25%; centrifugare a 4°C x 5 min a 11000 g (Eppendorf Centrifuge 5424 R) e rimuovere il surnatante.

L'RNA totale è stato, quindi, estratto dai tessuti inclusi in OCT tramite *PureLink FFPE RNA Isolation Kit (Ambion)*, secondo il protocollo fornito dal *kit*.

L'RNA totale è stato estratto dalle biopsie tumorali fresche utilizzando il *TriZol® (Invitrogen, Carlsbad, CA, USA)* seguendo il protocollo fornito dal *kit*.

In entrambi i casi, l'RNA estratto è stato quantificato utilizzando lo spettrofotometro *GenQuant pro (Biochrom)*.

6.6 Disegno dei primers per Real Time PCR

Per analizzare l'espressione dei lncRNA selezionati nello studio e dei loro mRNA *target* sono stati disegnati specifici *primers* tramite il *tool online* Primer-Blast (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>). È stata, inoltre, disegnata una coppia di *primers* specifica per il gene *housekeeping* PPIA (*peptidylprolyl isomerase A*). Onde evitare di amplificare eventuale DNA genomico contaminante, si è reso necessario disegnare coppie di *primers* che riconoscessero solo l'RNA trascritto da ogni gene di interesse. Pertanto, ove possibile, sono state scelte coppie di *primers* specifiche per amplimeri non monoesonici, cioè coppie di cui *forward* e *reverse* cadessero su esoni differenti e separati da introni di grandi dimensioni, o di cui almeno uno dei due *primers* cadesse a cavallo di due esoni confinanti. Nonostante ciò, c'è da annotare che due geni, RUNX1-IT1 e MCF2L-AS1, sono monoesonici.

Le coppie di *primers* utilizzate sono riportate nella Tabella 4.

GENE	FORWARD PRIMER	REVERSE PRIMER	PRODUCT LENGHT
ACVR1B	TGCAACAGGATCGACTTGAGG	CGGCGATGATGCCTACCAG	97
ADD3	ATGTGCAACTAGCCTGTGAGA	ACATGGAGATTGTCTACTCCACC	70
ARL4A	TATCCCTGCGTAGAAACGCC	TACAACTTCTCTGGTAGCTGATCC	97
CC2D1A	AGTTCITGGCTTTGGTCGGG	CAGGCTGGCCATCTTCTCAA	98
CD9	GCGGCCATCTGGGGATATT	GGTCATCCTTGGTTTTTCAGC	98
CTNNA1	TTATCGATGCTTCCCGCCTG	GAGTCATCCAACCTCCTCAGGG	91
DLEU1	CCAGCGCTTTACAGGGGTTA	TGACTCAATCCTGAAGGTTCTGT	100
DSP	AAAATCAAAGCCGACCTGCG	CCATCCTCTCAAAGGACGCT	91
EIF6	GCGGGATTCCCTCATIGACA	GCGGGAATGTGGAGAAGGTT	98
EPCAM	GTGGACATAGCTGATGTGGCT	CCCCATTTACTGTCAGGTCCA	94
GALNT3	TAGGAGCAACAGTCGCAACA	GCCAACAGAGGTTCTAGCCA	88
HIGD2A	AGAACCCGGTGGTACCCATA	TGGAAGGAGTAGAGGCCGTA	70
LINC00261	GCAGGAATCCCIGGATAAAGA	AGAGCCCGGCGGGTTA	85

LINC00483	ACCAGAACGAGTGGAGAAGATCA	CCACTTCCTGAAAGTCCCGC	92
LINC01207	TCCTAGCCACTCTCAGGGAC	CTCCATGAACCACCACCGAA	64
MCF2L-AS1	ATCACAGGGCCATATGCAAGA	AACCACAAAAGGATAGCTGCG	70
MEG3	CTCAGGCAGGATCTGGCATA	CCTGGAGTGCTGTGGAGAA	127
MIER3	ACACACACAAGGGACAATGAAC	GCCTTTCATTGCAGCAGTAT	98
PLS1	GGGAAGATGGACCTGCCATT	CATGAGTCCAGCACGCTTCA	77
PPIA	TCGTGCCGTTTTGCAGACGC	CAAACAGCTCAAAGGAGACGCGG	127
PRKAA1	ATGGCAGAAGTATGTAGAGCAATCA	AGGATTCTTCCTTCGTACACGC	90
RNF43	TGCCATCTGCTGGAGGAGTT	GATGTAACCAGGGGTCCACA	98
RUNX1-IT1	CCTGGGAAGTGGGAGATCTTTG	TCTGAGCTTTGAATCAGGAGTGATG	98
SH3YL1	GGAGGCAGCGGGATTGTAG	GCTATCCCAATGGCTGAGGG	71
SRPK1	ACCCCTGCTGACATTTGGAG	TCGAGTGTACTCTTCCCCTGA	96
TP73-AS1	ATTGTCCTCAACGACCAGCG	AACTTGCTTGIGCAAGAAGCTGG	60
TPD52	CCACAGAGACCTCTCGGAA	TTTGCTGCTAACACTTGAGACAG	100
USP7	TCTTTGTACAGGCGGATGCT	TTAAGCCGACGTAGCCTGTG	72
VDR	GCCACTGGCTTTCACCTCAAT	GGTGAATAGTGCCTTCCGCT	87
ZDHHC9	GGCGCTACCAGATGAAGCA	TTCTTGATACGAGGCGGTGG	96

Tabella 4. PCR *primers* per i lncRNA selezionati, gli mRNA *target* e per il gene *housekeeping* PPIA.

6.7 Analisi di espressione tramite Real-Time PCR

Abbiamo valutato l'espressione dei lncRNA selezionati e dei loro mRNA *target* nelle linee cellulari e nei tessuti di pazienti CRC tramite *Real-Time PCR* utilizzando il *kit* Power SYBR Green RNA-to-CT 1-Step (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA), seguendo il protocollo fornito dal *kit*. Tutte le reazioni sono state allestite nel termociclatore 7900HT Fast Real-Time PCR System (Applied Biosystems). I geni differenzialmente espressi (DE) sono stati identificati tramite il *software* SDS RQ Manager 1.2 (Applied Biosystems); i dati sono stati normalizzati utilizzando come "*reference gene*" il gene PPIA, sia per le linee cellulari che per i tessuti CRC. I *fold changes* dei geni DE sono

stati calcolati tramite il metodo $2^{-\Delta\Delta Ct}$. La significatività statistica dell'espressione differenziale dei geni è stata verificata tramite l'applicazione di un *T-test* accoppiato a una coda per comparare i valori di ΔCt dei campioni FFPE, mentre il *T-test* non accoppiato a una coda è stato utilizzato per comparare i valori di ΔCt dei campioni di tessuto incluso in OCT, delle biopsie fresche e dei campioni derivati dalle linee cellulari. Sono stati considerati statisticamente significativi i valori di *p-value* ≤ 0.05 .

6.8 Trattamenti delle linee cellulari

6.8.1 Trasfezione delle HCT-116 con LINC00483

È stata indotta la sovra-espressione *in vitro* di LINC00483 trasfettando le HCT-116 con il vettore di espressione pcDNA3.1(+) (Figura 22), in cui è stata clonata la sequenza del cDNA di LINC00483. La sequenza è stata inserita a valle del promotore CMV - nella stessa direzione del promotore - tramite l'enzima di restrizione EcoR V. La molecola di plasmide senza l'inserito è stata utilizzata come controllo negativo di trasfezione (vettore *scramble*). Entrambi i costrutti sono stati sintetizzati e acquistati presso la ditta GeneScript.

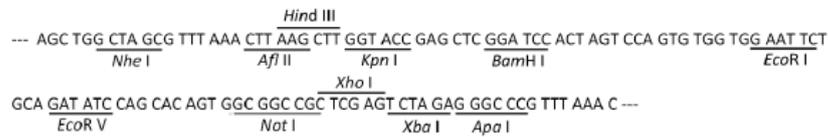
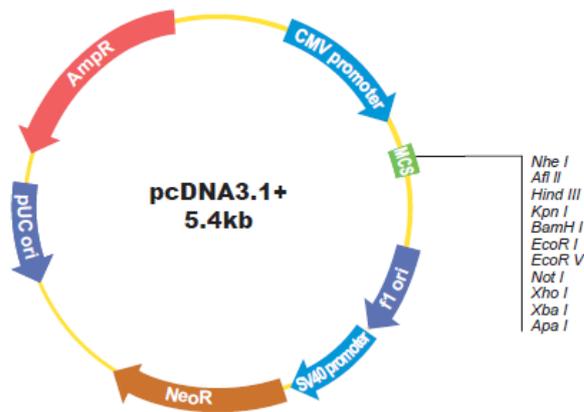


Figura 22. Rappresentazione schematica del vettore di espressione pcDNA3.1 (+).

10 µg di plasmide sono stati risospesi in 40 µl di TE buffer 1X. Per la trasfezione, 2 x 10⁵ cellule per pozzetto sono state seminate in *plates* da 24-well e trasfettate con 750 ng di plasmide insieme al reagente *Lipofectamine*TM 2000 (*Invitrogen*). È stato applicato il metodo di trasfezione diretto seguendo le indicazioni fornite dal protocollo. Dopo 48 h dalla trasfezione le cellule sono state lisate con *TriZol* per l'estrazione dell'RNA. L'RNA totale estratto è stato utilizzato sia per confermare l'efficienza di trasfezione (>99%) sia per analizzare, tramite *Real Time PCR*, l'espressione dei geni *target*, correlati positivamente a LINC00483. Tutti gli esperimenti sono stati eseguiti in triplicato biologico.

6.8.2 Saggio della vitalità cellulare

È stata valutata la vitalità/proliferazione cellulare a 48 h dalla trasfezione utilizzando il “Cell Counting kit-8” (CCK-8), secondo il protocollo fornito dal kit. I valori di assorbanza sono stati misurati tramite il Varioskan™ LUX (Thermo Scientific™). Per il saggio CCK-8 le HCT-116 sono state trasfettate con il vettore di espressione e il rispettivo vettore *scramble* seminando 2.5×10^4 cellule per pozzetto in *plates* da 96-well e scalando in maniera appropriata le quantità di reagenti di trasfezione. L'esperimento è stato eseguito in triplicato biologico.

6.8.3 Saggio della migrazione cellulare

È stato valutato il tasso di migrazione cellulare a seguito della trasfezione utilizzando il kit *Oris™ Universal Cell Migration Assembly* (Platypus Technologies), secondo le istruzioni fornite dal protocollo. In breve, le HCT-116 sono state trasfettate con il vettore di espressione e il rispettivo vettore *scramble* seminando 3.5×10^4 cellule per pozzetto in *plates* da 96-well - dopo l'inserzione degli *Oris Cell Seeding Stoppers* - e scalando in maniera appropriata le quantità di reagenti di trasfezione.

Per creare la zona di *detection* al centro di ciascun pozzetto gli *stoppers* sono stati rimossi dopo 24h (0 h). Le zone di *detection* di ciascun pozzetto sono state fotografate con il microscopio ottico *Leitz FLUOVERT* (Leica Microsystems) a 0 h (*premigration reference wells*) e a 24, 48, 72, 96 e 144 h dalla trasfezione. Le cellule

migrate all'interno della zona di *detection* sono state quantificate tramite l'utilizzo del *software* ImageJ. L'analisi ci ha permesso di quantificare le cellule migrate all'interno della zona di *detection* nel corso dei diversi *time points*. L'esperimento è stato eseguito in triplicato biologico.

6.8.4 Trattamento delle HCT-116 con TGF β -1

Per verificare il coinvolgimento dei lncRNA in studio nei processi della proliferazione cellulare e della formazione di metastasi, è stato indotto *in vitro* uno degli eventi cruciali della cascata metastatica: la transizione epitelio-mesenchima (EMT). Il TGF β -1 (*Transforming Growth Factor beta 1*) è una citochina secreta dalle cellule tumorali e dai fibroblasti dello stroma nel tessuto tumorale ed è considerato il principale induttore della EMT¹⁶⁷. Per il trattamento delle HCT-116 con TGF β -1, 5 x 10⁴ cellule per pozzetto sono state seminate in *plates* da 24-well e coltivate in condizioni di *serum starvation* (0.5 % FBS) per 24 h. Successivamente, le cellule sono state trattate con 20 ng/mL di TGF β -1 ed esposte all'azione della molecola per 24 h. I controlli sono stati mantenuti in solo RPMI-1640 *medium* (Gibco), senza l'aggiunta di FBS. Tutti gli esperimenti sono stati eseguiti in triplicato biologico. L'RNA totale estratto è stato utilizzato per valutare l'espressione dei geni tipicamente associati alla transizione epitelio-mesenchima (geni *markers* della EMT) e dei geni *target* dei lncRNA in studio. Specifiche coppie di *primers* sono state disegnate per valutare l'espressione dei geni *markers* tramite *Real-Time PCR* (Tabella 5). PPIA è stato utilizzato come "*reference gene*".

GENE	FORWARD PRIMER	REVERSE PRIMER	PRODUCT LENGTH
MMP7	AAGTGGTCACCTACAGGATCG	GGGATCTCTTTGCCCCACAT	103
VIM	CGAAAACACCCTGCAATCTT	CTGGATTTCCTCTTCGTGGA	133
ZEB1	GAGCACTTAAGAATTCACAGTGGAGAGA	ACTGAGATGTCTTGAGTCCTGTCTTG	163

Tabella 5. PCR *primers* per i geni *markers* della EMT.

6.8.5 *Trattamento delle HCT-116 con IL-6*

Per rafforzare e confermare i dati relativi al coinvolgimento dei lncRNA in studio nei processi tumorali associati alla EMT, le HCT-116 sono state trattate con IL-6 (*Inteleukin-6*), una citochina multifunzionale la cui via di segnalazione è iperattivata in molte malattie umane ed è stata associata con l'insorgenza e l'evoluzione del cancro¹⁶⁸. Per il trattamento delle HCT-116 con IL-6, 3×10^4 cellule per pozzetto sono state seminate in *plates* da *24-well* e coltivate in condizioni di *serum starvation* (no FBS) per 24 h. Successivamente, le cellule sono state trattate con 200 ng/mL di IL-6 ed esposte all'azione della molecola per 24 e 48 h. I campioni controllo sono stati mantenuti in RPMI-1640 *medium* (Gibco) (10% FBS). Tutti gli esperimenti sono stati eseguiti in triplicato biologico. L'RNA totale estratto è stato utilizzato per valutare l'espressione (tramite *Real Time PCR*) dei geni *markers* della transizione epitelio-mesenchima e dei geni *target* dei lncRNA in studio. PPIA è stato utilizzato come controllo endogeno.

6.8.6 Blocco del ciclo cellulare

Per verificare se l'espressione dei lncRNA in studio venga modulata durante fenomeni di blocco artificiale del ciclo cellulare, è stata valutata la risposta delle HCT-116 alla condizione di assenza di siero: la deprivazione di siero per 12 - 24 ore blocca il ciclo cellulare e porta le cellule in fase G₀. 5 x 10⁴ cellule sono state seminate in *plates* da 24-well. L'esperimento è stato strutturato in due fasi: un gruppo di campioni è stato sottoposto al trattamento tramite deprivazione di siero, mantenendo le cellule in RPMI-1640 *medium* (Gibco) (no FBS) per 24h. I controlli sono stati mantenuti in RPMI-1640 (10% FBS). Un secondo gruppo di cellule è stato sincronizzato tramite *serum starvation* (no FBS) per 24 h e successivamente incubato con terreno fresco senza siero (deprivazione di siero) per ulteriori 24 h. I controlli sono stati incubati con RPMI-1640 fresco con l'aggiunta del 10% di FBS. Tutti gli esperimenti sono stati eseguiti in triplicato biologico.

6.8.7 Trattamento delle HCT-116 con l'inibitore delle MAPK U0126

Per valutare il coinvolgimento dei lncRNA in studio nel *signalling* che sta alla base dell'iperproliferazione delle cellule di CRC, è stata effettuata sul nostro modello *in vitro* una inibizione farmacologica della via di segnalazione delle MAPK. Nelle HCT-116 la *pathway* delle MAPK è iperattivata, essendo queste cellule mutate nel gene *K-ras*. U0126 è un inibitore altamente selettivo delle chinasi MEK1 e MEK2, che

agisce inibendo la fosforilazione e l'attivazione *in vitro* di MEK1/MEK2 legando gli enzimi allo stato inattivo ed impedendo il reclutamento della chinasi ERK.

Per il trattamento, 3.2×10^4 cellule per pozzetto sono state seminate in *plates* da 24-well e coltivate in condizioni di *serum starvation* (no FBS) per 24 h (sincronizzazione); successivamente, le cellule sono state trattate con 25 mM U0126 (*MEK1/2 inhibitor*, Merck, Darmstadt, Germany). Le cellule sono state esposte al farmaco per 12 e 24 h. I controlli sono stati trattati con un equivalente volume di DMSO (il solvente del farmaco utilizzato per il trattamento). L'esperimento è stato eseguito in triplicato biologico. La vitalità cellulare è stata valutata tramite il saggio MTT [3-(4,5-dimethylthiazol-2-yl) -2,5-diphenyltetrazolium bromide] dopo 24 h dal trattamento. Il saggio valuta la funzionalità mitocondriale e, nello specifico, la capacità dell'enzima succinato deidrogenasi (attivo solo nelle cellule vitali) di ridurre l'MTT a formazano, conferendo alla sostanza un colore blu/violaceo. Tale reazione è valutata e misurata mediante la lettura spettrofotometrica del campione, alla lunghezza d'onda di 570 nm.

7. Risultati

7.1 Analisi Computazionale

Come precedentemente esposto, il presente lavoro è iniziato tramite la consultazione di “R2 Genomics” (<https://hgserver1.amc.nl/cgi-bin/r2/main.cgi?&species=hs>) per la ricerca dei *dataset* contenenti dati di espressione relativi al carcinoma del colon-retto, al fine di identificare quei geni che fossero differenzialmente espressi in associazione alle caratteristiche più gravi del fenotipo tumorale [e.g., presenza di metastasi, instabilità dei microsatelliti, stadio TNM avanzato, mutazioni di *k-ras*, *CpG island methylator phenotype (CIMP) status*].

Da ciascuno dei *dataset* selezionati sono stati estrapolati i geni che fossero *down*-regolati e *up*-regolati nel CRC in relazione alle suddette condizioni.

Sono state incrociate le liste dei geni *down*-regolati e di quelli *up*-regolati, estrapolate da ciascun *dataset*, ottenendo una lista unica di geni *up*-regolati (Tabella 6) e una lista unica di geni *down*-regolati (Tabella 7), la cui alterazione fosse comune in almeno 3 dei *dataset* consultati.

GENE
CYP1B1
NPR3
RGL2
SLIT2
TSPAN2

Tabella 6. Lista dei geni *up*-regolati nel CRC in relazione alle caratteristiche clinico-patologiche più gravi del fenotipo tumorale.

GENE
ACOT7
AGPAT5
ATP5B
AURKAIP1
CASP1
CEP55
CXCL3
FUT4
GSR
HNRNPAB
IDO1
KIF11
MCM5
PBK
PIGR
RANBP1
SCO2
TOE1
TLL12

Tabella 7. Lista dei geni *down-regolati* nel CRC in relazione alle caratteristiche clinico-patologiche più gravi del fenotipo tumorale.

Come dettagliatamente descritto nei Metodi, i geni deregolati nelle condizioni prese in esame sono stati utilizzati come “*bait genes*” (geni esca) (BG) per individuare i lncRNA che fossero correlati positivamente o negativamente con tali geni, tramite calcolo del coefficiente di *Pearson* (p -value <0.01) tra i valori di espressione dei geni deregolati e i valori di espressione di tutti i lncRNA presenti in tutti i *dataset* consultati.

Sono state incrociate le liste dei lncRNA associati a ciascun BG per ciascun *dataset* e selezionati solo quei lncRNA presenti in almeno 3 diversi *dataset*.

Sono stati, quindi, incrociati i lncRNA correlati positivamente/negativamente con tutti i geni *down*-regolati e, allo stesso modo, i lncRNA correlati positivamente/negativamente con tutti i geni *up*-regolati in associazione alle caratteristiche clinico-patologiche più gravi del CRC.

Infine, sono state generate 4 liste di lncRNA correlati ai BG associati alle manifestazioni più gravi del CRC (Tabelle 8 - 11).

1. lncRNA correlati positivamente ai BG *down*-regolati (Tabella 8);
2. lncRNA correlati negativamente ai BG *down*-regolati (Tabella 9);
3. lncRNA correlati positivamente ai BG *up*-regolati (Tabella 10);
4. lncRNA correlati negativamente ai BG *up*-regolati (Tabella 11);

lncRNA
DLEU1
EP300-AS1
HCP5
LINC00261
LINC00483
LINC00675
LINC01207
MCF2L-AS1
TFAP2A-AS1
THAP9-AS1
TTC28-AS1
USP30-AS1

Tabella 8. lncRNA correlati positivamente ai BG *down*-regolati.

lncRNA
A2M-AS1
ADAMTS9-AS2
ASAP1-IT1
CCDC144NL-AS1
CCDC18-AS1
DAPK1-IT1
DLEU2
EGOT
HYMAI
KLF3-AS1
LINC00312
LINC00622
LINC00623
LINC00663
LINC00702
LINC00865
LINC00869
LINC00893
LINC01138
LINC01279
LINC01410
LINC-PINT
MAGI2-AS3
MEG3
MIR99AHG
MLLT4-AS1
NR2F1-AS1
PSMA3-AS1
PSMD5-AS1
RUNX1-IT1
SERTAD4-AS1
STX17-AS1
TAPT1-AS1
TP53TG1
TP73-AS1
TUG1
ZNF561-AS1
ZNF667-AS1

Tabella 9. lncRNA correlati negativamente ai BG *down-regolati*.

lncRNA
LINC00312
MEG3
RUNX1-IT1
TP73-AS1

Tabella 10. lncRNA correlati positivamente ai BG *up*-regolati.

lncRNA
DLEU1
LINC00261
LINC00483
TMPO-AS1
LINC01207
MCF2L-AS1

Tabella 11. lncRNA correlati negativamente ai BG *up*-regolati.

Sono state sovrapposte le liste 1 e 4 e le liste 2 e 3, poiché ci si aspetta, in linea del tutto teorica, che i lncRNA correlati positivamente ai geni *down*-regolati fossero presenti anche nella lista dei lncRNA correlati negativamente ai geni *up*-regolati, e che i lncRNA correlati negativamente ai geni *down*-regolati fossero presenti anche nella lista dei lncRNA correlati positivamente ai geni *up*-regolati.

I lncRNA comuni alle due liste confrontate e selezionati per lo studio sono riportati nella Tabella 12.

lncRNA
DLEU1
LINC00261
LINC00483
LINC01207
MCF2L-AS1
MEG3
RUNX1-IT1
TP73-AS1

Tabella 12. lncRNA selezionati per lo studio.

7.2 Espressione dei lncRNA nelle biopsie CRC FFPE

L'espressione dei lncRNA selezionati è stata valutata, tramite *Real-Time PCR*, su 20 campioni di RNA da biopsie paraffinate di pazienti affetti da CRC, e sulle rispettive parti sane, prese dallo stesso paziente. I risultati hanno mostrato che 4 lncRNA (*i.e.*, MEG3, MCF2L-AS1, LINC00483, TP73-AS1) sono sotto-espressi in maniera statisticamente significativa nei tessuti tumorali contro le rispettive parti sane (Figura 23).

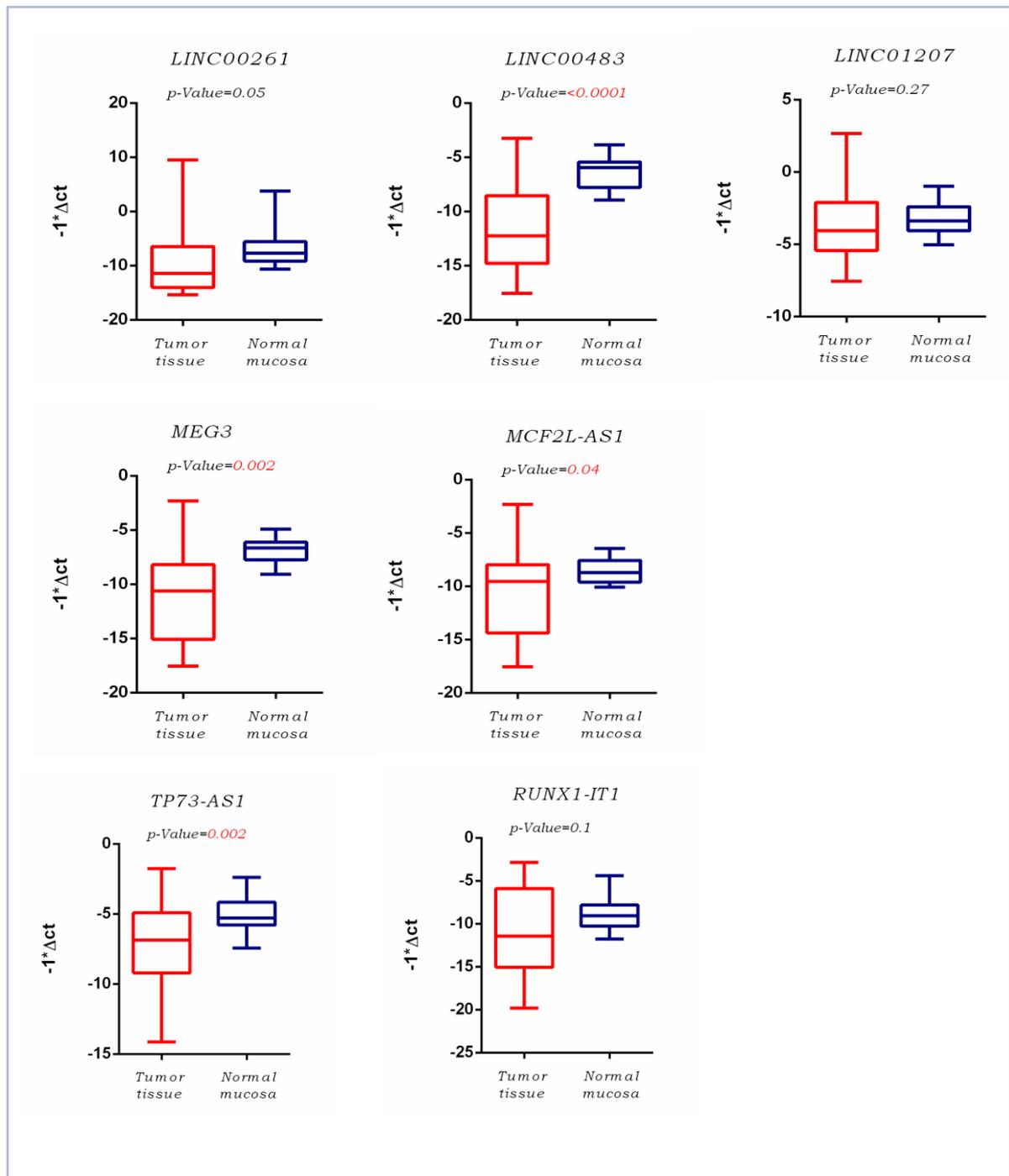


Figura 23. I box plot mostrano l'espressione differenziale dei lncRNA selezionati in 20 biopsie CRC FFPE contro le rispettive parti sane. I valori di p-value statisticamente significativi (<0.05) sono riportati in rosso. La coppia di primers disegnata per DLEU1 non ha dato alcun risultato, per cui è stata disegnata una seconda coppia di primers (dato non mostrato); anche quest'ultima, tuttavia, non produceva alcuna amplificazione. Per tale motivo i dati di DLEU1 non vengono mostrati.

7.3 Trattamenti delle linee cellulari

7.3.1 Il trattamento con $TGF\beta$ -1 induce la EMT *in vitro* nelle cellule HCT-116

L'analisi di espressione dei lncRNA selezionati per il nostro studio sulle biopsie CRC FFPE ci ha permesso di identificare 4 lncRNA differenzialmente espressi (DE lncRNA) (*i.e.*, MCF2L-AS1, MEG3, LINC00483, TP73-AS1) in maniera statisticamente significativa nei tessuti CRC rispetto alle parti adiacenti sane e sui quali abbiamo incentrato le successive analisi. Per verificare il coinvolgimento di tali lncRNA nei processi della proliferazione cellulare e della formazione di metastasi, abbiamo indotto *in vitro* uno degli eventi chiave della cascata metastatica: la transizione epitelio- mesenchima (EMT). La transizione epitelio-mesenchima è un processo fisiologico che si verifica durante i cambiamenti morfogenetici che coinvolgono lo sviluppo embrionale. In questo processo, le cellule epiteliali perdono le loro caratteristiche fenotipiche, quali la polarità apico-basale, i contatti cellula-cellula, ed acquisiscono un fenotipo mesenchimale, insieme ad una maggiore capacità invasiva e migratoria¹⁶⁹. Sebbene sia un processo fisiologico dello sviluppo, la EMT è stata osservata anche in alcune condizioni patologiche, tra cui la formazione di metastasi dei tumori¹⁵⁴. Il $TGF\beta$ -1 ha un ruolo chiave nella EMT. Si tratta di una citochina secreta dalle cellule tumorali e dai fibroblasti dello stroma nel tessuto tumorale ed è considerato il principale induttore *in vitro* della EMT¹⁷⁰. Abbiamo, quindi, valutato l'effetto del $TGF\beta$ -1 sulle linee cellulari HCT-116, e saggiato l'espressione dei DE lncRNA (tramite *Real Time PCR*) dopo 24 h dal trattamento con $TGF\beta$ -1.

Abbiamo, inoltre, valutato (tramite osservazione microscopia) la morfologia cellulare dopo 24 h dal trattamento. Come è possibile apprezzare dalle immagini mostrate in Figura 24, le cellule HCT-116 avevano cambiato morfologia, passando dalla loro tipica forma per lo più rotondeggiante ad una fusiforme, suggerendo che si stava verificando un processo di trasformazione cellulare.

TGF β -1 Treatment

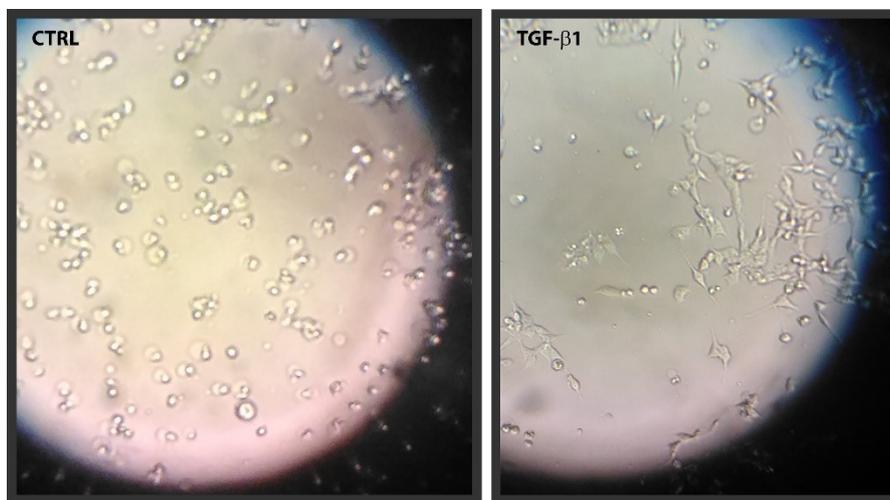


Figura 24. Morfologia cellulare dei controlli e dei trattati dopo 24 h dal trattamento con TGF β -1. Le cellule trattate con TGF β -1 presentano una morfologia più allungata e fusiforme rispetto ai controlli, in cui è possibile apprezzare la forma rotondeggiante che caratterizza le HCT-116.

Per verificare che il TGF β -1 avesse realmente indotto *in vitro* il processo della EMT abbiamo saggiato, tramite *Real Time PCR*, l'espressione di alcuni dei geni *markers* della EMT (*i.e.*, MMP7, VIM, ZEB1). I risultati hanno mostrato che l'espressione di MMP7, VIM e ZEB1 aumentava dopo 24 h dal trattamento (Tabella 13), confermandone l'efficacia.

GENE	FC MEDIANO 24 h
MMP7	2.28
VIM	2.53
ZEB1	1.45

Tabella 13. L'espressione dei geni *markers* della EMT (valutata tramite *Real Time PCR*) aumentava dopo 24 h dal trattamento con TGF β -1. Per ogni gene è mostrato il *fold change* mediano. Dati statisticamente significativi (*p-value* <0.05).

Abbiamo, quindi, valutato l'espressione dei DE lncRNA (tramite *Real Time PCR*) dopo 24 h dal trattamento con TGF β -1. I risultati hanno mostrato che il trattamento con TGF β -1 influenzava l'espressione dei DE lncRNA (Figura 25). Nello specifico, l'espressione di tre lncRNA (*i.e.*, MCF2L-AS1, MEG3, TP73-AS1) aumentava, mentre l'espressione di LINC00483 diminuiva dopo 24 h dal trattamento, anche se non in maniera statisticamente significativa.

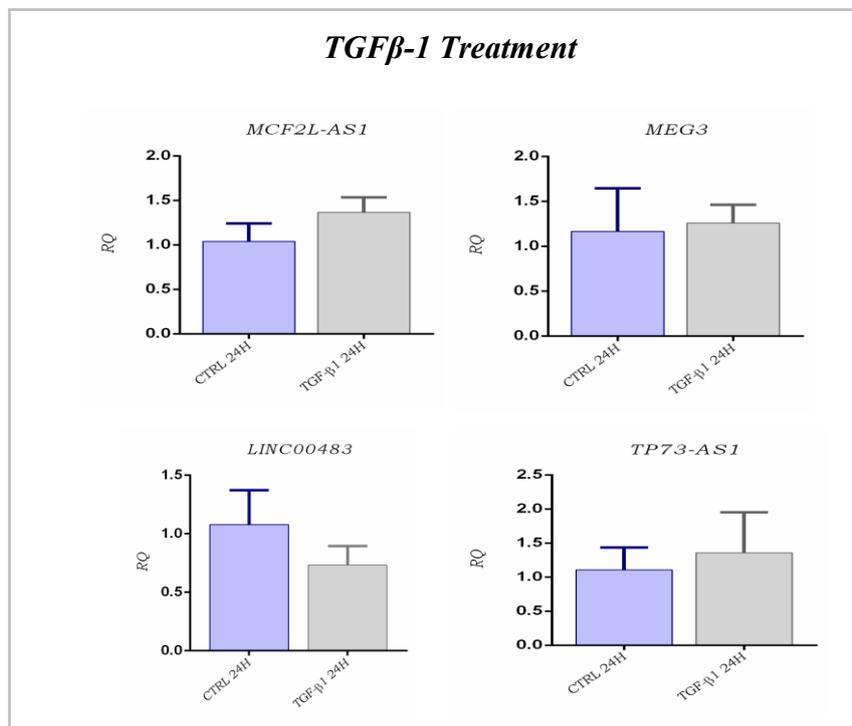


Figura 25. Il trattamento con TGF β -1 influenza l'espressione dei DE lncRNA. L'espressione di 3 lncRNA (MEG3, MCF2L-AS1, TP73-AS1) aumentava, mentre l'espressione di LINC00483 diminuiva dopo 24 h dal trattamento. I CTRL sono stati mantenuti in RPMI-1640, senza l'aggiunta di FBS.

7.3.2 Inibizione *in vitro* della pathway delle MAPK

7.3.2.1 L'inibizione delle MAPK influenza l'espressione dei DE lncRNA

Sulla base dei risultati ottenuti dal trattamento delle linee cellulari con $TGF\beta-1$, abbiamo deciso di incentrare le successive analisi su 3 dei DE lncRNA (*i.e.*, MCF2L-AS1, MEG3, LINC00483) selezionati nella prima fase dello studio. Per valutare il coinvolgimento dei lncRNA di nostro interesse nel *signalling* che sta alla base della iperproliferazione delle cellule di CRC, è stata effettuata sul nostro modello *in vitro* una inibizione farmacologica della via di segnalazione delle MAPK. L'espressione dei DE lncRNA è stata valutata sulle linee cellulari HCT-116 dopo 12 e 24 h dal trattamento con l'inibitore delle MAPK U0126. I risultati hanno mostrato che l'espressione dei DE lncRNA è influenzata dalla inibizione della *pathway* delle MAPK (Figura 26). Nello specifico, l'espressione di MCF2L-AS1, MEG3 e LINC00483 aumentava in maniera statisticamente significativa dopo 24 h dal trattamento.

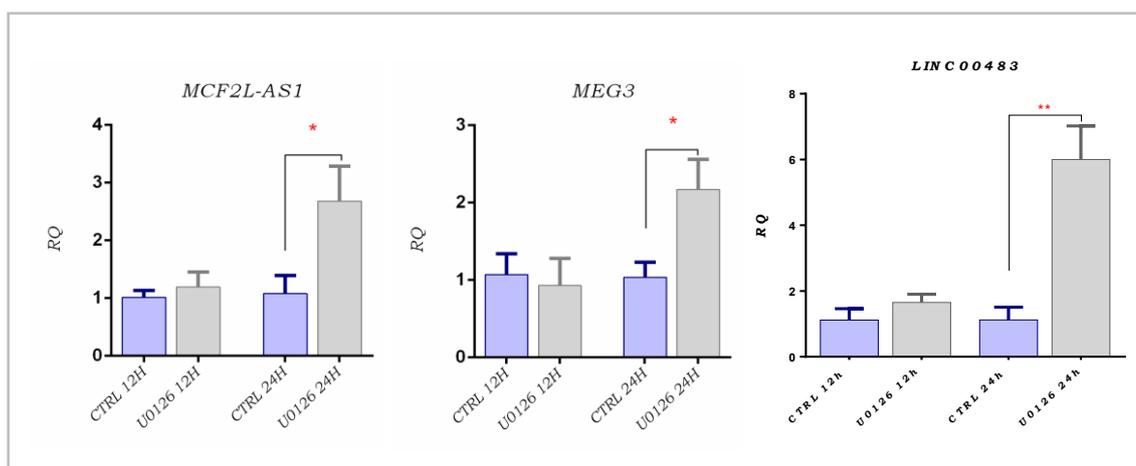


Figura 26. Gli istogrammi in figura mostrano l'espressione dei DE lncRNA dopo trattamento con U0126. Nello specifico, l'espressione di MCF2L-AS1, MEG3 e LINC00483 aumentava in maniera statisticamente significativa dopo 24h dal trattamento. *: $p\text{-value} < 0.05$; **: $p\text{-value} < 0.01$.

7.3.2.2 Saggio della vitalità cellulare

È stata valutata la vitalità cellulare dopo 24 h dal trattamento con l'inibitore delle MAPK U0126. I risultati hanno mostrato che il farmaco riduceva significativamente la proliferazione delle cellule trattate rispetto ai controlli negativi, trattati con un equivalente volume di DMSO (Figura 27), confermando l'efficacia del trattamento e il corrispondente fenotipo cellulare al quale era stata osservata un'alterazione dei lncRNA (nota: l'analisi di espressione dei lncRNA è stata condotta utilizzando un trattamento con 20 μ M di U0126).

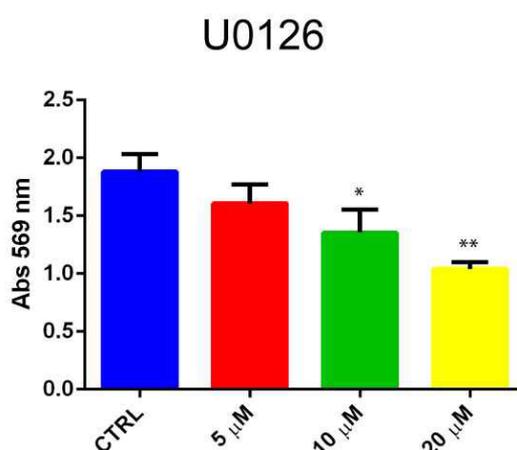


Figura 27: il trattamento con l'inibitore delle MAPK U0126 riduceva significativamente la vitalità cellulare, misurata tramite il saggio MTT; le cellule trattate con un equivalente volume di DMSO sono state usate come controllo.

*p-value < 0.05; **: p-value < 0.005.

7.4 Identificazione dei target di LINC00483

I trattamenti *in vitro* sulle linee cellulari HCT-116 hanno portato alla nostra attenzione un lncRNA i cui dati si sono mostrati particolarmente interessanti: LINC00483. Questo lncRNA risulta *down*-regolato in maniera quasi specifica nel CRC e la sua alterazione è

spesso associata a fattori prognostici favorevoli, come dimostrato dall'analisi computazionale che è stata eseguita su tutti i *dataset* presenti in *R2 Genomics* (1313 *dataset*), come mostrato in Tabella 14.

DATASET	CLASS	EXPRESSION	<i>P-value</i>
GSE3629 (121)	class: colorectal cancer vs nonneoplastic rectal mucosa	<i>down</i> in tumor	1.50E-19
GSE3629 (121)	group: sporadic ca vs uc-nonca	<i>down</i> in sporadic ca	1.50E-17
GSE21510 (148)	tissue: cancer-lcm vs normal-homogenized	<i>down</i> in cancer	5.80E-16
GSE8671 (64)	tissue: adenoma vs normal	<i>down</i> in adenoma	4.70E-10
GSE20916 (145)	tissue: adenocarcinoma vs normal_colon	<i>down</i> in adenocarcinoma	1.20E-09
GSE3629 (121)	group: sporadic ca vs uc-ca	<i>down</i> in sporadic ca	1.50E-09
GSE76124 (198)	tumor_grade: poorly_differentiated vs well_differentiated	<i>down</i> in poorly differentiated	5.70E-08
GSE102484 (683)	stage: 1 vs 4	<i>down</i> in 1	1.30E-07
GSE20916 (145)	malignancy: carcinoma vs distant_normal_colon	<i>down</i> in carcinoma	4.50E-07
GSE8671 (64)	nt_size: 0.2_cm_(x3) vs 2.0_cm	<i>down</i> in 2 cm	4.60E-07
GSE35896 (6 2)	braf_mutation: n vs y	<i>down</i> in y	5.30E-07
GSE8671 (64)	nt_size: 0.2_cm_(x3) vs 1.5_cm	<i>down</i> 1.5 cm	8.60E-07
GSE2109 (315)	nt_pathological_grade: 2 vs 3	<i>down</i> in 3	3.10E-06
GSE39582 (566)	cimp_status: neg vs pos	<i>down</i> in pos	3.30E-06
GSE35896 (6 2)	braf_mutation: n vs y	<i>down</i> in y	7.50E-06
GSE20916 (145)	tissue: carcinoma vs normal_colon	<i>down</i> in carcinoma	1.90E-05
GSE6791 (84)	sample_source: cervical normal vs head and neck cancer	<i>down</i> in cancer	1.90E-05
GSE6791 (84)	type: cervical normal vs head and neck cancer	<i>down</i> in cancer	1.90E-05
GSE4554 (84)	ms_status: msi vs mss	<i>down</i> in msi	2.00E-05

GSE8671 (64)	nt_location_tissue: sigmoid_colon_adenoma vs sigmoid_colon_normal	<i>down</i> in adenoma	2.60E-05
GSE76124 (198)	tumor_grade: moderately_differentiated vs well_differentiated	<i>down</i> in moderately differentiated	3.80E-05
GSE23878 (59)	group: colon tumour vs normal paired tissue	<i>down</i> in colon tumor	4.70E-05
GSE29271 (210)	survival_time_(months): 35 vs 9	<i>down</i> in 35	5.60E-05
GSE20916 (145)	malignacy: adenoma vs carcinoma	<i>down</i> in adenoma	6.10E-05
GSE13294 (155)	microsatellite_status: msi vs mss	<i>down</i> in msi	1.20E-04
GSE8671 (64)	nt_location_tissue: descending_colon_adenoma vs rectum_normal	<i>down</i> in adenoma	5.30E-04
GSE8671 (64)	nt_location_tissue: descending_colon_adenoma vs descending_colon_normal	<i>down</i> in adenoma	5.40E-04
GSE20916 (145)	tissue: colon_tumor vs norma_colon	<i>down</i> in tumor	6.10E-04
GSE3629 (121)	group: uc-associated ca vs uc-nonca	<i>down</i> in ca	6.90E-04
GSE41258 (390)	tissue: normal colon vs polyp	<i>down</i> in polyp	2.20E-03
GSE17538 (232)	ajcc_stage: 2 vs 4	<i>down</i> in 2	4.20E-03
GSE35896 (62)	pten_mutation: n vs y	<i>down</i> in y	4.20E-03
GSE29271 (210)	survival_time_(months): 36 vs 6	<i>down</i> in 36	4.60E-03
GSE42363 (14)	tumor_grade: moderately vs poorly	<i>down</i> in poorly differentiated	5.20E-03
GSE4183 (53)	group: colon_adenoma vs healthy_control	<i>down</i> in adenoma	5.20E-03
GSE35896 (62)	pten_mutation: n vs y	<i>down</i> in y	5.40E-03
GSE50948 (156)	invasive_tumor_area_size: 15 vs 6	<i>down</i> in 15	5.80E-03
GSE2109 (351)	nt_pathological_grade: 1 vs 3	<i>down</i> in 3	7.50E-03
GSE33114 (108)	sample_descr: normal colon mucosa vs primary tumor resection	<i>down</i> in tumor	8.30E-03
GSE33114 (108)	sample_source: normal colon mucosa vs primary tumor resection	<i>down</i> in tumor	8.30E-03
GSE33114 (108)	type: normal vs tumor	<i>down</i> in tumor	8.30E-03
GSE21510 (148)	tissue: cancer-homogenized vs normal-homogenized	<i>down</i> in cancer	8.30E-03
GSE4107 (22)	tissue: mucosa_control vs mucosa_patient	<i>down</i> in patients	8.60E-03

GSE9891 (285)	stagecode: ia vs iii	<i>down</i> in ia	9.50E-03
GSE35896 (62)	kras_mutation: n vs y	<i>down</i> in no	9.60E-03
GSE17538 (232)	ajcc_stage: 1 vs 2	<i>down</i> in 2	9.70E-03

Tabella 14. La tabella mostra i dati di espressione di LINC00483 [estrapolati dai *dataset* tramite R2 Genomics (<https://hgserver1.amc.nl/cgi-bin/r2/main.cgi?&species=hs>)], in diversi tipi di tumori umani. Per ogni *dataset* è riportato il GEO ID (il numero di campioni è indicato tra parentesi), le categorie a confronto, la relativa espressione di LINC00483 e il valore di *p-value* associato.

Secondo i nostri dati, LINC00483 risultava *down*-regolato nelle biopsie tumorali FFPE contro le rispettive parti adiacenti sane, e la sua espressione diminuiva dopo trattamento con *TGFβ-1* (che induce la capacità invasiva e migratoria delle cellule tumorali), mentre aumentava quando venivano spenti i segnali proliferativi indotti dalla via delle MAPK (trattamento con U0126). Questi risultati suggeriscono un potenziale ruolo di oncosoppressore per LINC00483. Abbiamo, quindi, focalizzato il nostro lavoro sulla comprensione dei meccanismi molecolari alla base della funzione di questo lncRNA. A tal fine, abbiamo ricercato i miRNA che potessero legare contemporaneamente il LINC00483 e i geni correlati positivamente con esso. Abbiamo, infatti, ipotizzato un potenziale ruolo di “*miRNA sponge*” per LINC00483, in quanto questa è nota essere la funzione più comune e caratterizzata dei lncRNA. Secondo questo modello, i lncRNA legano e sequestrano i miRNA, impedendo a questi ultimi di degradare i loro mRNA *target* e inibirne l’espressione. Tra lncRNA, miRNA e mRNA si viene, dunque, a creare un circuito di interazione in cui alla *down*-regolazione del lncRNA corrisponde la *up*-regolazione del miRNA *target* (e viceversa) e, di conseguenza, la *down*-regolazione degli mRNA *target* del miRNA, poiché vengono degradati da questi ultimi. Sono stati, quindi, ricercati i miRNA

correlati negativamente a LINC00483 estrapolando - tramite UCSC XENA (<https://xenabrowser.net/datapages/>) - i valori di espressione dei miRNA dalla coorte TCGA Colon Cancer (COAD) (28 *datasets*) e - tramite "R2 Genomics" - i valori di espressione di LINC00483, dallo stesso *dataset*. È stato, quindi, calcolato il coefficiente di correlazione di *Pearson* tra i valori di espressione dei miRNA e di LINC00483, selezionando quei miRNA che fossero correlati negativamente a LINC00483, e il cui valore di *p-value* associato fosse ≤ 0.001 (Tabella 15).

miRNA	R-value	P-value
hsa-let-7e-5p	-0.206	0.001
hsa-miR-205-5p	-0.261	< 0.0001
hsa-miR-181a-3p	-0.221	0.000
hsa-miR-34c-5p	-0.226	0.000
hsa-miR-302c-3p	-0.223	0.000
hsa-miR-511-5p	-0.209	0.001
hsa-miR-544a	-0.281	< 0.0001
hsa-miR-615-3p	-0.211	0.001
hsa-miR-218-2-3p	-0.219	0.000
hsa-miR-628-5p	-0.217	0.001
hsa-miR-944	-0.244	< 0.0001
hsa-miR-1231	-0.224	0.000
hsa-miR-1911-3p	-0.235	0.000
hsa-miR-205-3p	-0.223	0.000
hsa-miR-3125	-0.309	< 0.0001
hsa-miR-3190-5p	-0.202	0.001
hsa-miR-4464	-0.216	0.001
hsa-miR-4524a-5p	-0.279	< 0.0001
hsa-miR-4520-5p	-0.249	< 0.0001
hsa-miR-4658	-0.217	0.001
hsa-miR-4777-5p	-0.217	0.001
hsa-miR-4778-3p	-0.185	0.003
hsa-miR-5586-5p	-0.223	0.000
hsa-miR-5683	-0.205	0.001
hsa-miR-1185-2-3p	-0.221	0.000
hsa-miR-6502-3p	-0.249	< 0.0001
hsa-miR-6765-3p	-0.244	< 0.0001

hsa-miR-6831-3p	-0.249	< 0.0001
hsa-miR-7702	-0.308	< 0.0001
hsa-miR-7978	-0.206	0.001

Tabella 15. miRNA correlati negativamente a LINC00483.
Per ogni miRNA è mostrato il valore di *r-value* e il *p-value* associato (≤ 0.001).

Sono stati, quindi, ricercati i miRNA aventi siti di legame per LINC00483 tramite i *tool* *LncBase* e *miRcode* e confrontati con i miRNA correlati negativamente a LINC00483.

Gli *output* ottenuti da *LncBase* e *miRcode* e dall'analisi di correlazione sono stati incrociati ottenendo così una lista univoca di miRNA predetti da entrambi i *tool*. Tali miRNA presentano, dunque, siti di legame per LINC00483 e sono, al tempo stesso, negativamente correlati con esso (Tabella 16).

MiRNA
hsa-miR-302c-3p
hsa-miR-30a-5p
hsa-miR-3619-5p
hsa-miR-4464
hsa-miR-4660
hsa-miR-4743-3p
hsa-miR-511-5p
hsa-miR-7978

Tabella 16. miRNA ottenuti dal confronto incrociato tra gli *output* di *LncBase* e *miRcode* e gli *output* dell'analisi di correlazione. Tali miRNA presentano siti di legame per LINC00483 e sono, al tempo stesso, negativamente correlati con esso.

Questa analisi è stata rifinita tramite il *tool* RNA22 v.2. Più specificamente, sono stati selezionati i miRNA che presentavano almeno 2 siti di legame per LINC00483 ed un valore di *Folding Energy* (energia libera) ≤ -12 (tanto più negativo è il valore di *Folding Energy*

tanto più forte è l'entità dell'interazione lncRNA-miRNA) (Tabella 17). Il miRNA miR-544a presenta un solo sito di legame per LINC00483, tuttavia è stato selezionato poiché ha valore di *folding energy* molto negativo ed un valore di *p-value* notevolmente significativo.

miRNA	N. Binding Sites	Folding Energy	p-value
hsa-miR-205-3p	4	-12.1	3.74E-02
		-12.8	1.42E-03
		-12.8	1.42E-03
		-13.1	1.32E-01
hsa-miR-511-5p	2	-12.5	1.32E-01
		-15.4	8.50E-03
hsa-miR-330-5p	2	-12.8	3.53E-03
		-18.4	1.32E-01
hsa-miR-1231	2	-15.4	2.88E-02
		-16.1	1.42E-03
hsa-miR-544a	1	-16.2	1.42E-03

Tabella 17. miRNA selezionati tramite il tool RNA22 v.2 (<https://cm.jefferson.edu/rna22/Interactive/>). Per ogni miRNA è riportato il numero di siti di legame per LINC00483, il valore di *folding energy* (< -12) e il *p-value* associato.

Per tutti i miRNA così identificati abbiamo ricercato gli mRNA *target* tramite *TarBase v.8*. Per il miRNA miR-544a la ricerca dei *target* è stata implementata tramite consultazione del database *StarBase*. La ricerca dei *target* dei miRNA - miR-4464, miR-4660, miR-4743-3p, miR-511-5p, miR-7978 - non ha prodotto alcun risultato.

Per completare l'asse molecolare dell'interazione lncRNA-miRNA-mRNA sono stati consultati nuovamente i *dataset* selezionati nella fase iniziale dello studio al fine di identificare quei geni che mostrassero la correlazione positiva più forte con LINC00483. Tali liste di geni (per ciascun *dataset*) sono state confrontate con le liste degli mRNA *target* dei miRNA correlati negativamente ai lncRNA e aventi siti di legame per questi ultimi. Sono state unificate le liste dei geni *target* di ciascun miRNA di tutti i *dataset* consultati, ottenendo così, per ciascun miRNA, un'unica lista di geni *target* legati dai miRNA associati

a LINC00483 e, al tempo stesso, correlati positivamente a LINC00483 in accordo con la supposta funzione di “*miRNA sponge*” (Tabelle 18 -25).

Sono stati consultati i *database mirCancer* (<http://mircancer.ecu.edu/>) e *mir2Disease* (<http://www.mir2disease.org/>) per verificare che tali miRNA associati sia a LINC00483 che ai geni positivamente correlati con quest’ultimo fossero disregolati nel CRC.

Le tabelle di seguito (Tabelle 18-25) riportano i miRNA aventi siti di legame per LINC00483 e correlati negativamente con esso. Per ogni miRNA sono mostrati gli mRNA *target*, allo stesso tempo positivamente correlati a LINC00483.

MiRNA	miRNA/ LINC00483 correlation		miRNA expression (mirCancer - mir2Disease)	mRNA target	LINC00483/target correlation		Target expression	miRNA/target correlation (panCancer)	
	<i>R-value</i>	<i>p-value</i>			<i>R-value</i>	<i>p-value</i>		<i>R-value</i>	<i>p-value</i>
miR-30a-5p	-0.152	0.015	Down- regulated	CMAS	0.441	1.36E-14	Oncogene	-0.165	4.25E-04
				CNNM4	0.512	6.331E-25	Tumor suppressor	-0.071	1.32E-01
				DSP	0.406	2.684E-12	Tumor suppressor	-0.11	2.00E-02
				EPB41L4B	0.591	2.083E-28	Oncogene	-0.001	9.88E-01
				FBXO34	0.419	4.181E-13		-0.094	4.58E-02
				PLS1	0.569	7.92E-26	Oncogene	-0.185	7.57E-05
				RPS6KA3	0.469	1.067E-16	Tumor suppressor	-0.09	5.67E-02
				SPATA2	0.431	6.986E-14	Oncogene	0.035	4.56E-01
				TPD52	0.428	1.015E-13	Oncogene in CRC	-0.271	5.39E-09

Tabella 18. La tabella mostra i valori di correlazione tra il miR-30a-5p e il LINC00483, tra il LINC00483 e gli mRNA *target* del miR-30a-5p (allo stesso tempo positivamente correlati a LINC00483), e tra il miR-30a-5p e ciascun mRNA *target*, con i rispettivi valori di *p-value* associati. Viene, anche, riportata l’espressione del miR-30a-5p nel CRC (mirCancer, mir2Disease) e l’espressione dei geni *target*.

miRNA	miRNA/ LINC00483 correlation		miRNA expression (mirCancer - mir2Disease)	mRNA target	LINC00483/ target correlation		Target expression	miRNA/target correlation (panCancer)	
	<i>R-value</i>	<i>p-value</i>			<i>R-value</i>	<i>p-value</i>		<i>R-value</i>	<i>p-value</i>
miR-205-3p	-0.222	0.0003	Down-regulated	PRKAA1	0.207	2.27E-04	Oncogene	-0.098	3.81E-02
				RTCB	0.174	0.002		0.003	9.56E-01
				TMEM59	0.179	0.009	Oncogene	-0.083	8.04E-02

Tabella 19. La tabella mostra i valori di correlazione tra il miR-205-3p e il LINC00483, tra il LINC00483 e gli mRNA *target* del miR-205-3p (allo stesso tempo positivamente correlati a LINC00483), e tra il miR-205-3p e ciascun mRNA *target*, con i rispettivi valori di *p-value* associati. Viene, anche, riportata l'espressione del miR-205-3p nel CRC (mirCancer, mir2Disease) e l'espressione dei geni *target*.

miRNA	miRNA/ LINC00483 correlation		miRNA expression (mirCancer - mir2Disease)	mRNA target	LINC00483/ target correlation		Target expression	miRNA/target correlation (panCancer)	
	<i>R-value</i>	<i>p-value</i>			<i>R-value</i>	<i>p-value</i>		<i>R-value</i>	<i>p-value</i>
miR-302c-3p	-0.222	0.0003	Down-regulated	ARHGEF5	0.45	3.05E-15	Oncogene in CRC	-0.061	1.96E-01
				ARL4A	0.471	7.746E-17	Oncogene in CRC	-0.087	6.47E-02
				DCTPP1	0.452	2.184E-15	Oncogene	-0.054	2.49E-01
				FZD5	0.629	3.987E-33	Oncogene	0.002	9.72E-01
				VDR	0.559	1.612E-30	Tumor suppressor in CRC	-0.072	1.25E-01
				ADD3	0.437	1.15E-17	Tumor suppressor	-0.085	7.21E-02

Tabella 20. La tabella mostra i valori di correlazione tra il miR-302c-3p e il LINC00483, tra il LINC00483 e gli mRNA *target* del miR-302c-3p (allo stesso tempo positivamente correlati a LINC00483), e tra il miR-302c-3p e ciascun mRNA *target*, con i rispettivi valori di *p-value* associati. Viene, anche, riportata l'espressione del miR-302c-3p nel CRC (mirCancer, mir2Disease) e l'espressione dei geni *target*.

miRNA	miRNA/ LINC00483 correlation		miRNA expression (mirCancer - mir2Disease)	mRNA target	LINC00483/ target correlation		Target expression	miRNA/target correlation (panCancer)	
	<i>R-value</i>	<i>p-value</i>			<i>R-value</i>	<i>p-value</i>		<i>R-value</i>	<i>p-value</i>
miR-330-5p	-0.179	0.004	Down-regulated	ABCC3	0.453	1.99E-15	Oncogene	-0.139	3.21E-03
				ACVR1B	0.442	1.14E-4	Tumor suppressor in CRC	-0.106	2.47E-02
				APPL2	0.404	6.33E-15	Tumor suppressor in CRC	0.07	1.36E-01
				CC2D1A	0.32	1.02E-07		-0.206	1.09E-05
				CEBPA	0.462	4.77E-15	Oncogene in CRC	-0.34	1.21E-13
				FUT4	0.486	5.15E-18	Oncogene in CRC	0.061	1.96E-01
				RNF43	0.451	2.56E-15	Oncogene in CRC	-0.238	3.37E-07
				SMAP1	0.369	1.99E-12	Oncogene in CRC	-0.141	2.68E-03
				HIGD2A	0.363	5.07E-12	Tumor suppressor	-0.158	7.55E-04
				MGLL	0.321	8.88E-08	Tumor suppressor in CRC	-0.073	1.23E-01
				TSPAN3	0.449	1.04E-18	Oncogene	0.109	2.05E-02
ZDHHC9	0.389	3.10E-11		-0.102	2.98E-02				

Tabella 21. La tabella mostra i valori di correlazione tra il miR-330-5p e il LINC00483, tra il LINC00483 e gli mRNA *target* del miR-330-5p (allo stesso tempo positivamente correlati a LINC00483), e tra il miR-330-5p e ciascun mRNA *target*, con i rispettivi valori di *p-value* associati. Viene anche riportata l'espressione del miR-330-5p nel CRC (mirCancer, mir2Disease) e l'espressione dei geni *target*.

miRNA	miRNA/ LINC00483 correlation		miRNA expression (mirCancer - mir2Disease)	mRNA target	LINC00483/ target correlation		Target expression	miRNA/target correlation (panCancer)	
	<i>R-value</i>	<i>p-value</i>			<i>R-value</i>	<i>p-value</i>		<i>R-value</i>	<i>p-value</i>
miR-511-5p	-0.209	0.001	Down-regulated	USP7	0.372	2.78E-10	Oncogene in CRC	-0.122	9.56E-03
				SH3YL1	0.349	3.843E-11	Oncogene	-0.084	6.27E-01

Tabella 22. La tabella mostra i valori di correlazione tra il miR-511-5p e il LINC00483, tra il LINC00483 e gli mRNA *target* del miR-511-5p (allo stesso tempo positivamente correlati a LINC00483), e tra il miR-511-5p e ciascun mRNA *target*, con i rispettivi valori di *p-value* associati. Viene anche riportata l'espressione del miR-511-5p nel CRC (mirCancer, mir2Disease) e l'espressione dei geni *target*.

miRNA	miRNA/ LINC00483 correlation		miRNA expression (mirCancer - mir2Disease)	mRNA target	LINC00483/ target correlation		Target expression	miRNA/target correlation (panCancer)	
	<i>R-value</i>	<i>p-value</i>			<i>R-value</i>	<i>p-value</i>		<i>R-value</i>	<i>p-value</i>
miR-544a	-0.280	< 0.0001	Up-regulated	BCL11B	0.361	1.72E-08		0.029	5.46E-01
				CD9	0.391	1.89E-10	Oncogene in CRC	0.014	7.73E-01
				CEBPG	0.554	2.47E-24		-0.018	6.99E-01
				CMAS	0.441	1.35E-14	Oncogene	-0.024	6.06E-01
				CXCL14	0.387	9.42E-10	Tumor suppressor	0.037	4.27E-01
				DSP	0.406	2.68E-12	Tumor suppressor	-0.007	8.80E-01
				EPCAM	0.479	1.92E-17	Oncogene	-0.019	6.95E-01
				RAPGEFL1	0.569	7.23E-32		-0.019	6.81E-01
				SPATA2	0.431	6.97E-14	Oncogene	-0.027	5.66E-01
				ST14	0.488	3.21E-18	Oncogene	-0.076	1.09E-01
				TMEM106C	0.438	2.18E-14	Tumor suppressor	-0.07	1.39E-01
				TOX3	0.452	2.26E-15	Tumor suppressor	-0.097	4.05E-02
				AXIN2	0.4	6.24E-12	Tumor suppressor in CRC	-0.009	8.49E-01
				CHMP4B	0.394	1.48E-11		-0.065	1.70E-01
				FAM171A1	0.451	2.47E-15	Oncogene	0.01	8.30E-01
				GALNT3	0.422	2.66E-13	Oncogene in CRC	0.041	3.84E-01
				KIAA1161	0.505	1.02E-19		-0.031	5.07E-01
				LRRC1	0.478	2.38E-17	Oncogene	0.039	4.04E-01
				MIER3	0.376	1.53E-10	Tumor suppressor in CRC	0.015	7.46E-01
NRARP	0.553	3.65E-24	Oncogene	0.046	3.31E-01				
SLC35A3	0.465	2.37E-16		-0.019	6.93E-01				

Tabella 23. La tabella mostra i valori di correlazione tra il miR-544a e il LINC00483, tra il LINC00483 e gli mRNA *target* del miR-544a (allo stesso tempo positivamente correlati a LINC00483), e tra il miR-544a e ciascun mRNA *target*, con i rispettivi valori di *p-value* associati. Viene riportata anche l'espressione del miR-544a nel CRC (mirCancer, mir2Disease) e l'espressione dei geni *target*.

miRNA	miRNA/ LINC00483 correlation		miRNA expression (mirCancer - mir2Disease)	mRNA target	LINC00483/ target correlation		Target expression	miRNA/target correlation (panCancer)	
	<i>R-value</i>	<i>p-value</i>			<i>R-value</i>	<i>p-value</i>		<i>R-value</i>	<i>p-value</i>
miR-1231	-0.224	0.0003		CTNNA1	0.321	2.00E-09	Tumor suppressor in CRC	-0.238	3.21E-07
				EIF6	0.44	1.6849E-14	Oncogene in CRC	-0.045	3.45E-01
				FAM120A	0.333	2.5905E-08	Oncogene in CRC	0.008	8.60E-01
				RNF103	0.245	8.1358E-05		-0.074	1.18E-01

Tabella 24. La tabella mostra i valori di correlazione tra il miR-1231 e il LINC00483, tra il LINC00483 e gli mRNA *target* del miR-1231 (allo stesso tempo positivamente correlati a LINC00483), e tra il miR-1231 e ciascun mRNA *target*, con i rispettivi valori di *p-value* associati. Viene anche riportata l'espressione del miR-1231 nel CRC (mirCancer, mir2Disease) e l'espressione dei geni *target*.

miRNA	miRNA/ LINC00483 correlation		miRNA expression (mirCancer - mir2Disease)	mRNA target	LINC00483/ target correlation		Target expression	miRNA/target correlation (panCancer)	
	<i>R-value</i>	<i>p-value</i>			<i>R-value</i>	<i>p-value</i>		<i>R-value</i>	<i>p-value</i>
miR-3619-5p	-0.189	0.003		ACVR1B	0.442	1.1E-14	Tumor suppressor in CRC	-0.133	4.64E-03
				JUP	0.43	7.6E-13	Tumor suppressor	-0.09	5.73E-02
				SRPK1	0.38	1.0E-10	Oncogene	-0.063	1.81E-01
				ADD3	0.437	1.1E-17	Tumor suppressor	-0.02	6.76E-01

Tabella 25. La tabella mostra i valori di correlazione tra il miR-3619-5p e LINC00483, tra LINC00483 e gli mRNA *target* del miR-3619-5p (allo stesso tempo positivamente correlati a LINC00483), e tra il miR-3619-5p e ciascun mRNA *target*, con i rispettivi valori di *p-value* associati. Viene anche riportata l'espressione del miR-3619-5p nel CRC (mirCancer e mir2Disease) e l'espressione dei geni *target*.

Da queste liste di geni sono stati selezionati gli mRNA *target* che presentavano i coefficienti di correlazione con LINC00483 più elevati (Tabella 26).

LINC00483 Target	<i>R-value</i>	<i>p-value</i>
ACVR1B	0.442	1.14E-04
ADD3	0.437	1.10E-17
ARL4A	0.471	7.75E-17
CC2D1A	0.32	1.02E-07
CD9	0.391	1.89E-10
CTNNA1	0.321	2.00E-09
DSP	0.406	2.68E-12
EIF6	0.44	1.68E-14
EPCAM	0.479	1.92E-17
GALNT3	0.422	2.66E-13
HIGD2A	0.363	5.07E-12
MIER3	0.376	1.53E-10
PLS1	0.569	7.92E-26
PRKAA1	0.207	2.27E-04
RNF43	0.451	2.56E-15
SH3YL1	0.349	3.84E-11
SRPK1	0.38	1.00E-10
TPD52	0.428	1.02E-13
USP7	0.372	2.78E-10
VDR	0.559	1.61E-30
ZDHHC9	0.389	3.10E-11

Tabella 26. Lista finale dei geni positivamente correlati a LINC00483 selezionati per lo studio. Per ciascun gene *target* è mostrato il valore del coefficiente di correlazione di Pearson (*r-value*) e il valore di *p-value* associato.

7.5 Analisi di espressione di LINC00483 e dei suoi geni target

7.5.1 Analisi di espressione di LINC00483 sui tessuti CRC inclusi in OCT e sulle biopsie fresche

L'analisi di espressione di LINC00483 effettuata tramite *Real-Time PCR* su 20 campioni di RNA da biopsie paraffinate di pazienti CRC e sulle rispettive parti

sane, prese dallo stesso paziente, ha mostrato che il LINC00483 è sotto-espresso in maniera statisticamente significativa nei tessuti tumorali contro le rispettive parti adiacenti sane. Per rafforzare questo dato e valutare l'espressione di LINC00483 anche in relazione alle caratteristiche clinico-patologiche più gravi del CRC, quali la formazione di metastasi, abbiamo analizzato l'espressione di questo lncRNA (tramite *Real Time PCR*) su 25 campioni di tessuto tumorale primitivo e 11 campioni di tessuto tumorale primitivo metastatico inclusi in OCT e su 3 biopsie tumorali fresche. I risultati hanno mostrato che LINC00483 è *down-regolato* in maniera statisticamente significativa nei tumori CRC metastatici rispetto ai tumori primitivi non metastatici (Figura 28).

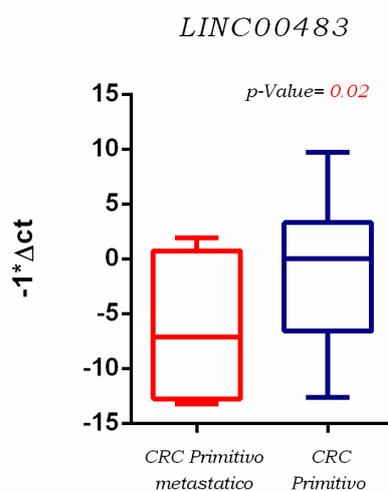


Figura 28. Il box-plot mostra l'espressione di LINC00483 in 11 tessuti CRC primitivi metastatici contro 28 tessuti CRC primitivi non metastatici. I risultati hanno mostrato che LINC00483 è *down-regolato* in maniera statisticamente significativa (p -value=0.02) nei tumori CRC metastatici contro i tumori primitivi non metastatici.

7.5.2 Trasfezione delle HCT-116 con LINC00483 e analisi di espressione dei geni *target*

Le analisi di espressione tramite *Real Time PCR* hanno mostrato che LINC00483 è costantemente *down*-regolato nei tessuti tumorali contro le adiacenti parti sane e nei tessuti tumorali metastatici contro i CRC primitivi non metastatici. Data la ipotetica funzione di questo lncRNA quale oncosoppressore nel CRC e, di conseguenza, i bassi livelli di espressione del suo trascritto nei tessuti tumorali, abbiamo indotto *in vitro* la sovra-espressione di LINC00483, trasfettando le HCT-116 con il vettore di espressione pcDNA3.1(+) (in cui è stata clonata la sequenza del cDNA di LINC00483) al fine di verificare come varia l'espressione dei suoi presunti geni *target* (positivamente correlati con LINC00483) identificati tramite approccio computazionale. L'RNA totale estratto dalle cellule HCT-116 è stato utilizzato sia per confermare l'efficienza di trasfezione (>99%) sia per analizzare, tramite *Real Time PCR*, l'espressione dei geni *target*. I risultati sono mostrati in Figura 29. L'espressione di tutti i geni *target* positivamente correlati a LINC00483 aumentava nei campioni trasfettati con LINC00483 rispetto ai campioni controllo, trasfettati con la molecola di plasmide priva dell'inserito di cDNA. Tutte le successive analisi sono state incentrate solo su quei geni *target* la cui espressione aumentava in maniera statisticamente significativa nei campioni trasfettati rispetto ai campioni controllo (*i.e.*, ACVR1B, ARL4A, CTNNA1, EIF6, EPCAM, HIGD2A, MIER3, PLS1, SRPK1, TPD52, VDR, ZDHHC9).

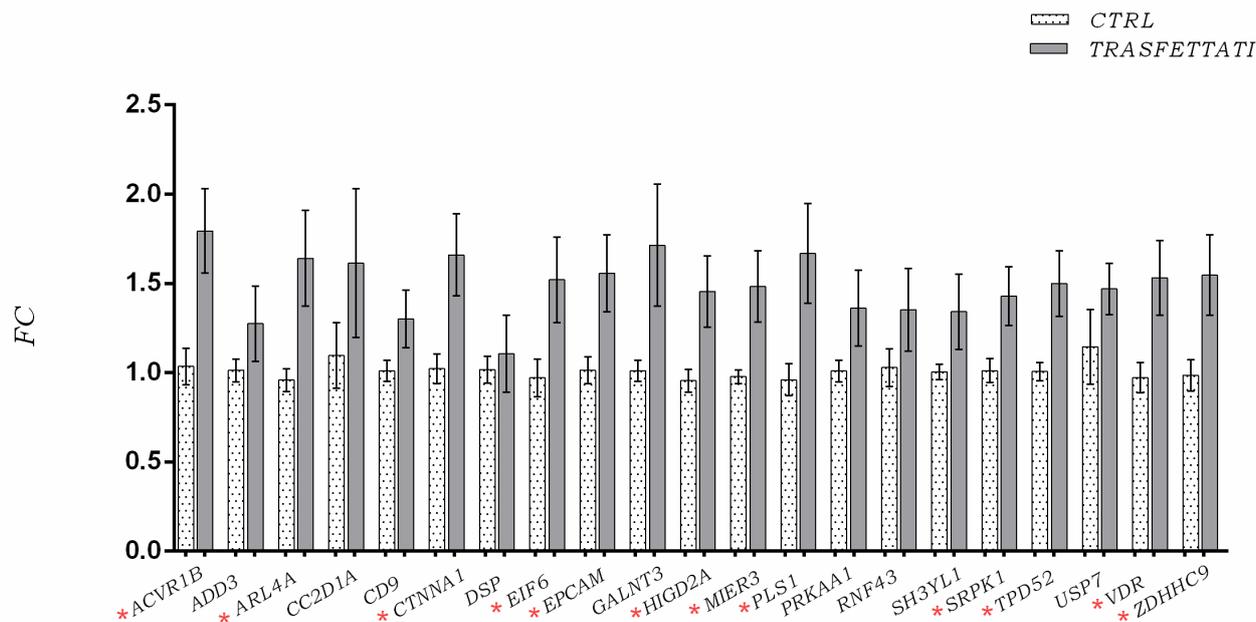


Figura 29. L'espressione di tutti i geni target (positivamente correlati a LINC00483) aumentava nelle cellule HCT-116 trasfettate con il vettore di espressione pcDNA3.1(+), contenente la sequenza del cDNA di LINC00483, rispetto ai controlli negativi di trasfezione. La molecola di plasmide priva dell'inserto è stata utilizzata come controllo negativo.
*: p-value < 0.05

7.5.3 Analisi di espressione di LINC00483 e dei geni target sulle linee cellulari

trattate con l'inibitore delle MAPK U0126

L'espressione di LINC00483 e dei geni target la cui espressione aumentava in maniera statisticamente significativa nelle cellule HCT-116 trasfettate rispetto ai controlli negativi è stata analizzata, tramite *Real Time PCR*, sulle stesse linee cellulari trattate con l'inibitore delle MAPK U0126, dopo 12 e 24 ore dal trattamento. I risultati hanno mostrato che l'espressione di LINC00483 aumentava dopo 12 ore dal trattamento, sebbene una variazione statisticamente significativa sia stata osservata dopo 24 ore dal trattamento (Figura 30). L'espressione di 4 geni target (i.e., ACVR1B, ARL4A, CTNNA1, PLS1) aumentava dopo 12 e 24 h dal trattamento (sebbene la

variazione di espressione dopo 12 ore non fosse statisticamente significativa), mentre l'espressione di 5 geni *target* (EIF6, EPCAM, HIGD2A, MIER3, ZDHHC9) aumentava dopo 24 ore dal trattamento con U0126, in maniera statisticamente significativa. L'espressione di tutti i geni *target* a 12 e 24 ore dal trattamento con U0126 è mostrata nelle Figure 31 e 32, rispettivamente.

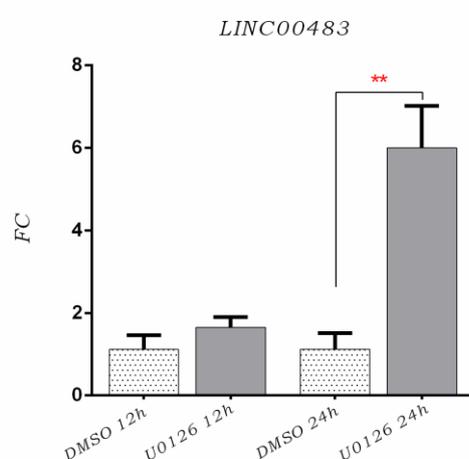


Figura 30. Espressione di LINC00483 nelle cellule HCT-116 dopo 12 e 24 h dal trattamento con U0126. I controlli sono stati trattati con un equivalente volume di DMSO. **:p-value<0.01

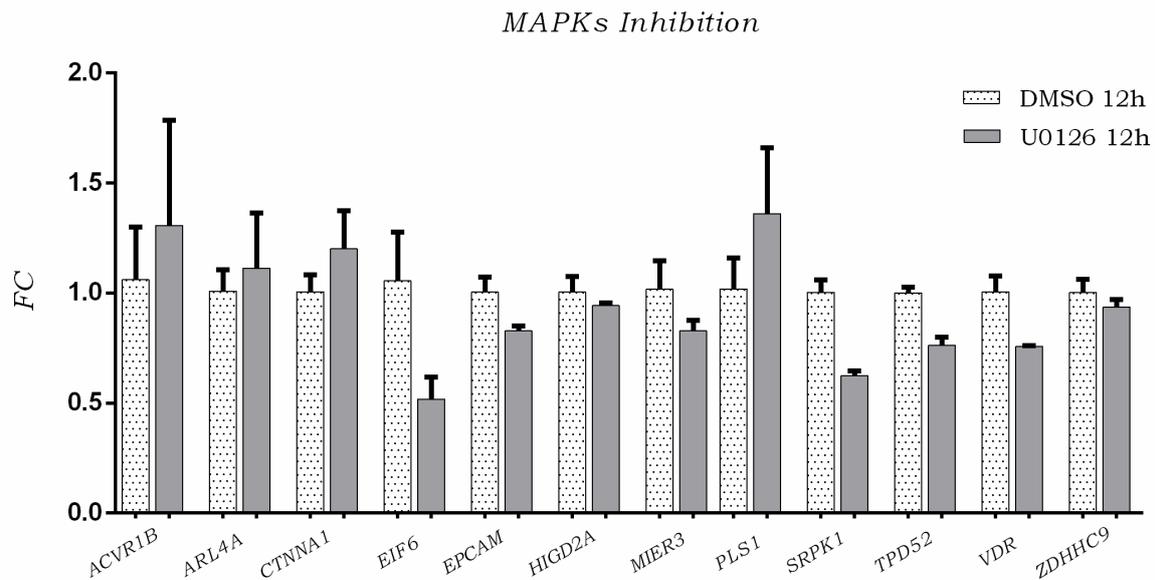


Figura 31. Espressione dei geni target correlati positivamente a LINC00483 nelle cellule HCT-116 dopo 12 h dal trattamento con U0126. L'espressione di 4 geni target (i.e., ACVR1B, ARL4A, CTNNA1, PLS1) aumentava dopo 12 h dal trattamento, sebbene non in maniera statisticamente significativa. I controlli sono stati trattati con un equivalente volume di DMSO.

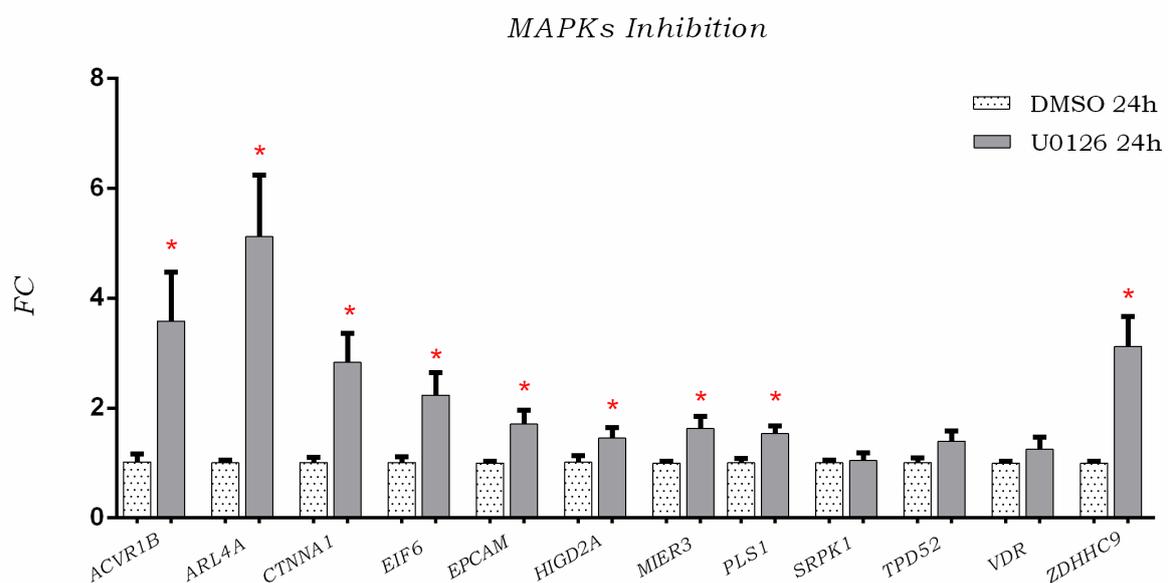


Figura 32. Espressione dei geni target correlati positivamente a LINC00483 nelle cellule HCT-116 dopo 24 h dal trattamento con U0126. I controlli sono stati trattati con un equivalente volume di DMSO.

*:p-value<0.05

Tutte le successive analisi di espressione sono state incentrate solo su quei geni *target*, positivamente correlati a LINC00483, la cui espressione aumentava in maniera statisticamente significativa dopo trattamento con l'inibitore delle MAPK U0126 (*i.e.*, ACVR1B, ARL4A, CTNNA1, EIF6, EPCAM, HIGD2A, MIER3, PLS1, ZDHHC9). Questa scelta è coerente con il modello di interazione miRNA-lncRNA-mRNA da noi ipotizzato secondo cui la molecola di lncRNA e gli RNA messaggeri che interagiscono indirettamente con essa sono legati da una correlazione positiva. Alla *up*-regolazione del lncRNA deve corrispondere, dunque, la *up*-regolazione dell'RNA messaggero *target*. Abbiamo, pertanto, selezionato solo quei geni la cui espressione aumentasse nelle linee cellulari trattate con l'inibitore della MAPK U0126 rispetto ai controlli negativi, proprio come aumentava l'espressione di LINC00483 negli stessi campioni.

7.5.4 Blocco del ciclo cellulare: analisi di espressione di LINC00483 e dei geni target

Per verificare se l'espressione di LINC00483 e dei suoi geni *target* venga modulata durante fenomeni di arresto del ciclo cellulare, è stata valutata la risposta delle HCT-116 alla condizione di assenza di siero: la deprivazione di siero per 12 - 24 ore blocca il ciclo cellulare e porta le cellule in fase G₀. L'espressione di LINC00483 e dei geni *target* selezionati (*i.e.*, ACVR1B, ARL4A, CTNNA1, EIF6, EPCAM, HIGD2A, MIER3, PLS, ZDHHC9) è stata valutata tramite *Real Time PCR* in due gruppi di campioni: un gruppo di campioni che è stato sottoposto al trattamento tramite

deprivazione di siero per 24 h e un secondo gruppo di cellule che è stato sincronizzato tramite *serum starvation* (no FBS) per 24 h e successivamente incubato con terreno fresco senza siero (deprivazione di siero) per ulteriori 24 h. I risultati dell'analisi sono mostrati in Figura 33. I dati dimostrano che LINC00483 aumentava la sua espressione in maniera statisticamente significativa dopo 24 h di assenza di sostanze nutritive, fenomeno che, nei modelli *in vitro*, induce un blocco momentaneo del ciclo cellulare (fase G0). Un simile *trend* di *up*-regolazione lo si osserva anche nel secondo modello sperimentale nel quale abbiamo indotto la deprivazione da siero in cellule che erano già in condizioni di blocco del ciclo cellulare. È interessante notare che due dei geni potenziali *target* di LINC00483 hanno mostrato in questi modelli sperimentali di blocco del ciclo cellulare un *trend* di espressione sovrapponibile a quello di LINC00483: ARL4A, HIGD2A.

Blocco del ciclo cellulare

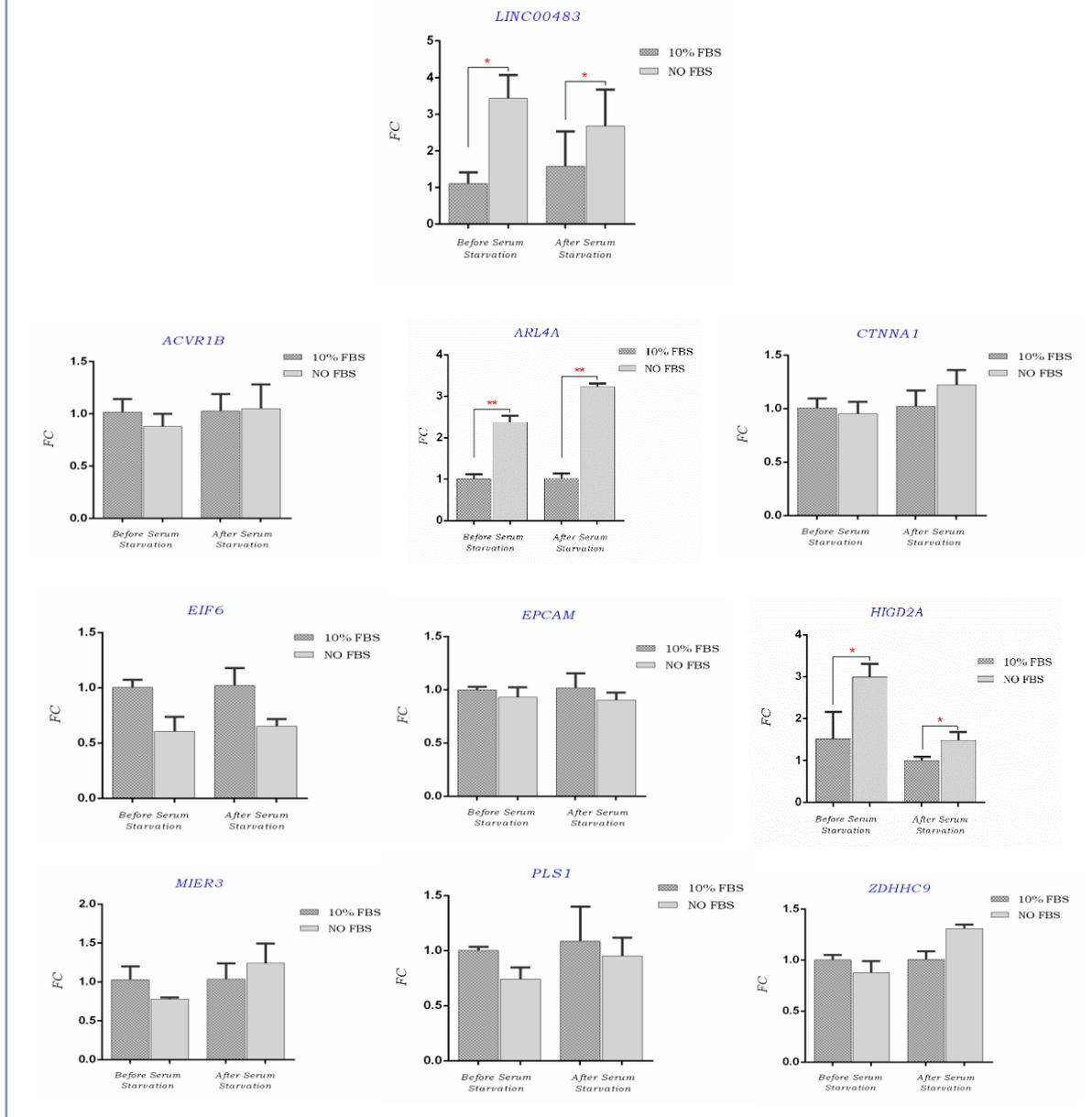


Figura 33. Espressione di LINC00483 e dei geni target selezionati (ACVR1B, ARL4A, CTNNA1, EIF6, EPCAM, HIGD2A, MIER3, PLS, ZDHHC9) dopo blocco artificiale del ciclo cellulare tramite deprivazione di siero. L'espressione di LINC00483 e dei geni target è stata valutata, tramite Real Time PCR, in due gruppi di campioni. Un gruppo di campioni è stato sottoposto al trattamento tramite deprivazione di siero per 24h. I controlli sono stati mantenuti in RPMI-1640 (10% FBS). Un secondo gruppo di campioni è stato sincronizzato tramite serum starvation (no FBS) per 24 h e successivamente incubato con terreno fresco senza siero (deprivazione di siero) per ulteriori 24 h. I controlli sono stati incubati con RPMI-1640 (10% FBS). * p -value \leq 0.05; ** p -value \leq 0.001

7.5.5 Analisi di espressione di LINC00483 e dei geni target sulle linee cellulari trattate con IL-6

Per rafforzare e confermare i dati relativi al coinvolgimento di LINC00483 nei processi cellulari associati alla inibizione delle proprietà invasive e metastatiche delle cellule tumorali, abbiamo analizzato (tramite *Real Time PCR*) l'espressione di LINC00483 e dei suoi geni *target* (i.e., ACVR1B, ARL4A, CTNNA1, EIF6, EPCAM, HIGD2A, MIER3, PLS, ZDHHC9) sulle cellule HCT-116 trattate con IL-6 (*Inteleukin-6*), una citochina multifunzionale la cui via di segnalazione è iper-attivata in molte malattie umane ed è stata associata con l'insorgenza e l'evoluzione del cancro¹⁶⁸. L'RNA totale estratto è stato utilizzato per valutare l'espressione di LINC00483 e dei geni *target* dopo 24 e 48 ore dal trattamento con IL-6, e di ZEB1 (un gene *marker* tipicamente associato alla transizione epitelio-mesenchima) dopo 48 h dal trattamento. I risultati hanno mostrato che l'espressione di ZEB1 aumentava dopo 48 h dal trattamento (Tabella 27), confermandone l'efficacia.

GENE	FC MEDIANO 48 h	p-value
ZEB1	2.45	0.03

Tabella 27. L'espressione di ZEB1 (valutata tramite *Real Time PCR*) aumentava dopo 48 h dal trattamento con IL-6. È mostrato il *fold change* mediano e il *T-test p-value*.

I risultati hanno mostrato che l'espressione di LINC00483 diminuiva in maniera statisticamente significativa dopo 48 h dal trattamento con IL-6 (Figura 34).

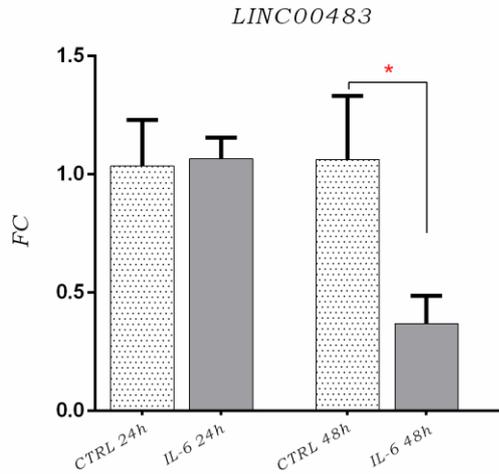


Figura 34. L'espressione di LINC00483 diminuisce in maniera statisticamente significativa dopo 48 h dal trattamento con IL-6. I controlli sono stati mantenuti in RPMI-1640 (10% FBS).

*:p-value<0.05

L'espressione di tutti i geni *target* dopo 24 e 48 ore dal trattamento è mostrata nelle Figure 35 e 36, rispettivamente.

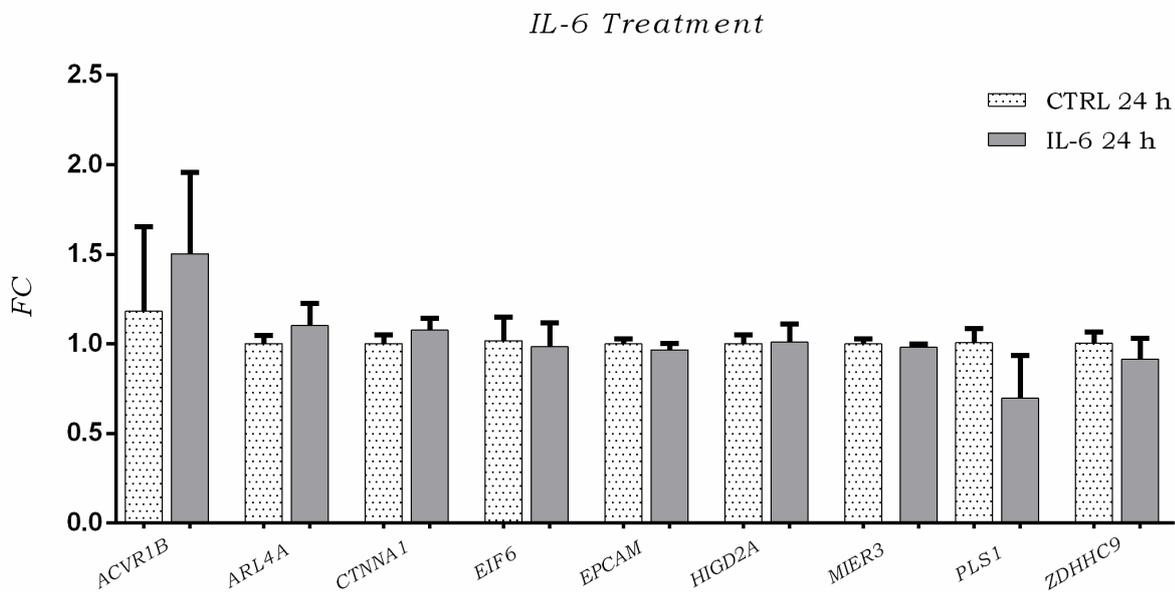


Figura 35. Espressione dei geni *target* di LINC00483 (valutata tramite Real Time PCR) dopo 24 h dal trattamento con IL-6. I controlli sono stati mantenuti in RPMI-1640 (10% FBS).

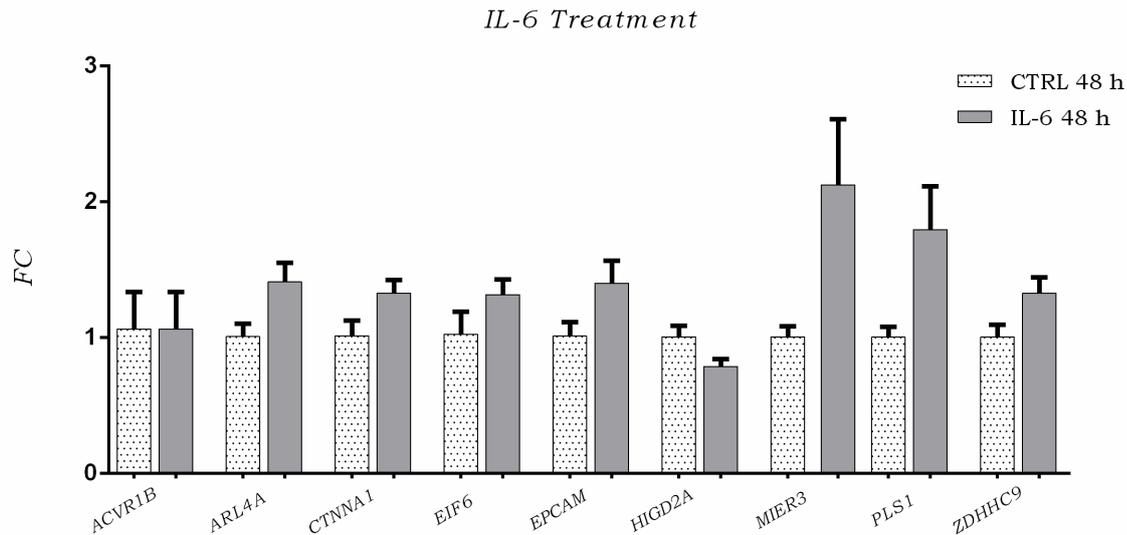


Figura 36. Espressione dei geni target di LINC00483 (valutata tramite Real Time PCR) dopo 48 h dal trattamento con IL-6. I controlli sono stati mantenuti in RPMI-1640 (10% FBS).

Tra tutti i geni *target* saggiati, solo un gene - HIGD2A - mostrava lo stesso *trend* di espressione di LINC00483: la sua espressione, infatti, diminuiva in maniera statisticamente significativa dopo 48 h dal trattamento con IL-6 (Figura 37).

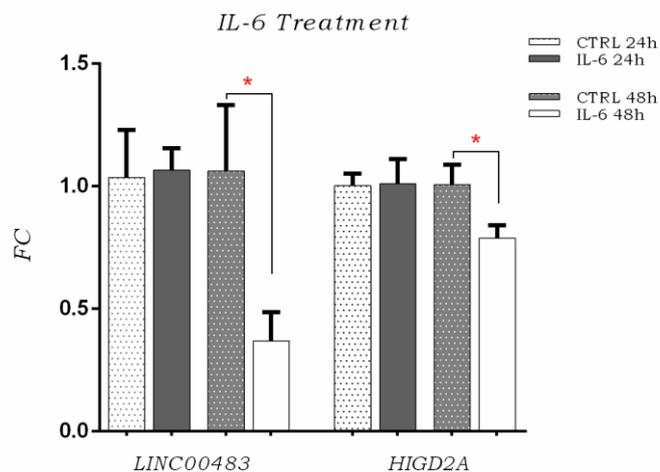


Figura 37. L'espressione di LINC00483 e del gene target HIGD2A (valutata tramite Real Time PCR) diminuiva dopo 48 h dal trattamento con IL-6. I controlli sono stati mantenuti in RPMI-1640 (10% FBS).

*: $p\text{-value} < 0.05$.

7.6 LINC00483 inibisce la capacità migratoria delle cellule HCT-116

È stato valutato il tasso di migrazione cellulare a seguito della trasfezione delle HCT-116 con LINC00483, al fine di saggiare il potenziale ruolo di oncosoppressore di LINC00483 e, dunque, la sua capacità di inibire la capacità migratoria delle cellule HCT-116. Il saggio della migrazione cellulare è stato allestito utilizzando il *kit Oris™ Universal Cell Migration Assembly (Platypus Technologies)*, secondo le istruzioni fornite dal protocollo. Abbiamo quantificato, tramite l'uso del *software ImageJ*, le cellule che hanno migrato all'interno delle zone di *detection* nel corso dei diversi *time points* [24 h (*reference wells*) 48, 72, 96 e 144 ore dopo la rimozione degli *Oris Cell seeding stoppers*]. I risultati hanno mostrato che il tasso di migrazione cellulare diminuiva in maniera statisticamente significativa nelle cellule trasfettate rispetto ai controlli negativi (MOCK), nel corso di tutti i *time points*, con il decremento più significativo a 96 h dalla trasfezione (Figura 38).

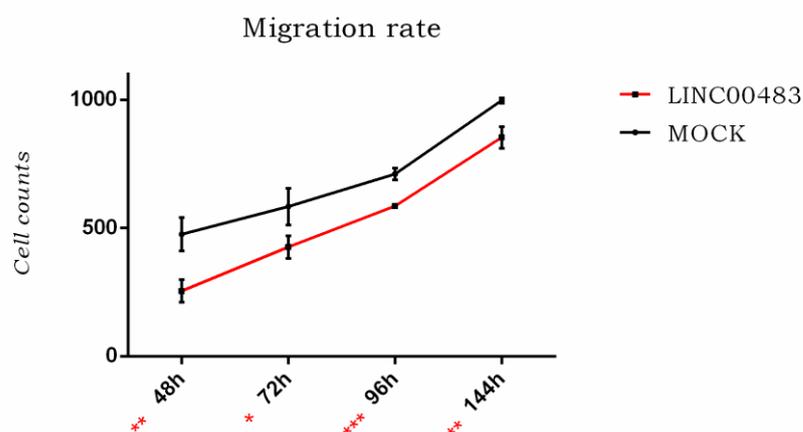


Figura 38. Numero di cellule migrate all'interno della zona di detection nelle HCT-116 trasfettate e nei rispettivi controlli negativi (MOCK), dopo 48, 72, 96 e 144 ore dalla trasfezione.

*: $p\text{-value} < 0.05$; **: $p\text{-value} < 0.01$; ***: $p\text{-value} < 0.001$

7.7 Saggio della vitalità cellulare: LINC00483 non altera la vitalità delle cellule HCT-116

Abbiamo valutato se la vitalità/proliferazione delle HCT-116 fosse influenzata da LINC00483 misurando la vitalità cellulare dopo 48 h dalla trasfezione, utilizzando il "Cell Counting kit-8" (CCK-8) e misurando i valori di assorbanza (tramite il Varioskan™ LUX) delle HCT-116 trasfettate con il vettore di espressione e con il rispettivo vettore *scramble*. I risultati non hanno evidenziato alcuna differenza significativa tra i due gruppi in esame (Figura 39). I dati ottenuti, pertanto, suggeriscono che la *down*-regolazione di LINC00483 non influenzi in maniera diretta la proliferazione cellulare.

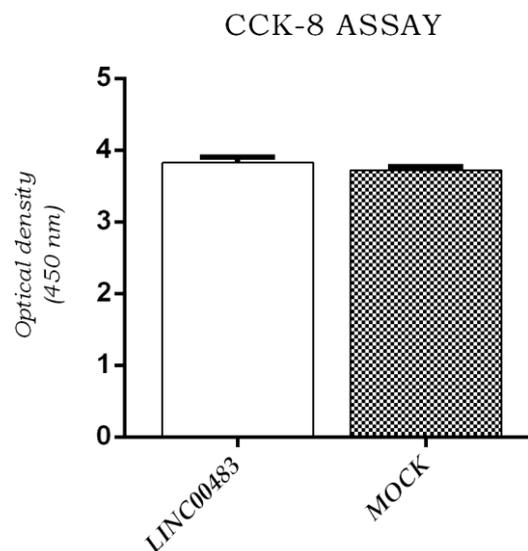


Figura 39. La proliferazione cellulare delle HCT-116 trasfettate con LINC00483 e dei rispettivi controlli negativi (MOCK) dopo 48 h dalla trasfezione è stata determinata tramite il saggio CCK-8. Nessuna differenza statisticamente significativa è stata rilevata tra i due gruppi in esame.

7.8 LINC00483 e HIGD2A: analisi di correlazione

I nostri dati, ottenuti dalle analisi di espressione di LINC00483 e dei suoi geni *target* (identificati computazionalmente) sulle linee cellulari di CRC HCT-116, hanno messo in evidenza un particolare gene *target*, HIGD2A (HIG1 hypoxia inducible domain family member 2A), la cui espressione veniva modulata in maniera simile a quella di LINC00483 nel corso di tutti i trattamenti sulle linee cellulari (Figura 40), suggerendo una loro potenziale interazione all'interno di un ipotetico circuito regolativo RNA-RNA.

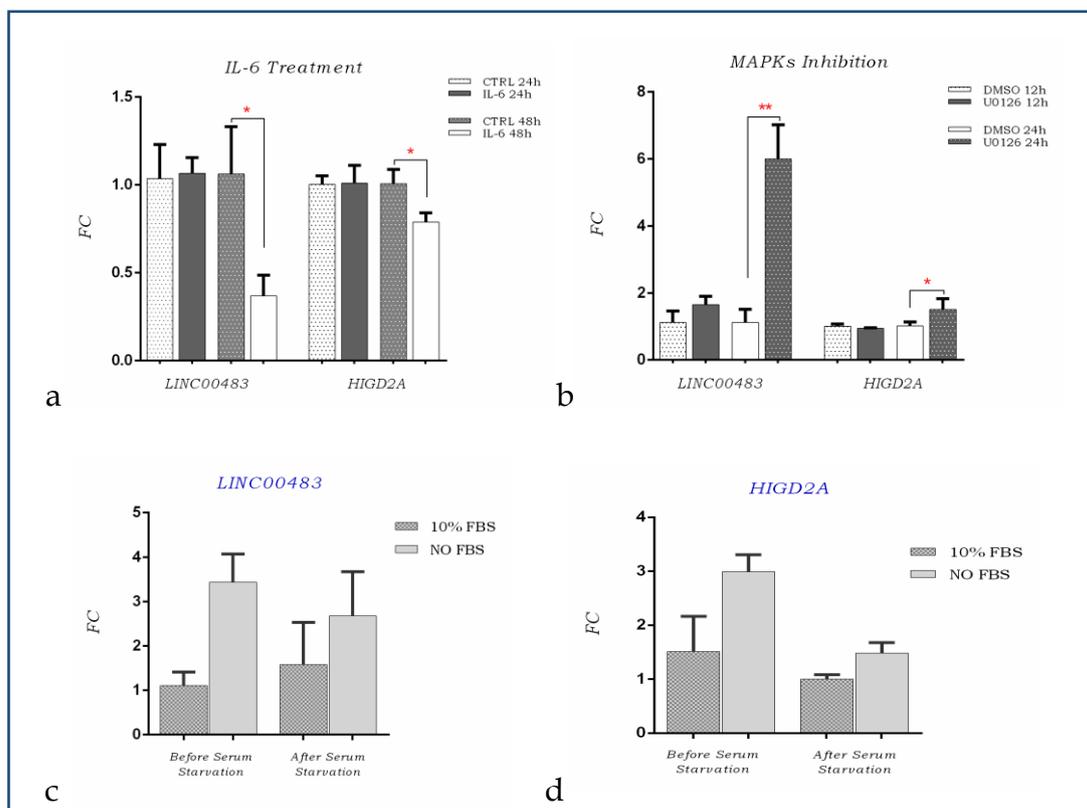


Figura 40. **a)** HIGD2A mostra un trend di espressione sovrapponibile a quello di LINC00483 (down-regolazione) nelle HCT-116 trattate con IL-6 rispetto ai controlli negativi, a 24 e 48h ; **b)** HIGD2A mostra un trend di espressione sovrapponibile a quello di LINC00483 (up-regolazione) nelle HCT-116 trattate con l'inibitore delle MAPK U0126 rispetto ai controlli negativi, a 12 e 24 h; **c-d)** espressione di LINC00483 e di HIGD2A nelle HCT-116 dopo blocco artificiale del ciclo cellulare, tramite deprivazione di siero.

*: $p\text{-value} < 0.05$; **: $p\text{-value} < 0.01$

Abbiamo verificato, tramite “R2 Genomics” (<https://hgserver1.amc.nl/cgi-bin/r2/main.cgi?&species=hs>), quale fosse l’espressione di HIGD2A in tutti i *dataset* contenenti dati di espressione relativi al CRC. I risultati hanno mostrato che questo gene è costantemente *down*-regolato nei tessuti tumorali rispetto ai controlli sani, e in tutte le manifestazioni più gravi del fenotipo tumorale (Tabella 28).

DATASET	CLASS	EXPRESSION	P-value
GSE3629 (121)	class: colorectal cancer vs nonneoplastic rectal mucosa	Down in colorectal cancer	1.30E-49
GSE3629 (121)	group: sporadic ca vs uc-nonca	Down in sporadic ca	3.40E-42
GSE20916 (145)	tissue: adenocarcinoma vs colon_tumor	Down in colon tumor	4.00E-26
GSE20916 (145)	tissue: colon_tumor vs normal_colon	Down in colon tumor	4.70E-21
GSE20916 (145)	tissue: carcinoma vs colon_tumor	Down in colon tumor	7.60E-12
GSE20916 (145)	malignacy: adenoma vs nd	Down in adenoma	6.90E-10
GSE41258 (390)	tissue: polyp vs primary tumor	Down in primary tumor	1.50E-09
GSE39582 (566)	cit-ccms: c2 vs c4	Down in c4	1.70E-09
GSE41258 (390)	tissue: lung metastasis vs polyp	Down in lung metastasis	1.10E-07
GSE41258 (390)	tissue: lung metastasis vs polyp	Down in lung metastasis	1.10E-07
GSE21510 (148)	tissue: cancer-lcm vs normal-homogenized	Down in cancer	8.30E-05
GSE41258 (390)	tissue: normal colon vs primary tumor	Down in primary tumor	1.10E-04
GSE39582 (566)	braf_mutation: na vs wt	Down in na	6.20E-04
GSE75316 (59)	braf_v600e_status: mut vs wt	Down in mut	8.50E-04
GSE18105 (111)	metastasis: metastatic recurrence vs nd	Down in metastatic recurrence	1.10E-03
GSE39582 (566)	tnm_stage: 1 vs 4	Down in 4	8.60E-03
GSE39582 (566)	cit-ccms: c1 vs c5	Down in c5	9.80E-03

Tabella 28. Espressione di HIGD2A in diversi *dataset* contenenti dati di espressione relativi al CRC (R2 Genomics). La tabella mostra il GEO ID di ciascun *dataset* (il numero dei campioni è indicato tra parentesi), le categorie a confronto, la relativa espressione di HIGD2A e il valore di *p-value* associato.

Abbiamo, anche, verificato (tramite “R2 Genomics”) quale fosse la correlazione tra l’espressione di LINC00483 e l’espressione di HIGD2A in diversi *dataset* contenenti dati di espressione relativi al CRC (Tabella 29). I risultati hanno mostrato che i due geni risultano correlati positivamente in tutti i *database* consultati.

DATASET	<i>R-value</i>	<i>p-value</i>
GSE45270 (13)	0.642	0.018
GSE39395 (24)	0.588	2.49E-03
GSE2109 (38)	0.491	1.75E-03
GSE37892 (130)	0.44	1.63E-07
GSE20916 (145)	0.387	1.54E-06
GSE41568 (133)	0.375	8.89E-06
GSE2109 (315)	0.358	6.07E-11
GSE92921 (59)	0.351	6.35E-03
GSE4107 (22)	0.324	0.141
GSE23878 (59)	0.318	0.014
GSE8671 (32)	0.31	0.085
GSE4183 (53)	0.263	0.057
GSE75316 (59)	0.251	0.055
GSE2109 (39)	0.242	0.138
GSE28702 (83)	0.236	0.032
GSE14333 (290)	0.217	2.01E-04
GSE13294 (155)	0.217	6.66E-03
GSE4554 (84)	0.197	0.072
GSE33114 (108)	0.196	0.042
GSE14333a	0.185	4.72E-04
GSE21510 (18)	0.177	0.483
GSE21510 (148)	0.158	0.055
GSE17538 (232)	0.157	0.017
COAD (286)	0.156	8.27E-03
GSE18088 (53)	0.136	0.33
GSE35896 (62)	0.127	0.327
GSE18105 (111)	0.124	0.193
GSE8671 (64)	0.109	0.391

Tabella 29. I dati in tabella mostrano che i due geni, LINC00483 e HIGD2A, sono positivamente correlati in tutti i *dataset* consultati (R2 Genomics). La tabella mostra il GEO ID di ciascun *dataset* (il numero di campioni è indicato tra parentesi), il valore del coefficiente di correlazione di *Pearson* e il valore di *p-value* associato.

Per verificare sperimentalmente il dato dell'analisi di correlazione ottenuto per via computazionale, abbiamo analizzato, tramite *Real Time PCR*, l'espressione di HIGD2A su 20 campioni di RNA da biopsie paraffinate di pazienti affetti da CRC, e sulle rispettive parti sane, prese dallo stesso paziente. I risultati hanno mostrato che HIGD2A è *down-regolato* in maniera statisticamente significativa nei tessuti tumorali contro le rispettive parti sane (Figura 41).

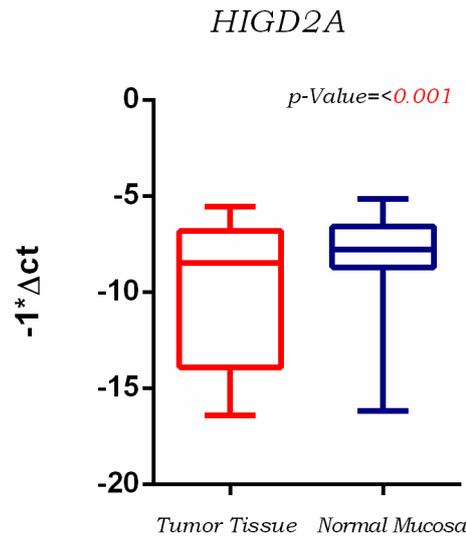


Figura 41. Il box plot in figura mostra la down-regolazione di HIGD2A in 20 biopsie CRC FFPE contro le rispettive parti adiacenti sane. PPIA è stato utilizzato come controllo endogeno.

Abbiamo, quindi calcolato il coefficiente di correlazione di *Pearson* tra i valori di espressione di LINC00483 nei tessuti tumorali CRC FFPE e i valori di espressione di HIGD2A negli stessi campioni. I risultati hanno mostrato che i due geni sono correlati positivamente, in maniera statisticamente significativa (Tabella 30).

<i>R-value</i>	<i>p-value</i>
0.41	0.05

Tabella 30. La tabella mostra il coefficiente di correlazione di *Pearson* tra i valori di espressione di LINC00483 e quelli di HIGD2A (con il valore di *p-value* associato), in 20 tessuti tumorali CRC FFPE contro le rispettive parti adiacenti sane. I due geni risultano positivamente correlati, in maniera statisticamente significativa.

8. *Discussione*

Il carcinoma del colon retto (CRC) rappresenta uno dei tumori più comuni al mondo ed è considerato la terza causa di morte correlata al cancro¹⁷¹. Il tasso di incidenza e di mortalità del CRC è notevolmente diminuito negli ultimi 10 anni, grazie ad un maggiore *screening* e alla diagnosi precoce. Tuttavia, una volta che il cancro ha metastatizzato, la sopravvivenza a 5 anni rimane molto bassa (<http://seer.cancer.gov/statfacts/html/colorect.html>). La comprensione dei meccanismi cellulari e molecolari che sottostanno al fenotipo aggressivo e metastatico delle cellule di CRC è, pertanto, fondamentale al fine di sviluppare nuove strategie terapeutiche attraverso le quali agire a monte del processo di metastatizzazione, inibendo l'evoluzione del tumore verso gli stadi più avanzati e incurabili della malattia. In questa prospettiva, nel presente lavoro abbiamo esplorato l'esistenza di lncRNA la cui alterata espressione fosse associata alle caratteristiche più gravi del CRC. Numerose evidenze sperimentali hanno, infatti, mostrato il ruolo fondamentale di queste molecole in tutti gli stadi della carcinogenesi e della formazione di metastasi¹⁵³.

In questo lavoro abbiamo utilizzato un approccio combinato di biologia computazionale e sperimentale, consultando inizialmente i *dataset* che contenessero dati di espressione relativi al CRC e identificando quei geni che fossero differenzialmente espressi in relazione alle caratteristiche clinico-patologiche più gravi di questo tumore. Tali geni sono stati utilizzati come "*bait genes*" (geni esca) per individuare i lncRNA ad essi associati. L'analisi di espressione, tramite *Real time PCR*, dei lncRNA selezionati mediante l'analisi

computazionale su biopsie di CRC e sulle rispettive parti adiacenti sane, ci ha permesso di identificare un lncRNA - LINC00483 - che risultava sotto-espresso in maniera statisticamente significativa nei tessuti tumorali rispetto alla adiacente mucosa sana. Inoltre, questo lncRNA risultava anche *down*-regolato nei tumori metastatici rispetto ai primitivi, suggerendo una associazione tra la sotto-espressione di LINC00483 e le caratteristiche più gravi del CRC, quali la formazione di metastasi. Questi risultati erano confermati dai dati computazionali, in quanto LINC00483 risultava costantemente *down*-regolato in tutti i *dataset* analizzati.

Per corroborare la nostra ipotesi relativa ad un potenziale coinvolgimento di LINC00483 nei meccanismi della inibizione dell'invasione cellulare e della formazione delle metastasi abbiamo indotto *in vitro* uno degli eventi chiave della cascata metastatica: la transizione epitelio-mesenchima (EMT).

È comunemente accettato che la transizione epitelio-mesenchima rappresenti un processo fondamentale per la formazione di metastasi distali. Si tratta di un complesso processo attraverso il quale le cellule epiteliali acquisiscono un fenotipo mesenchimale. La EMT si verifica in condizioni fisiologiche in diversi momenti dello sviluppo embrionale, come la formazione del mesoderma e la migrazione delle creste neurali. Nell'adulto, la EMT viene riattivata in determinate condizioni patologiche quali la rimarginazione delle ferite, la fibrosi, e la progressione tumorale^{172, 173}. In letteratura è ampiamente documentato il coinvolgimento dei lncRNA nella regolazione di questo processo fondamentale nell'evoluzione dei tumori. Zhang *et al.* hanno dimostrato che il lncRNA NEAT1 è

strettamente associato alla progressione del carcinoma al seno promuovendo la proliferazione cellulare e la EMT¹⁶⁴. Il lncRNA LINC01186 inibisce la migrazione e l'invasione tramite la transizione epitelio-mesenchima nel cancro ai polmoni¹⁶³, mentre la sovra-espressione del lncRNA linc-UBC1 promuove la proliferazione cellulare, la migrazione e l'invasione nel cancro al colon retto¹⁷⁴. Una associazione tra LINC00483 e la transizione epitelio-mesenchima è stata recentemente riportata per il tumore ai polmoni: Yang *et al.* hanno dimostrato che LINC00483 induce la EMT nell'adenocarcinoma polmonare¹⁷⁵.

Il *TGF-β1* (*Transforming Growth Factor- beta 1*) è considerato uno dei principali induttori *in vitro* della EMT¹⁷⁰ ed è stato dimostrato che il trattamento delle cellule di CRC con *TGF-β1* induce la EMT¹⁷⁶.

Per verificare il coinvolgimento di LINC00483 nei processi della EMT abbiamo, quindi, trattato le nostre linee cellulari di CRC - le HCT116 - con *TGF-β1*. L'espressione dei *markers* mesenchimali - MMP7, VIM e ZEB1 - aumentava dopo 24 ore dal trattamento, confermandone l'efficacia. I nostri risultati hanno mostrato che l'espressione di LINC00483 diminuiva nelle cellule trattate con *TGF-β1* rispetto ai controlli negativi dopo 24 ore dal trattamento, confermando i nostri dati preliminari relativi alla sua funzione di gene oncosoppressore nel CRC e alla associazione tra la sua diminuita espressione e le caratteristiche più aggressive delle cellule di CRC.

La *pathway* delle MAPK è cruciale per l'insorgenza e l'evoluzione del CRC^{177, 178}. È stato osservato che i pazienti affetti da CRC presentano frequentemente mutazioni puntiformi

nel gene K-ras, la proteina G che agisce a monte di questa *pathway*, rendendola costitutivamente attiva e portando, quindi, ad una iper-attivazione del segnale di trasduzione. Abbiamo indagato gli effetti della inibizione di questa via di segnalazione sull'espressione di LINC00483, trattando le cellule HCT-116 (mutate in K-ras) con l'inibitore delle MAPK U0126. Come dimostrato dal saggio MTT, U0126 riduceva la proliferazione delle cellule trattate rispetto ai controlli negativi, confermando l'efficacia del trattamento. L'analisi di espressione, tramite *Real Time PCR*, di LINC00483 ha dimostrato che la sua espressione aumentava in maniera statisticamente significativa nelle cellule trattate con U0126 rispetto ai controlli negativi, suggerendo così un'associazione tra blocco di una via di *signalling* iper-proliferativa e l'aumentata espressione di LINC00483. L'associazione tra LINC00483 e la *pathway* delle MAPK è stata già riportata per il carcinoma gastrico: LINC00483 promuoveva la proliferazione cellulare tramite *up-regolazione* di SPAG9 (una proteina *scaffold* nella cascata delle MAPK) e attivazione delle MAPK nei modelli *in vivo* e *in vitro* di carcinoma gastrico¹⁷⁹. Nel nostro modello tumorale i meccanismi che sottostanno alla funzione di questo lncRNA potrebbero essere contesto-dipendenti e, pertanto, determinare un effetto specifico sulla proliferazione cellulare e sulle vie di segnalazione che la controllano, differente da quello osservato in altri modelli tumorali. Infatti, è interessante notare che l'espressione di LINC00483 sembra essere particolarmente spiccata nel tessuto intestinale fisiologico e quasi completamente assente negli altri distretti corporei, suggerendo che questo lncRNA abbia un ruolo specifico e fondamentale proprio a livello dell'epitelio intestinale.

Negli ultimi anni, evidenze sempre maggiori hanno confermato che un gran numero di lncRNA esplica le proprie funzioni agendo da ceRNA (*competing endogenous RNA*) per inibire l'azione dei miRNA. Abbiamo, quindi, ipotizzato un potenziale ruolo di “*miRNA sponge*” per LINC00483. Secondo questo modello, i lncRNA legano e sequestrano i miRNA, impedendo a questi ultimi di degradare i loro mRNA *target* e inibirne l'espressione. Tra lncRNA, miRNA e mRNA si viene, dunque, a creare un circuito di interazione in cui alla *down*-regolazione del lncRNA corrisponde la *up*-regolazione del miRNA *target* (e viceversa) e, di conseguenza, la *down*-regolazione degli mRNA *target* del miRNA, poiché vengono degradati da questi ultimi. Ad esempio, LINC00974 aumenta l'espressione del gene KRT19 - *target* del miR-642 - sequestrando il miR-642 e promuovendo la proliferazione cellulare nel carcinoma epatocellulare¹⁸⁰. LINC00668 promuove la tumorigenesi e la progressione tumorale sequestrando il miR-188-5p e *up*-regolando USP47 nel cancro al colon retto¹⁸¹. lncRNA-ATB può *up*-regolare l'espressione di ZEB1 e ZEB2 competendo per il legame ai membri della famiglia del miR-200, promuovendo così la transizione epitelio-mesenchima nel tumore epatocellulare¹⁸². Un ruolo di ceRNA è stato documentato anche per LINC00483: nel carcinoma gastrico LINC00483 agisce da spugna molecolare per il miRNA oncosoppressore miR-30a-3p. Inoltre, esso ripristina l'espressione di SPAG9, che è negativamente regolato dal miR-30a-3p, e attiva il *signalling* delle MAPK nelle cellule del tumore gastrico¹⁷⁹. Yang *et al.* hanno, invece, indagato l'interazione tra LINC00483, il miR-144 ed il gene homeobox A10 (HOXA10) nell'adenocarcinoma polmonare (LAD) e l'effetto di tale interazione sulla EMT. LINC00483 e HOXA10 erano *up*-regolati, mentre il miR-144 era *down*-regolato in LAD. Il

silenziamento di LINC00483 determinava una riduzione dei livelli di HOXA10, MMP-2, MMP-9, vimentina ed N-caderina, ma un aumento dei livelli del miR-144 e della E-caderina¹⁷⁵.

In accordo con la supposta funzione di “*miRNA sponge*” per LINC00483 abbiamo identificato, per via computazionale, i miRNA che avessero siti di legame per LINC00483, che fossero negativamente correlati con LINC00483 e che avessero come *target* i geni positivamente correlati a LINC00483. Abbiamo elucidato il significato della *down-regolazione* di LINC00483 nel CRC e il suo ruolo nella progressione tumorale tramite esperimenti di sovra-espressione *in vitro*. I nostri risultati hanno suggerito che la sovra-espressione di LINC00483 potrebbe inibire le capacità migratorie delle cellule di CRC. Il saggio della migrazione cellulare ha, infatti, dimostrato che il tasso di migrazione diminuiva in maniera statisticamente significativa nelle cellule trasfettate con LINC00483 rispetto ai controlli negativi, nel corso di tutti i *time points*, con il decremento più significativo a 96 ore dalla trasfezione. Il saggio CCK8 ha, tuttavia, mostrato che la sovra-espressione di LINC00483 non sembra avere un effetto fenotipico diretto sulla proliferazione cellulare, in quanto non è stata osservata alcuna differenza significativa nel tasso di proliferazione tra il gruppo delle cellule trasfettate e il gruppo controllo. Questi risultati suggeriscono che LINC00483 sia coinvolto nei processi della proliferazione e della migrazione, sebbene una sua modulazione diretta sembri influenzare direttamente solo quest’ultimo processo. Più verosimilmente, la sua funzione si esprime in maniera concertata a quella di altre molecole o fattori di regolazione.

Le analisi di espressione dei geni potenziali *target* di LINC00483, identificati computazionalmente, sulle linee cellulari di CRC trasfettate hanno confermato la correlazione positiva tra l'espressione di LINC00483 e questi trascritti, in quanto tutti i geni *target* saggiati risultavano *up*-regolati nelle cellule HCT-116 trasfettate rispetto ai controlli negativi. Abbiamo, pertanto, verificato se questi geni fossero coinvolti nel *signalling* regolativo nel quale LINC00483 veniva modulato: EMT e controllo del ciclo cellulare. A tal scopo, abbiamo saggiato l'espressione dei *target* su linee cellulari di CRC trattate con IL-6, U0126 e sottoposte a deprivazione di siero, rispettivamente. Sebbene un certo numero di *target* venisse modulato in maniera simile a quella di LINC00483 nel corso dei vari trattamenti, i nostri esperimenti hanno messo in evidenza un particolare gene *target*, HIGD2A (*HIG1 hypoxia inducible domain family member 2a*), la cui espressione era strettamente associata a quella di LINC00483, in tutti i trattamenti sulle linee cellulari. Infatti, il trattamento delle HCT-116 con IL-6 (Interleuchina-6) mostrava che l'espressione di HIGD2A così come l'espressione di LINC00483 diminuiva dopo 48 ore dal trattamento. Precedenti studi hanno dimostrato il ruolo dell'Interleuchina-6 nella induzione della EMT, in diversi modelli tumorali^{183, 184}.

Dato altresì interessante, l'espressione di entrambi i geni - LINC00483 e HIGD2A - aumentava dopo 24 ore dal trattamento con l'inibitore delle MAPK U0126 e nelle HCT-116 sottoposte a deprivazione di siero per 24 e 48 ore. La *serum starvation* porta a cambiamenti nell'espressione di numerosi geni (*e.g.*, geni associati all'autofagia, geni coinvolti nel metabolismo degli acidi grassi) e determina l'uscita delle cellule dal ciclo cellulare, insieme ad una ridotta espressione dei geni responsabili della progressione attraverso il ciclo

cellulare, come H-ras, Myc e i componenti della *pathway* delle MAPK e di PI3K¹⁸⁵. I risultati del trattamento delle linee cellulari di CRC con l'inibitore delle MAPK U0126 e del trattamento tramite deprivazione di siero, presi insieme, potrebbero suggerire che LINC00483 sia controllato negativamente da qualche gene del ciclo cellulare, ipotesi che necessita di ulteriori approfondimenti.

Più verosimilmente, LINC00483 opera attraverso una complessa rete regolativa - basata sulla interazione tra molecole di RNA - nella quale si trova coinvolto il gene HIGD2A (così come anche altri dei geni *target* analizzati in questo studio) e attraverso la quale vengono inibite le *pathway* coinvolte nella proliferazione ed invasione cellulare. In letteratura la funzione di HIGD2A è poco documentata; tuttavia, Ameri *et al.* hanno dimostrato che HIGD1A (*HIG1 hypoxia inducible domain family member 1a*), un membro appartenente alla stessa famiglia, interagisce con la catena di trasporto degli elettroni, modula il consumo di ossigeno, la produzione di ROS e l'attività di AMPK per promuovere la sopravvivenza cellulare durante la *starvation* da glucosio, mentre reprime simultaneamente la crescita tumorale *in vivo*¹⁸⁶. È verosimile che HIGD2A svolga la stessa funzione di gene oncosoppressore nel nostro modello tumorale e che sia coinvolto nei processi della migrazione e della proliferazione cellulare. Le nostre analisi hanno infatti mostrato che l'espressione di HIGD2A era ridotta nei tessuti di CRC rispetto alla adiacente mucosa sana e tale espressione correlava positivamente e in maniera statisticamente significativa con quella di LINC00483.

Nonostante HIGD2A sia stato scelto come candidato di elezione, le nostre indagini hanno coinvolto altri geni (*i.e.*, ACVR1B, ARL4A, CTNNA1, EIF6, EPCAM, HIGD2A, MIER3, PLS1, SRPK1, TPD52, VDR, ZDHHC9) positivamente correlati a LINC00483, per alcuni dei quali è ampiamente documentata la funzione di geni oncosoppressori e la loro associazione con i processi della EMT. Ad esempio, Peng *et al.* hanno dimostrato che l'espressione di MIER3 (MIER *family member 3*) era significativamente ridotta nei CRC primari ed era associata con le metastasi e la prognosi negativa. Inoltre, la sovra-espressione di MIER3 inibiva significativamente la proliferazione delle cellule CRC, la migrazione e l'invasione *in vitro* e reprimeva la crescita del tumore e le metastasi *in vivo*. MIER3 inibiva la proliferazione cellulare e l'invasione tramite riduzione dei livelli di Sp1 e la conseguente repressione della transizione epitelio-mesenchima¹⁸⁷.

La deregolazione di VDR (*Vitamin D Receptor*) è correlata alla progressione e alla ricorrenza del CRC¹⁸⁸. La proteina VDR è espressa in quasi tutti i tipi cellulari e i tessuti umani normali, ma anche nelle linee cellulari e nei tumori di diversa origine¹⁸⁹. Elevati livelli di espressione di VDR sono associati all'assenza di linfonodi e alla prognosi positiva nel cancro al colon^{190, 191}, e la *down*-regolazione di VDR è stata osservata in una parte dei melanomi e dei tumori al colon, al seno, ai polmoni e alle ovaie¹⁹²⁻¹⁹⁵. Li *et al.* hanno dimostrato che la forma attiva della vitamina D3 - 1,25(OH)2D3 - attenua la EMT indotta da IL-1 β attivando il *signalling* di VDR nelle linee cellulari di CRC. Le analisi *Western Blot* mostravano che l'espressione di VDR diminuiva dopo stimolazione con IL-1 β , mentre i livelli dei *markers* epiteliali e mesenchimali - E-caderina e vimentina - diminuivano ed aumentavano, rispettivamente. Larriba *et al.* hanno dimostrato che SNAIL2 reprime

l'espressione di VDR e blocca l'induzione del fenotipo epiteliale promossa dalla 1,25(OH)2D3 nelle cellule di cancro al colon¹⁹⁶.

ACVR1B (*Activin A receptor type 1B*) codifica per il recettore 1B dell'attivina, che appartiene alla superfamiglia del TGF- β . Il TGF- β e l'attivina sono coinvolti nella regolazione della proliferazione cellulare, della migrazione e dell'apoptosi ed utilizzano uno specifico complesso recettore tipo I/tipo II per la trasduzione del segnale¹⁹⁷. Recenti studi hanno dimostrato che il *signalling* dell'attivina inibisce la crescita delle cellule tumorali di colon¹⁹⁸. La distruzione del gene ACVR1B tramite la tecnologia CRISPR-Cas9 in organoidi derivati da tumori CRC ha confermato la sua funzione di gene oncosoppressore nel CRC. Inoltre, la dimensione media dei tumori derivati dagli organoidi in cui era stato introdotto il gRNA ACVR1B era significativamente maggiore della dimensione dei tumori generati dagli organoidi controllo, confermando che ACVR1B rappresenta un TSG (*Tumor suppressor gene*) per il CRC¹⁹⁹. In un altro studio Bauer *et al.* hanno indagato il potenziale contributo dell'attivina e del suo *signalling* a valle, tramite ACVR1B, alla EMT nel cancro al colon. 72 ore dopo il trattamento delle cellule FET con attivina, l'espressione della E-caderina era ridotta rispetto ai controlli, mentre l'espressione della vimentina era aumentata, in accordo con la EMT. Inoltre si osservava la *down-regolazione* di p21 in tutti i *time points*¹⁹⁷. La delezione di p21 aumentava il tasso di tumorigenesi indotta da Ras o c-Myc, ed era associato con il profilo di espressione genica e le caratteristiche immunoistochimiche della EMT²⁰⁰. Questo effetto si esplicava tramite l'interazione di ACVR1B con la subunità p85 della fosfatidilinositolo-3-chinasi (PI3K) e conseguente attivazione del *signalling* PI3K/akt¹⁹⁷.

Diversi studi hanno dimostrato il ruolo oncosoppressore di CTNNA1 (*Catenin alpha 1*) in vari tumori²⁰¹⁻²⁰³. Nel CRC, i livelli di espressione di CTNNA1 erano marcatamente più bassi nei tessuti tumorali rispetto ai tessuti adiacenti sani e la sua sovra-espressione inibiva la proliferazione e la crescita delle cellule SW480 e SW620²⁰⁴.

Nelle nostre indagini, l'espressione dei suddetti geni correlava con l'espressione di LINC00483 nelle linee cellulari trasfettate, suggerendo che anche questi, come LINC00483, possano svolgere una funzione di oncosoppressori nel CRC, in accordo con i dati già presenti in letteratura. Tuttavia, poiché l'associazione di espressione tra LINC00483 ed i suoi presunti *target* non è stata costantemente osservata in tutti i processi cellulari modulati *in vitro* (tranne che per HIGD2A), questo tipo di relazione potrebbe essere strettamente associata o esclusiva di specifiche *pathway* molecolari e non ubiquitaria. Questi risultati non escludono, tuttavia, che tali geni possano far parte di una *network* regolativa che vede coinvolto LINC00483 ma di cui questo lncRNA non è il principale e diretto controllore.

In conclusione, i nostri risultati suggeriscono che LINC00483 possa essere identificato come un nuovo gene oncosoppressore nel CRC e possa spiegare la sua funzione tramite una inibizione dei processi della transizione epitelio-mesenchima e della invasione e proliferazione cellulare, sebbene i meccanismi molecolari alla base della sua azione debbano essere ancora chiariti. In questa prospettiva, un'altra ipotesi che si potrebbe verificare ma che non è stata presentata in questa tesi, è quella di una interazione diretta tra LINC00483 e i trascritti. In questo modello, il tipo di relazione tra le due molecole di

RNA (lncRNA e mRNA) non cambia, poiché alla *up/down*-regolazione dell'una corrisponde la *up/down*-regolazione dell'altra, tuttavia diverso è il meccanismo attraverso cui si esplica la loro interazione; in questo caso, infatti, il lncRNA non lega il miRNA impedendo a quest'ultimo di degradare gli mRNA *target*, ma tramite complementarità di sequenza lega in maniera diretta l'mRNA *target* proteggendolo dall'azione del miRNA, poiché ne maschera i siti di legame. Wu *et al.* ad esempio, hanno dimostrato che il lncRNA THOR promuove la staminalità e la migrazione delle cellule di osteosarcoma (OS) legando in maniera diretta la regione centrale del 3'UTR del gene SOX9, aumentando la stabilità del suo mRNA e la sua espressione nelle cellule OS²⁰⁵. Lo stesso tipo di interazione diretta tra lncRNA THOR e il 3'UTR di SOX9 era stata riportata da Song *et al.* nel carcinoma gastrico²⁰⁶. Attualmente è in corso una identificazione dei potenziali *target* diretti di LINC00483, in quanto, da nostre indagini computazionali preliminari, risulta che la sequenza di questo lncRNA contiene un elemento ALU che risulta essere complementare alle sequenze di molti trascritti. È verosimile, pertanto, che LINC00483 possa legare tali trascritti e promuoverne la stabilità e la protezione contro la degradazione mediata dai miRNA. Tra i potenziali *target* diretti di LINC00483, abbiamo identificato un gene - PIGR - che mostra una regione di complementarità molto ampia con LINC00483 in corrispondenza del 3'UTR del suo trascritto. L'mRNA di PIGR è altamente espresso nell'epitelio sano del colon, ma numerosi *reports* hanno dimostrato la *down*-regolazione della sua espressione durante la progressione del cancro al colon²⁰⁷⁻²⁰⁹. Un dato notevolmente interessante è relativo alla recente caratterizzazione di una regione del 3'UTR di PIGR che include un segnale di poliadenilazione, diversi elementi ripetuti in

tandem, nonché sequenze consenso di legame a proteine coinvolte nella regolazione della stabilità degli mRNA²¹⁰. Potremmo, pertanto, ipotizzare che LINC00483 leghi questa regione del 3'UTR di PIGR, promuovendo la stabilità del suo mRNA. I meccanismi molecolari attraverso cui si esplica questa funzione dovranno, tuttavia, essere elucidati.

Un'altra ricerca parallela è quella relativa alla identificazione dei fattori di trascrizione (TF) che potrebbero potenzialmente regolare l'espressione di LINC00483. In particolare, vorremmo approfondire la funzione di un potenziale TF di LINC00483 - HNF4A - la cui espressione risulta essere positivamente correlata a quella di LINC00483 in tutti i *dataset* analizzati. Questo TF ha suscitato, altresì, il nostro interesse in quanto viene fosforilato e inibito da ERK1²¹¹, corroborando il nostro dato relativo alla aumentata espressione di LINC00483 nelle linee cellulari trattate con l'inibitore delle MAPK U0126. Verificheremo, pertanto, il controllo trascrizionale di HNF4A su LINC00483 trattando le linee cellulari con un inibitore specifico di questo TF e verificando come viene modulata l'espressione di LINC00483. Ricostruire la rete regolativa all'interno della quale opera questo lncRNA è, infatti, fondamentale per comprenderne le dinamiche di azione nel CRC e sviluppare nuovi approcci terapeutici per il trattamento dei CRC metastatici.

9. Bibliografia

1. Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS. **Non-coding RNAs: regulators of disease.** *J Pathol.* 2010;220(2):126-39.
2. Prasanth KV, Spector DL. **Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum.** *Genes Dev.* 2007;21(1):11-42.
3. Mattick JS. **Non-coding RNAs: the architects of eukaryotic complexity.** *EMBO Rep.* 2001;2(11):986-91.
4. Szymanski M, Barciszewski J. **RNA regulation in mammals.** *Ann N Y Acad Sci.* 2006;1067:461-8.
5. Prensner JR, Chinnaiyan AM. **The emergence of lncRNAs in cancer biology.** *Cancer Discov.* 2011;1(5):391-407.
6. Dhanoa JK, Sethi RS, Verma R, Arora JS, Mukhopadhyay CS. **Long non-coding RNA: its evolutionary relics and biological implications in mammals: a review.** *J Anim Sci Technol.* 2018;60:25.
7. Eddy SR. **Non-coding RNA genes and the modern RNA world.** *Nat Rev Genet.* 2001;2(12):919-29.
8. Brosnan CA, Voinnet O. **The long and the short of noncoding RNAs.** *Curr Opin Cell Biol.* 2009;21(3):416-25.
9. Guan D, Zhang W, Zhang W, Liu GH, Belmonte JC. **Switching cell fate, ncRNAs coming to play.** *Cell Death Dis.* 2013;4:e464.

10. Horiuchi T, Aigaki T. **Alternative trans-splicing: a novel mode of pre-mRNA processing.** *Biol Cell.* 2006;98(2):135-40.
11. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. **Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals.** *Nature.* 2009;458(7235):223-7.
12. Ragusa M, Barbagallo C, Statello L, Condorelli AG, Battaglia R, Tamburello L, et al. **Non-coding landscapes of colorectal cancer.** *World J Gastroenterol.* 2015;21(41):11709-39.
13. Johnsson P, Lipovich L, Grander D, Morris KV. **Evolutionary conservation of long non-coding RNAs; sequence, structure, function.** *Biochim Biophys Acta.* 2014;1840(3):1063-71.
14. Lapidot M, Pilpel Y. **Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms.** *EMBO Rep.* 2006;7(12):1216-22.
15. Dieci G, Fiorino G, Castelnuovo M, Teichmann M, Pagano A. **The expanding RNA polymerase III transcriptome.** *Trends Genet.* 2007;23(12):614-22.
16. Goodrich JA, Kugel JF. **Non-coding-RNA regulators of RNA polymerase II transcription.** *Nat Rev Mol Cell Biol.* 2006;7(8):612-6.
17. Gonzalez I, Munita R, Agirre E, Dittmer TA, Gysling K, Misteli T, et al. **A lncRNA regulates alternative splicing via establishment of a splicing-specific chromatin signature.** *Nat Struct Mol Biol.* 2015;22(5):370-6.
18. 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.

19. Gardini A, Shiekhattar R. **The many faces of long noncoding RNAs.** *FEBS J.* 2015;282(9):1647-57.
20. Wilusz JE. **Long noncoding RNAs: Re-writing dogmas of RNA processing and stability.** *Biochim Biophys Acta.* 2016;1859(1):128-38.
21. Louro R, Smirnova AS, Verjovski-Almeida S. **Long intronic noncoding RNA transcription: expression noise or expression choice?** *Genomics.* 2009;93(4):291-8.
22. Gomes AQ, Nolasco S, Soares H. **Non-coding RNAs: multi-tasking molecules in the cell.** *Int J Mol Sci.* 2013;14(8):16010-39.
23. Romero-Barrios N, Legascue MF, Benhamed M, Ariel F, Crespi M. **Splicing regulation by long noncoding RNAs.** *Nucleic Acids Res.* 2018;46(5):2169-84.
24. Khandelwal A, Bacolla A, Vasquez KM, Jain A. **Long non-coding RNA: A new paradigm for lung cancer.** *Mol Carcinog.* 2015;54(11):1235-51.
25. Kino T, Hurt DE, Ichijo T, Nader N, Chrousos GP. **Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor.** *Sci Signal.* 2010;3(107):ra8.
26. 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(7128):666-70.
27. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. **Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis.** *Nature.* 2010;464(7291):1071-6.

28. Luke B, Lingner J. **TERRA: telomeric repeat-containing RNA.** *EMBO J.* 2009;28(17):2503-10.
29. Mercer TR, Dinger ME, Sunkin SM, Mehler MF, Mattick JS. **Specific expression of long noncoding RNAs in the mouse brain.** *Proc Natl Acad Sci U S A.* 2008;105(2):716-21.
30. Rinn JL, Chang HY. **Genome regulation by long noncoding RNAs.** *Annu Rev Biochem.* 2012;81:145-66.
31. Bhaskaran M, Mohan M. **MicroRNAs: history, biogenesis, and their evolving role in animal development and disease.** *Vet Pathol.* 2014;51(4):759-74.
32. Saini HK, Griffiths-Jones S, Enright AJ. **Genomic analysis of human microRNA transcripts.** *Proc Natl Acad Sci U S A.* 2007;104(45):17719-24.
33. Hutvagner G, Zamore PD. **A microRNA in a multiple-turnover RNAi enzyme complex.** *Science.* 2002;297(5589):2056-60.
34. Aukerman MJ, Sakai H. **Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes.** *Plant Cell.* 2003;15(11):2730-41.
35. Wu L, Fan J, Belasco JG. **MicroRNAs direct rapid deadenylation of mRNA.** *Proc Natl Acad Sci U S A.* 2006;103(11):4034-9.
36. Lytle JR, Yario TA, Steitz JA. **Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR.** *Proc Natl Acad Sci U S A.* 2007;104(23):9667-72.
37. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. **Prediction of mammalian microRNA targets.** *Cell.* 2003;115(7):787-98.

38. Bartel DP. **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell.* 2004;116(2):281-97.
39. Lai EC. **Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation.** *Nat Genet.* 2002;30(4):363-4.
40. Qu S, Yang X, Li X, Wang J, Gao Y, Shang R, et al. **Circular RNA: A new star of noncoding RNAs.** *Cancer Lett.* 2015;365(2):141-8.
41. Suzuki H, Tsukahara T. **A view of pre-mRNA splicing from RNase R resistant RNAs.** *Int J Mol Sci.* 2014;15(6):9331-42.
42. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, et al. **Circular RNAs are abundant, conserved, and associated with ALU repeats.** *RNA.* 2013;19(2):141-57.
43. Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, et al. **Circular intronic long noncoding RNAs.** *Mol Cell.* 2013;51(6):792-806.
44. Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, et al. **Exon-intron circular RNAs regulate transcription in the nucleus.** *Nat Struct Mol Biol.* 2015;22(3):256-64.
45. Meng S, Zhou H, Feng Z, Xu Z, Tang Y, Li P, et al. **CircRNA: functions and properties of a novel potential biomarker for cancer.** *Mol Cancer.* 2017;16(1):94.
46. Hansen TB, Kjems J, Damgaard CK. **Circular RNA and miR-7 in cancer.** *Cancer Res.* 2013;73(18):5609-12.
47. Han L, Zhang G, Zhang N, Li H, Liu Y, Fu A, et al. **Prognostic potential of microRNA-138 and its target mRNA PDK1 in sera for patients with non-small cell lung cancer.** *Med Oncol.* 2014;31(9):129.

48. Hansen TB, Wiklund ED, Bramsen JB, Villadsen SB, Statham AL, Clark SJ, et al. **miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA.** *EMBO J.* 2011;30(21):4414-22.
49. Ragusa M, Barbagallo C, Brex D, Caponnetto A, Cirnigliaro M, Battaglia R, et al. **Molecular Crosstalking among Noncoding RNAs: A New Network Layer of Genome Regulation in Cancer.** *Int J Genomics.* 2017;2017:4723193.
50. Xue M, Li X, Wu W, Zhang S, Wu S, Li Z, et al. **Upregulation of long non-coding RNA urothelial carcinoma associated 1 by CCAAT/enhancer binding protein alpha contributes to bladder cancer cell growth and reduced apoptosis.** *Oncol Rep.* 2014;31(5):1993-2000.
51. Zhang L, Cao X, Zhang L, Zhang X, Sheng H, Tao K. **UCA1 overexpression predicts clinical outcome of patients with ovarian cancer receiving adjuvant chemotherapy.** *Cancer Chemother Pharmacol.* 2016;77(3):629-34.
52. Ji P, Diederichs S, Wang W, Boing S, Metzger R, Schneider PM, et al. **MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer.** *Oncogene.* 2003;22(39):8031-41.
53. West JA, Davis CP, Sunwoo H, Simon MD, Sadreyev RI, Wang PI, et al. **The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites.** *Mol Cell.* 2014;55(5):791-802.
54. Tsang FH, Au SL, Wei L, Fan DN, Lee JM, Wong CC, et al. **Long non-coding RNA HOTTIP is frequently up-regulated in hepatocellular carcinoma and is targeted by tumour suppressive miR-125b.** *Liver Int.* 2015;35(5):1597-606.

55. Chiyomaru T, Yamamura S, Fukuhara S, Yoshino H, Kinoshita T, Majid S, et al. **Genistein inhibits prostate cancer cell growth by targeting miR-34a and oncogenic HOTAIR.** *PLoS One*. 2013;8(8):e70372.
56. Chiyomaru T, Fukuhara S, Saini S, Majid S, Deng G, Shahryari V, et al. **Long non-coding RNA HOTAIR is targeted and regulated by miR-141 in human cancer cells.** *J Biol Chem*. 2014;289(18):12550-65.
57. Tsang WP, Ng EK, Ng SS, Jin H, Yu J, Sung JJ, et al. **Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer.** *Carcinogenesis*. 2010;31(3):350-8.
58. Zhou X, Ye F, Yin C, Zhuang Y, Yue G, Zhang G. **The Interaction Between MiR-141 and lncRNA-H19 in Regulating Cell Proliferation and Migration in Gastric Cancer.** *Cell Physiol Biochem*. 2015;36(4):1440-52.
59. Barbagallo D, Condorelli A, Ragusa M, Salito L, Sammito M, Banelli B, et al. **Dysregulated miR-671-5p / CDR1-AS / CDR1 / VSNL1 axis is involved in glioblastoma multiforme.** *Oncotarget*. 2016;7(4):4746-59.
60. Song P, Yin SC. **Long non-coding RNA EWSAT1 promotes human nasopharyngeal carcinoma cell growth in vitro by targeting miR-326/-330-5p.** *Aging (Albany NY)*. 2016;8(11):2948-60.
61. Xia T, Chen S, Jiang Z, Shao Y, Jiang X, Li P, et al. **Long noncoding RNA FER1L4 suppresses cancer cell growth by acting as a competing endogenous RNA and regulating PTEN expression.** *Sci Rep*. 2015;5:13445.

62. Deng L, Yang SB, Xu FF, Zhang JH. **Long noncoding RNA CCAT1 promotes hepatocellular carcinoma progression by functioning as let-7 sponge.** *J Exp Clin Cancer Res.* 2015;34:18.
63. Su DN, Wu SP, Chen HT, He JH. **HOTAIR, a long non-coding RNA driver of malignancy whose expression is activated by FOXC1, negatively regulates miRNA-1 in hepatocellular carcinoma.** *Oncol Lett.* 2016;12(5):4061-7.
64. Song B, Guan Z, Liu F, Sun D, Wang K, Qu H. **Long non-coding RNA HOTAIR promotes HLA-G expression via inhibiting miR-152 in gastric cancer cells.** *Biochem Biophys Res Commun.* 2015;464(3):807-13.
65. Jin C, Yan B, Lu Q, Lin Y, Ma L. **Reciprocal regulation of Hsa-miR-1 and long noncoding RNA MALAT1 promotes triple-negative breast cancer development.** *Tumour Biol.* 2016;37(6):7383-94.
66. Tan J, Qiu K, Li M, Liang Y. **Double-negative feedback loop between long non-coding RNA TUG1 and miR-145 promotes epithelial to mesenchymal transition and radioresistance in human bladder cancer cells.** *FEBS Lett.* 2015;589(20 Pt B):3175-81.
67. Ma F, Wang SH, Cai Q, Jin LY, Zhou D, Ding J, et al. **Long non-coding RNA TUG1 promotes cell proliferation and metastasis by negatively regulating miR-300 in gallbladder carcinoma.** *Biomed Pharmacother.* 2017;88:863-9.
68. Xie CH, Cao YM, Huang Y, Shi QW, Guo JH, Fan ZW, et al. **Long non-coding RNA TUG1 contributes to tumorigenesis of human osteosarcoma by sponging miR-9-5p and regulating POU2F1 expression.** *Tumour Biol.* 2016;37(11):15031-41.

69. Wang Y, Liu Z, Yao B, Dou C, Xu M, Xue Y, et al. **Long non-coding RNA TUSC7 acts a molecular sponge for miR-10a and suppresses EMT in hepatocellular carcinoma.** *Tumour Biol.* 2016;37(8):11429-41.
70. Liu Q, Huang J, Zhou N, Zhang Z, Zhang A, Lu Z, et al. **LncRNA loc285194 is a p53-regulated tumor suppressor.** *Nucleic Acids Res.* 2013;41(9):4976-87.
71. Qi P, Xu MD, Shen XH, Ni SJ, Huang D, Tan C, et al. **Reciprocal repression between TUSC7 and miR-23b in gastric cancer.** *Int J Cancer.* 2015;137(6):1269-78.
72. Tuo YL, Li XM, Luo J. **Long noncoding RNA UCA1 modulates breast cancer cell growth and apoptosis through decreasing tumor suppressive miR-143.** *Eur Rev Med Pharmacol Sci.* 2015;19(18):3403-11.
73. Li HJ, Li X, Pang H, Pan JJ, Xie XJ, Chen W. **Long non-coding RNA UCA1 promotes glutamine metabolism by targeting miR-16 in human bladder cancer.** *Jpn J Clin Oncol.* 2015;45(11):1055-63.
74. Wang F, Ying HQ, He BS, Pan YQ, Deng QW, Sun HL, et al. **Upregulated lncRNA-UCA1 contributes to progression of hepatocellular carcinoma through inhibition of miR-216b and activation of FGFR1/ERK signaling pathway.** *Oncotarget.* 2015;6(10):7899-917.
75. Yang Y, Jiang Y, Wan Y, Zhang L, Qiu J, Zhou S, et al. **UCA1 functions as a competing endogenous RNA to suppress epithelial ovarian cancer metastasis.** *Tumour Biol.* 2016;37(8):10633-41.

76. Song P, Ye LF, Zhang C, Peng T, Zhou XH. **Long non-coding RNA XIST exerts oncogenic functions in human nasopharyngeal carcinoma by targeting miR-34a-5p.** *Gene*. 2016;592(1):8-14.
77. Chang S, Chen B, Wang X, Wu K, Sun Y. **Long non-coding RNA XIST regulates PTEN expression by sponging miR-181a and promotes hepatocellular carcinoma progression.** *BMC Cancer*. 2017;17(1):248.
78. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, et al. **Circular RNAs are a large class of animal RNAs with regulatory potency.** *Nature*. 2013;495(7441):333-8.
79. Zheng J, Liu X, Xue Y, Gong W, Ma J, Xi Z, et al. **TTBK2 circular RNA promotes glioma malignancy by regulating miR-217/HNF1beta/Derlin-1 pathway.** *J Hematol Oncol*. 2017;10(1):52.
80. Xie H, Ren X, Xin S, Lan X, Lu G, Lin Y, et al. **Emerging roles of circRNA_001569 targeting miR-145 in the proliferation and invasion of colorectal cancer.** *Oncotarget*. 2016;7(18):26680-91.
81. Huang G, Zhu H, Shi Y, Wu W, Cai H, Chen X. **circ-ITCH plays an inhibitory role in colorectal cancer by regulating the Wnt/beta-catenin pathway.** *PLoS One*. 2015;10(6):e0131225.
82. Li F, Zhang L, Li W, Deng J, Zheng J, An M, et al. **Circular RNA ITCH has inhibitory effect on ESCC by suppressing the Wnt/beta-catenin pathway.** *Oncotarget*. 2015;6(8):6001-13.

83. Yu L, Gong X, Sun L, Zhou Q, Lu B, Zhu L. **The Circular RNA Cdr1as Act as an Oncogene in Hepatocellular Carcinoma through Targeting miR-7 Expression.** *PLoS One.* 2016;11(7):e0158347.
84. Chen L, Zhang S, Wu J, Cui J, Zhong L, Zeng L, et al. **circRNA_100290 plays a role in oral cancer by functioning as a sponge of the miR-29 family.** *Oncogene.* 2017;36(32):4551-61.
85. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, et al. **Natural RNA circles function as efficient microRNA sponges.** *Nature.* 2013;495(7441):384-8.
86. Zhao ZJ, Shen J. **Circular RNA participates in the carcinogenesis and the malignant behavior of cancer.** *RNA Biol.* 2017;14(5):514-21.
87. Wang Q, Tang H, Yin S, Dong C. **Downregulation of microRNA-138 enhances the proliferation, migration and invasion of cholangiocarcinoma cells through the upregulation of RhoC/p-ERK/MMP-2/MMP-9.** *Oncol Rep.* 2013;29(5):2046-52.
88. Akkoca AN, Yanik S, Ozdemir ZT, Cihan FG, Sayar S, Cincin TG, et al. **TNM and Modified Dukes staging along with the demographic characteristics of patients with colorectal carcinoma.** *Int J Clin Exp Med.* 2014;7(9):2828-35.
89. Markowitz SD, Bertagnolli MM. **Molecular origins of cancer: Molecular basis of colorectal cancer.** *N Engl J Med.* 2009;361(25):2449-60.
90. Moran A, Ortega P, de Juan C, Fernandez-Marcelo T, Frias C, Sanchez-Pernaute A, et al. **Differential colorectal carcinogenesis: Molecular basis and clinical relevance.** *World J Gastrointest Oncol.* 2010;2(3):151-8.

91. Applegate CC, Lane MA. **Role of retinoids in the prevention and treatment of colorectal cancer.** *World J Gastrointest Oncol.* 2015;7(10):184-203.
92. 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(2):367-80.
93. Papadatos-Pastos D, Rabbie R, Ross P, Sarker D. **The role of the PI3K pathway in colorectal cancer.** *Crit Rev Oncol Hematol.* 2015;94(1):18-30.
94. Chen HM, Fang JY. **Epigenetic Biomarkers for the Early Detection of Gastrointestinal Cancer.** *Gastrointest Tumors.* 2014;1(4):201-8.
95. Hendon SE, DiPalma JA. **U.S. practices for colon cancer screening.** *Keio J Med.* 2005;54(4):179-83.
96. Bretthauer M. **Colorectal cancer screening.** *J Intern Med.* 2011;270(2):87-98.
97. Markowitz SD, Dawson DM, Willis J, Willson JK. **Focus on colon cancer.** *Cancer Cell.* 2002;1(3):233-6.
98. Meyerhardt JA, Mayer RJ. **Systemic therapy for colorectal cancer.** *N Engl J Med.* 2005;352(5):476-87.
99. Kimura Y, Morohashi S, Yoshizawa T, Suzuki T, Morohashi H, Sakamoto Y, et al. **Clinicopathological significance of vascular endothelial growth factor, thymidine phosphorylase and microvessel density in colorectal cancer.** *Mol Med Rep.* 2016;13(2):1551-7.
100. Ganetzky B, Wu CF. **Neurogenetics of membrane excitability in Drosophila.** *Annu Rev Genet.* 1986;20:13-44.

101. Michael MZ, SM OC, van Holst Pellekaan NG, Young GP, James RJ. **Reduced accumulation of specific microRNAs in colorectal neoplasia.** *Mol Cancer Res.* 2003;1(12):882-91.
102. Griffin CA, Hawkins AL, Packer RJ, Rorke LB, Emanuel BS. **Chromosome abnormalities in pediatric brain tumors.** *Cancer Res.* 1988;48(1):175-80.
103. Wu G, Sinclair C, Hinson S, Ingle JN, Roche PC, Couch FJ. **Structural analysis of the 17q22-23 amplicon identifies several independent targets of amplification in breast cancer cell lines and tumors.** *Cancer Res.* 2001;61(13):4951-5.
104. Asangani IA, Rasheed SA, Nikolova DA, Leupold JH, Colburn NH, Post S, et al. **MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdc4 and stimulates invasion, intravasation and metastasis in colorectal cancer.** *Oncogene.* 2008;27(15):2128-36.
105. Vicinus B, Rubie C, Faust SK, Frick VO, Ghadjar P, Wagner M, et al. **miR-21 functionally interacts with the 3'UTR of chemokine CCL20 and down-regulates CCL20 expression in miR-21 transfected colorectal cancer cells.** *Cancer Lett.* 2012;316(1):105-12.
106. Wang P, Zou F, Zhang X, Li H, Dulak A, Tomko RJ, Jr., et al. **microRNA-21 negatively regulates Cdc25A and cell cycle progression in colon cancer cells.** *Cancer Res.* 2009;69(20):8157-65.
107. Roy S, Yu Y, Padhye SB, Sarkar FH, Majumdar AP. **Difluorinated-curcumin (CDF) restores PTEN expression in colon cancer cells by down-regulating miR-21.** *PLoS One.* 2013;8(7):e68543.

108. Qu A, Du L, Yang Y, Liu H, Li J, Wang L, et al. **Hypoxia-inducible MiR-210 is an independent prognostic factor and contributes to metastasis in colorectal cancer.** *PLoS One*. 2014;9(3):e90952.
109. Koga Y, Yasunaga M, Takahashi A, Kuroda J, Moriya Y, Akasu T, et al. **MicroRNA expression profiling of exfoliated colonocytes isolated from feces for colorectal cancer screening.** *Cancer Prev Res (Phila)*. 2010;3(11):1435-42.
110. Feng X, Wang Z, Fillmore R, Xi Y. **MiR-200, a new star miRNA in human cancer.** *Cancer Lett*. 2014;344(2):166-73.
111. Lu YX, Yuan L, Xue XL, Zhou M, Liu Y, Zhang C, et al. **Regulation of colorectal carcinoma stemness, growth, and metastasis by an miR-200c-Sox2-negative feedback loop mechanism.** *Clin Cancer Res*. 2014;20(10):2631-42.
112. Hahn S, Jackstadt R, Siemens H, Hunten S, Hermeking H. **SNAIL and miR-34a feed-forward regulation of ZNF281/ZBP99 promotes epithelial-mesenchymal transition.** *EMBO J*. 2013;32(23):3079-95.
113. Lai MC, Yang Z, Zhou L, Zhu QQ, Xie HY, Zhang F, et al. **Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation.** *Med Oncol*. 2012;29(3):1810-6.
114. Yamada K, Kano J, Tsunoda H, Yoshikawa H, Okubo C, Ishiyama T, et al. **Phenotypic characterization of endometrial stromal sarcoma of the uterus.** *Cancer Sci*. 2006;97(2):106-12.

115. 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(6):851-8.
116. Yang L, Lin C, Liu W, Zhang J, Ohgi KA, Grinstein JD, et al. **ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs.** *Cell*. 2011;147(4):773-88.
117. 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(9):2289-94.
118. Wilusz JE, Freier SM, Spector DL. **3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA.** *Cell*. 2008;135(5):919-32.
119. 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(1):169-75.
120. Guo Q, Zhao Y, Chen J, Hu J, Wang S, Zhang D, et al. **BRAF-activated long non-coding RNA contributes to colorectal cancer migration by inducing epithelial-mesenchymal transition.** *Oncol Lett*. 2014;8(2):869-75.
121. Grisanzio C, Freedman ML. **Chromosome 8q24-Associated Cancers and MYC.** *Genes Cancer*. 2010;1(6):555-9.
122. Haiman CA, Le Marchand L, Yamamoto J, Stram DO, Sheng X, Kolonel LN, et al. **A common genetic risk factor for colorectal and prostate cancer.** *Nat Genet*. 2007;39(8):954-6.

123. Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, Doddapaneni H, et al. **The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer.** *Nat Genet.* 2009;41(8):882-4.
124. Tuupanen S, Turunen M, Lehtonen R, Hallikas O, Vanharanta S, Kivioja T, et al. **The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling.** *Nat Genet.* 2009;41(8):885-90.
125. Ling H, Spizzo R, Atlasi Y, Nicoloso M, Shimizu M, Redis RS, et al. **CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer.** *Genome Res.* 2013;23(9):1446-61.
126. Xiang JF, Yin QF, Chen T, Zhang Y, Zhang XO, Wu Z, et al. **Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus.** *Cell Res.* 2014;24(5):513-31.
127. Beck-Engeser GB, Lum AM, Huppi K, Caplen NJ, Wang BB, Wabl M. **Pvt1-encoded microRNAs in oncogenesis.** *Retrovirology.* 2008;5:4.
128. Tseng YY, Moriarity BS, Gong W, Akiyama R, Tiwari A, Kawakami H, et al. **PVT1 dependence in cancer with MYC copy-number increase.** *Nature.* 2014;512(7512):82-6.
129. 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(4):2509-19.
130. Wu Y, Zhang L, Wang Y, Li H, Ren X, Wei F, et al. **Long noncoding RNA HOTAIR involvement in cancer.** *Tumour Biol.* 2014;35(10):9531-8.
131. Li L, Liu B, Wapinski OL, Tsai MC, Qu K, Zhang J, et al. **Targeted disruption of Hotair leads to homeotic transformation and gene derepression.** *Cell Rep.* 2013;5(1):3-12.

132. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, et al. **Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression.** *Proc Natl Acad Sci U S A.* 2009;106(28):11667-72.
133. Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, et al. **lincRNAs act in the circuitry controlling pluripotency and differentiation.** *Nature.* 2011;477(7364):295-300.
134. Lustig O, Ariel I, Ilan J, Lev-Lehman E, De-Groot N, Hochberg A. **Expression of the imprinted gene H19 in the human fetus.** *Mol Reprod Dev.* 1994;38(3):239-46.
135. Kallen AN, Zhou XB, Xu J, Qiao C, Ma J, Yan L, et al. **The imprinted H19 lincRNA antagonizes let-7 microRNAs.** *Mol Cell.* 2013;52(1):101-12.
136. Rigoutsos I, Lee SK, Nam SY, Anfossi S, Pasculli B, Pichler M, et al. **N-BLR, a primate-specific non-coding transcript leads to colorectal cancer invasion and migration.** *Genome Biol.* 2017;18(1):98.
137. Han Y, Yang YN, Yuan HH, Zhang TT, Sui H, Wei XL, et al. **UCA1, a long non-coding RNA up-regulated in colorectal cancer influences cell proliferation, apoptosis and cell cycle distribution.** *Pathology.* 2014;46(5):396-401.
138. Neve B, Jonckheere N, Vincent A, Van Seuning I. **Epigenetic Regulation by lincRNAs: An Overview Focused on UCA1 in Colorectal Cancer.** *Cancers (Basel).* 2018;10(11).
139. Ni B, Yu X, Guo X, Fan X, Yang Z, Wu P, et al. **Increased urothelial cancer associated 1 is associated with tumor proliferation and metastasis and predicts poor prognosis in colorectal cancer.** *Int J Oncol.* 2015;47(4):1329-38.

140. Wu W, Zhang S, Li X, Xue M, Cao S, Chen W. **Ets-2 regulates cell apoptosis via the Akt pathway, through the regulation of urothelial cancer associated 1, a long non-coding RNA, in bladder cancer cells.** *PLoS One*. 2013;8(9):e73920.
141. Zuo ZK, Gong Y, Chen XH, Ye F, Yin ZM, Gong QN, et al. **TGFbeta1-Induced LncRNA UCA1 Upregulation Promotes Gastric Cancer Invasion and Migration.** *DNA Cell Biol*. 2017;36(2):159-67.
142. Li GY, Wang W, Sun JY, Xin B, Zhang X, Wang T, et al. **Long non-coding RNAs AC026904.1 and UCA1: a "one-two punch" for TGF-beta-induced SNAI2 activation and epithelial-mesenchymal transition in breast cancer.** *Theranostics*. 2018;8(10):2846-61.
143. Gou L, Liu M, Xia J, Wan Q, Jiang Y, Sun S, et al. **BMP9 Promotes the Proliferation and Migration of Bladder Cancer Cells through Up-Regulating lncRNA UCA1.** *Int J Mol Sci*. 2018;19(4).
144. Xue M, Chen W, Xiang A, Wang R, Chen H, Pan J, et al. **Hypoxic exosomes facilitate bladder tumor growth and development through transferring long non-coding RNA-UCA1.** *Mol Cancer*. 2017;16(1):143.
145. Xue M, Li X, Li Z, Chen W. **Urothelial carcinoma associated 1 is a hypoxia-inducible factor-1alpha-targeted long noncoding RNA that enhances hypoxic bladder cancer cell proliferation, migration, and invasion.** *Tumour Biol*. 2014;35(7):6901-12.
146. Barbagallo C, Brex D, Caponnetto A, Cirnigliaro M, Scalia M, Magnano A, et al. **LncRNA UCA1, Upregulated in CRC Biopsies and Downregulated in Serum Exosomes, Controls mRNA Expression by RNA-RNA Interactions.** *Mol Ther Nucleic Acids*. 2018;12:229-41.

147. Nie W, Ge HJ, Yang XQ, Sun X, Huang H, Tao X, et al. **LncRNA-UCA1 exerts oncogenic functions in non-small cell lung cancer by targeting miR-193a-3p.** *Cancer Lett.* 2016;371(1):99-106.
148. Zheng J, Yi D, Liu Y, Wang M, Zhu Y, Shi H. **Long noncoding RNA UCA1 regulates neural stem cell differentiation by controlling miR-1/Hes1 expression.** *Am J Transl Res.* 2017;9(8):3696-704.
149. Luo J, Chen J, Li H, Yang Y, Yun H, Yang S, et al. **LncRNA UCA1 promotes the invasion and EMT of bladder cancer cells by regulating the miR-143/HMGB1 pathway.** *Oncol Lett.* 2017;14(5):5556-62.
150. Wu H, Zhou C. **Long non-coding RNA UCA1 promotes lung cancer cell proliferation and migration via microRNA-193a/HMGB1 axis.** *Biochem Biophys Res Commun.* 2018;496(2):738-45.
151. Thiery JP. **Epithelial-mesenchymal transitions in tumour progression.** *Nat Rev Cancer.* 2002;2(6):442-54.
152. Valastyan S, Weinberg RA. **Tumor metastasis: molecular insights and evolving paradigms.** *Cell.* 2011;147(2):275-92.
153. Weidle UH, Birzele F, Kollmorgen G, Ruger R. **Long Non-coding RNAs and their Role in Metastasis.** *Cancer Genomics Proteomics.* 2017;14(3):143-60.
154. Yeung KT, Yang J. **Epithelial-mesenchymal transition in tumor metastasis.** *Mol Oncol.* 2017;11(1):28-39.
155. Egeblad M, Werb Z. **New functions for the matrix metalloproteinases in cancer progression.** *Nat Rev Cancer.* 2002;2(3):161-74.

156. Jafri MA, Al-Qahtani MH, Shay JW. **Role of miRNAs in human cancer metastasis: Implications for therapeutic intervention.** *Semin Cancer Biol.* 2017;44:117-31.
157. Park SM, Gaur AB, Lengyel E, Peter ME. **The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2.** *Genes Dev.* 2008;22(7):894-907.
158. Zhang P, Hong H, Sun X, Jiang H, Ma S, Zhao S, et al. **MicroRNA-10b regulates epithelial-mesenchymal transition by modulating KLF4/Notch1/E-cadherin in cisplatin-resistant nasopharyngeal carcinoma cells.** *Am J Cancer Res.* 2016;6(2):141-56.
159. Song Y, Li J, Zhu Y, Dai Y, Zeng T, Liu L, et al. **MicroRNA-9 promotes tumor metastasis via repressing E-cadherin in esophageal squamous cell carcinoma.** *Oncotarget.* 2014;5(22):11669-80.
160. Zaravinos A. **The Regulatory Role of MicroRNAs in EMT and Cancer.** *J Oncol.* 2015;2015:865816.
161. Tang R, Liang L, Luo D, Feng Z, Huang Q, He R, et al. **Downregulation of MiR-30a is Associated with Poor Prognosis in Lung Cancer.** *Med Sci Monit.* 2015;21:2514-20.
162. Kumarswamy R, Mudduluru G, Ceppi P, Muppala S, Kozlowski M, Niklinski J, et al. **MicroRNA-30a inhibits epithelial-to-mesenchymal transition by targeting Snai1 and is downregulated in non-small cell lung cancer.** *Int J Cancer.* 2012;130(9):2044-53.
163. Hao Y, Yang X, Zhang D, Luo J, Chen R. **Long noncoding RNA LINC01186, regulated by TGF-beta/SMAD3, inhibits migration and invasion through Epithelial-Mesenchymal-Transition in lung cancer.** *Gene.* 2017;608:1-12.

164. Zhang M, Wu WB, Wang ZW, Wang XH. **lncRNA NEAT1 is closely related with progression of breast cancer via promoting proliferation and EMT.** *Eur Rev Med Pharmacol Sci.* 2017;21(5):1020-6.
165. Rao DD, Vorhies JS, Senzer N, Nemunaitis J. **siRNA vs. shRNA: similarities and differences.** *Adv Drug Deliv Rev.* 2009;61(9):746-59.
166. Watrin M, Dausse E, Lebars I, Rayner B, Bugaut A, Toulme JJ. **Aptamers targeting RNA molecules.** *Methods Mol Biol.* 2009;535:79-105.
167. Xu J, Lamouille S, Derynck R. **TGF-beta-induced epithelial to mesenchymal transition.** *Cell Res.* 2009;19(2):156-72.
168. Yadav A, Kumar B, Datta J, Teknos TN, Kumar P. **IL-6 promotes head and neck tumor metastasis by inducing epithelial-mesenchymal transition via the JAK-STAT3-SNAIL signaling pathway.** *Mol Cancer Res.* 2011;9(12):1658-67.
169. Nieto MA. **Epithelial plasticity: a common theme in embryonic and cancer cells.** *Science.* 2013;342(6159):1234850.
170. Katsuno Y, Lamouille S, Derynck R. **TGF-beta signaling and epithelial-mesenchymal transition in cancer progression.** *Curr Opin Oncol.* 2013;25(1):76-84.
171. Siegel R, Ma J, Zou Z, Jemal A. **Cancer statistics, 2014.** *CA Cancer J Clin.* 2014;64(1):9-29.
172. Kalluri R, Weinberg RA. **The basics of epithelial-mesenchymal transition.** *J Clin Invest.* 2009;119(6):1420-8.
173. Thiery JP, Acloque H, Huang RY, Nieto MA. **Epithelial-mesenchymal transitions in development and disease.** *Cell.* 2009;139(5):871-90.

174. Gao X, Wen J, Gao P, Zhang G, Zhang G. **Overexpression of the long non-coding RNA, linc-UBC1, is associated with poor prognosis and facilitates cell proliferation, migration, and invasion in colorectal cancer.** *Onco Targets Ther.* 2017;10:1017-26.
175. Yang QS, Li B, Xu G, Yang SQ, Wang P, Tang HH, et al. **Long noncoding RNA LINC00483/microRNA-144 regulates radiosensitivity and epithelial-mesenchymal transition in lung adenocarcinoma by interacting with HOXA10.** *J Cell Physiol.* 2019;234(7):11805-21.
176. Mao L, Li Y, Zhao J, Li Q, Yang B, Wang Y, et al. **Transforming growth factor-beta1 contributes to oxaliplatin resistance in colorectal cancer via epithelial to mesenchymal transition.** *Oncol Lett.* 2017;14(1):647-54.
177. Knickelbein K, Zhang L. **Mutant KRAS as a critical determinant of the therapeutic response of colorectal cancer.** *Genes Dis.* 2015;2(1):4-12.
178. Zenonos K, Kyprianou K. **RAS signaling pathways, mutations and their role in colorectal cancer.** *World J Gastrointest Oncol.* 2013;5(5):97-101.
179. Li D, Yang M, Liao A, Zeng B, Liu D, Yao Y, et al. **Linc00483 as ceRNA regulates proliferation and apoptosis through activating MAPKs in gastric cancer.** *J Cell Mol Med.* 2018.
180. Tang J, Zhuo H, Zhang X, Jiang R, Ji J, Deng L, et al. **A novel biomarker Linc00974 interacting with KRT19 promotes proliferation and metastasis in hepatocellular carcinoma.** *Cell Death Dis.* 2014;5:e1549.

181. Yan S, Yue Y, Wang J, Li W, Sun M, Gu C, et al. **LINC00668 promotes tumorigenesis and progression through sponging miR-188-5p and regulating USP47 in colorectal cancer.** *Eur J Pharmacol.* 2019;858:172464.
182. Yuan JH, Yang F, Wang F, Ma JZ, Guo YJ, Tao QF, et al. **A long noncoding RNA activated by TGF-beta promotes the invasion-metastasis cascade in hepatocellular carcinoma.** *Cancer Cell.* 2014;25(5):666-81.
183. Gyamfi J, Lee YH, Min BS, Choi J. **Niclosamide reverses adipocyte induced epithelial-mesenchymal transition in breast cancer cells via suppression of the interleukin-6/STAT3 signalling axis.** *Sci Rep.* 2019;9(1):11336.
184. Wang C, Dou C, Wang Y, Liu Z, Roberts L, Zheng X. **TLX3 repressed SNAI1-induced epithelial-mesenchymal transition by directly constraining STAT3 phosphorylation and functionally sensitized 5-FU chemotherapy in hepatocellular carcinoma.** *Int J Biol Sci.* 2019;15(8):1696-711.
185. Wang F, Liang R, Soibam B, Yang J, Liu Y. **Coregulatory long non-coding RNA and protein-coding genes in serum starved cells.** *Biochim Biophys Acta Gene Regul Mech.* 2019;1862(1):84-95.
186. Ameri K, Jahangiri A, Rajah AM, Tormos KV, Nagarajan R, Pekmezci M, et al. **HIGD1A Regulates Oxygen Consumption, ROS Production, and AMPK Activity during Glucose Deprivation to Modulate Cell Survival and Tumor Growth.** *Cell Rep.* 2015;10(6):891-9.

187. Peng M, Hu Y, Song W, Duan S, Xu Q, Ding Y, et al. **MIER3 suppresses colorectal cancer progression by down-regulating Sp1, inhibiting epithelial-mesenchymal transition.** *Sci Rep.* 2017;7(1):11000.
188. Egan JB, Thompson PA, Ashbeck EL, Conti DV, Duggan D, Hibler E, et al. **Genetic polymorphisms in vitamin D receptor VDR/RXRA influence the likelihood of colon adenoma recurrence.** *Cancer Res.* 2010;70(4):1496-504.
189. Larriba MJ, Garcia de Herreros A, Munoz A. **Vitamin D and the Epithelial to Mesenchymal Transition.** *Stem Cells Int.* 2016;2016:6213872.
190. Vandewalle B, Adenis A, Hornez L, Revillion F, Lefebvre J. **1,25-dihydroxyvitamin D3 receptors in normal and malignant human colorectal tissues.** *Cancer Lett.* 1994;86(1):67-73.
191. Evans SR, Nolla J, Hanfelt J, Shabahang M, Nauta RJ, Shchepotin IB. **Vitamin D receptor expression as a predictive marker of biological behavior in human colorectal cancer.** *Clin Cancer Res.* 1998;4(7):1591-5.
192. Brozyna AA, Jozwicki W, Janjetovic Z, Slominski AT. **Expression of vitamin D receptor decreases during progression of pigmented skin lesions.** *Hum Pathol.* 2011;42(5):618-31.
193. Anderson MG, Nakane M, Ruan X, Kroeger PE, Wu-Wong JR. **Expression of VDR and CYP24A1 mRNA in human tumors.** *Cancer Chemother Pharmacol.* 2006;57(2):234-40.
194. Lopes N, Sousa B, Martins D, Gomes M, Vieira D, Veronese LA, et al. **Alterations in Vitamin D signalling and metabolic pathways in breast cancer progression: a study of**

VDR, CYP27B1 and CYP24A1 expression in benign and malignant breast lesions. *BMC Cancer*. 2010;10:483.

195. Thill M, Fischer D, Kelling K, Hoellen F, Dittmer C, Hornemann A, et al. **Expression of vitamin D receptor (VDR), cyclooxygenase-2 (COX-2) and 15-hydroxyprostaglandin dehydrogenase (15-PGDH) in benign and malignant ovarian tissue and 25-hydroxycholecalciferol (25(OH)D3) and prostaglandin E2 (PGE2) serum level in ovarian cancer patients.** *J Steroid Biochem Mol Biol*. 2010;121(1-2):387-90.

196. Larriba MJ, Martin-Villar E, Garcia JM, Pereira F, Pena C, de Herreros AG, et al. **Snail2 cooperates with Snail1 in the repression of vitamin D receptor in colon cancer.** *Carcinogenesis*. 2009;30(8):1459-68.

197. Bauer J, Ozden O, Akagi N, Carroll T, Principe DR, Staudacher JJ, et al. **Activin and TGFbeta use diverging mitogenic signaling in advanced colon cancer.** *Mol Cancer*. 2015;14:182.

198. Bauer J, Sporn JC, Cabral J, Gomez J, Jung B. **Effects of activin and TGFbeta on p21 in colon cancer.** *PLoS One*. 2012;7(6):e39381.

199. Takeda H, Kataoka S, Nakayama M, Ali MAE, Oshima H, Yamamoto D, et al. **CRISPR-Cas9-mediated gene knockout in intestinal tumor organoids provides functional validation for colorectal cancer driver genes.** *Proc Natl Acad Sci U S A*. 2019;116(31):15635-44.

200. Liu M, Casimiro MC, Wang C, Shirley LA, Jiao X, Katiyar S, et al. **p21CIP1 attenuates Ras- and c-Myc-dependent breast tumor epithelial mesenchymal transition**

- and cancer stem cell-like gene expression in vivo.** *Proc Natl Acad Sci U S A.* 2009;106(45):19035-9.
201. Vite A, Li J, Radice GL. **New functions for alpha-catenins in health and disease: from cancer to heart regeneration.** *Cell Tissue Res.* 2015;360(3):773-83.
202. Piao HL, Yuan Y, Wang M, Sun Y, Liang H, Ma L. **alpha-catenin acts as a tumour suppressor in E-cadherin-negative basal-like breast cancer by inhibiting NF-kappaB signalling.** *Nat Cell Biol.* 2014;16(3):245-54.
203. Zhou YN, Xu CP, Chen Y, Han B, Yang SM, Fang DC. **Alpha-catenin expression is decreased in patients with gastric carcinoma.** *World J Gastroenterol.* 2005;11(22):3468-72.
204. Chen X, Zhu H, Wu X, Xie X, Huang G, Xu X, et al. **Downregulated pseudogene CTNNAP1 promote tumor growth in human cancer by downregulating its cognate gene CTNNA1 expression.** *Oncotarget.* 2016;7(34):55518-28.
205. Wu H, He Y, Chen H, Liu Y, Wei B, Chen G, et al. **LncRNA THOR increases osteosarcoma cell stemness and migration by enhancing SOX9 mRNA stability.** *FEBS Open Bio.* 2019;9(4):781-90.
206. Song H, Xu Y, Shi L, Xu T, Fan R, Cao M, et al. **LncRNA THOR increases the stemness of gastric cancer cells via enhancing SOX9 mRNA stability.** *Biomed Pharmacother.* 2018;108:338-46.
207. Arends JW, Wiggers T, Thijs CT, Verstijnen C, Swaen GJ, Bosman FT. **The value of secretory component (SC) immunoreactivity in diagnosis and prognosis of colorectal carcinomas.** *Am J Clin Pathol.* 1984;82(3):267-74.

208. Isaacson P. **Immunoperoxidase study of the secretory immunoglobulin system in colonic neoplasia.** *J Clin Pathol.* 1982;35(1):14-25.
209. Koretz K, Schlag P, Quentmeier A, Moller P. **Evaluation of the secretory component as a prognostic variable in colorectal carcinoma.** *Int J Cancer.* 1994;57(3):365-70.
210. Traicoff JL, De Marchis L, Ginsburg BL, Zamora RE, Khattar NH, Blanch VJ, et al. **Characterization of the human polymeric immunoglobulin receptor (PIGR) 3'UTR and differential expression of PIGR mRNA during colon tumorigenesis.** *J Biomed Sci.* 2003;10(6 Pt 2):792-804.
211. Veto B, Bojcsuk D, Bacquet C, Kiss J, Sipeki S, Martin L, et al. **The transcriptional activity of hepatocyte nuclear factor 4 alpha is inhibited via phosphorylation by ERK1/2.** *PLoS One.* 2017;12(2):e0172020.