

UNIVERSITA' DEGLI STUDI DI  
GENOVA



CORSO DI LAUREA in BIOTECNOLOGIE

TESI DI LAUREA

*Studio della regolazione dell'espressione  
del gene ACVR1 a livello post-  
trascrizionale.*

Relatore:

Chiar.mo Prof. Roberto Ravazzolo

Candidato:

Serena Cappato

ANNO ACCADEMICO 2008-2009

## **PREMESSA**

Ho svolto la mia attività di tirocinio presso il Laboratorio di Genetica Molecolare dell'Istituto Giannina Gaslini, partecipando al progetto che riguarda lo studio di *ACVRI*, gene implicato nella Fibrodiplosia Ossificante Progressiva (FOP), una rara e severa malattia genetica ad ereditarietà autosomica dominante che causa ossificazioni eterotopiche.

## **RINGRAZIAMENTI**

Ringrazio il Prof. Roberto Ravazzolo per avermi dato la possibilità di fare tirocinio presso il suo laboratorio, la Dott.ssa Renata Bocciardi per avermi seguito in questo anno con pazienza e dedizione, le Dott.sse Marzia Mura, Monica Marini, Francesca Giacomelli e Lorella Covone per i loro preziosi consigli e l'*Associazione FOP Italia*, per l'importante collaborazione e il continuo incoraggiamento.

Dedico questo primo traguardo alla mia famiglia che mi ha sostenuto in questi anni e a Walter.

# INDICE

<b>1 INTRODUZIONE</b>	<b>pag. 4</b>
1.1 Il gene <i>ACVRI</i> e il suo prodotto proteico	pag. 5
1.2 La FOP e <i>ACVRI</i>	pag. 7
1.3 Le regioni "trascritte ma non tradotte" all'estremità 3'	pag. 9
1.4 I microRNA	pag.10
1.5 Le sequenze ricche in AU (ARE)	pag.13
<b>2 SCOPO</b>	<b>pag.15</b>
<b>3 RISULTATI</b>	<b>pag. 17</b>
Struttura di <i>ACVRI</i>	pag. 18
Il 3'UTR di <i>ACVRI</i>	pag. 28
- Siti di poliadenilazione	pag. 28
- Siti putativi di legame per i microRNA	pag. 29
- Le sequenze ricche in AU	pag. 31
Studi funzionali sul 3'UTR	pag. 32
- Analisi della stabilità del trascritto di <i>ACVRI</i>	pag. 35
- Azione dei microRNA sul 3'UTR di <i>ACVRI</i>	pag. 37
<b>4 CONCLUSIONI</b>	<b>pag. 39</b>
<b>5 MATERIALI E METODI</b>	<b>pag. 41</b>
<b>6 BIBLIOGRAFIA</b>	<b>pag. 54</b>

## 1 INTRODUZIONE

## 1.1 Il gene *ACVRI* e il suo prodotto proteico.

IL gene *ACVRI* (Activin A receptor, type I) conosciuto anche come *ALK2* (Activin receptor-like kinase 2) è localizzato sul cromosoma 2 *locus* q23- q24 e codifica per un recettore di tipo I delle proteine morfogenetiche dell'osso (BMPs) (Shore et al., 2006). Si tratta di una proteina di 509 aminoacidi altamente conservata durante l'evoluzione dei vertebrati, appartenente alla famiglia dei recettori del "Trasforming Growth Factor  $\beta$ " (TGF- $\beta$ ), composta da 7 membri (da ALK1 ad ALK7) (per una revisione Graham et al., 2006).

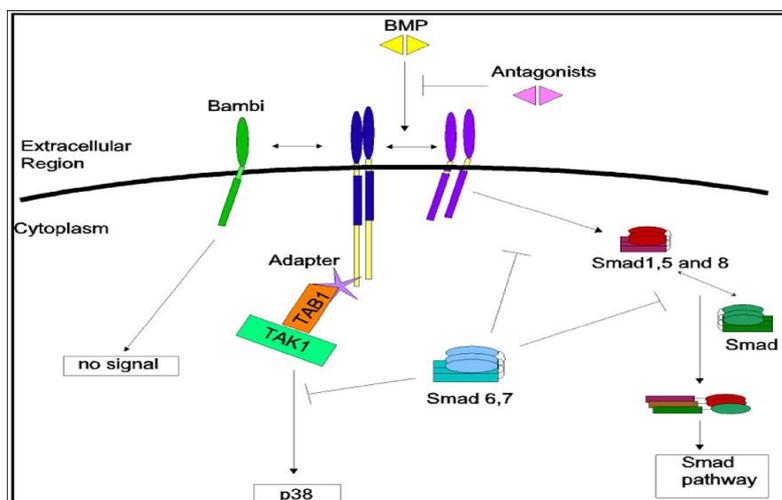
I recettori di tipo I sono presenti sulla superficie cellulare associati ai recettori di tipo II con i quali interagiscono nell'avvio della cascata di segnalazione interna. Entrambi i tipi recettoriali condividono l'organizzazione strutturale, costituita da un dominio extracellulare, di interazione con il ligando, un dominio transmembrana e un dominio citoplasmatico con funzione di Serina/Treonina chinasi. I recettori di tipo I, come ACVR1/ALK2, presentano un ulteriore dominio con funzione regolatoria definito dominio GS per la struttura ricca in residui di Glicina e Serina, localizzato nella porzione intracellulare prossima alla membrana e molto conservato durante l'evoluzione (per una revisione Shi et al., 2003). L'attivazione del complesso recettoriale indotta dal ligando porta alla fosforilazione del dominio GS dei recettori di tipo I da parte delle molecole di tipo II avviando la cascata di segnalazione interna. Il dominio GS è quindi funzionalmente molto importante e interviene sia nel mantenimento del recettore in stato inattivo, in presenza di concentrazioni subottimali di ligando, che al momento del suo spegnimento (Shi et al., 2003).

Come schematizzato in Figura 1, i recettori delle BMP possono essere coinvolti in differenti meccanismi di segnalazione intracellulare andando ad interagire con diverse vie di segnalazione intracellulare, tra cui quella delle Smad coinvolta nel differenziamento cellulare e quella di p38 implicata nell'avvio dell'apoptosi (Nohe et al., 2003).

Nella prima via dopo il legame delle BMP al complesso tetrameric preformato ("Preformed Hetero-oligomeric complex", PFC) dei recettori di tipo II con i recettori di tipo I, segue la fosforilazione dei recettori di tipo I a livello del dominio GS da parte del recettore di tipo II, questo consente di reclutare le proteine receptor-regulated-SMAD (R-Smad, Smad1,5 o 8) che in seguito a

fosforilazione formano un complesso eterodimerico con la Common Mediator Smad (CO-Smad, Smad4). che migra nel nucleo e va a regolare la trascrizione di specifici geni bersaglio. Nel nucleo le proteine Smad presentano un'attività trascrizionale legando il DNA direttamente o interagendo con proteine che lo legano a loro volta. (Nohe et al., 2003)

La via alternativa a quella appena descritta dipende da una differente interazione dei recettori di tipo I e II a livello della membrana cellulare. In questo caso i recettori si presentano come un complesso dimerico; il ligando viene reclutato da parte dei recettori di tipo II che in seguito a questo legame, reclutano i recettori di tipo I formando il Complesso di Segnalazione Indotto dalle BMP ("BMP induced signaling complex", BISC). La cascata di segnalazione interna è mediata da TAB1 ("Tak binding protein" ) che forma un complesso con TAK1, una MAP cinasi cinasi, ("Mitogen Activated Protein Kinase kinase", MAPKK) e questo determina l'attivazione di p38 coinvolto nel processo di apoptosi cellulare (Nohe et al., 2003).



**Figura 1** Rappresentazione schematica della via di segnalazione mediata dalle BMP.

La via di segnalazione delle BMP è regolata in modo molto preciso e a diversi livelli. A livello extra-cellulare sono noti diversi antagonisti specifici che possono legare le BMP sequestrandole dal mezzo e impedendo il legame con il complesso recettoriale come Noggin, Gremlin, Chordin e Follistatin (Nohe et al., 2003). Un controllo "alla membrana" può essere svolto da pseudorecettori

che svolgono una funzione tampone, legando e sequestrando il ligando, ma senza trasdurre alcun segnale all'interno della cellula, trattandosi di recettori sprovvisti del dominio enzimatico; un esempio è Bambi ("BMP and activin membrane-bound inhibitor") una glicoproteina transmembrana evolutivamente conservata nei vertebrati, dai pesci all'uomo, e strettamente relazionata a recettori di tipo I della famiglia dei TGF $\beta$  nel dominio extracellulare nonostante a livello intracellulare sia priva del dominio cinasico (Figura 1). La sua funzione è quella di interferire con la formazione del complesso tra recettori di tipo I e II e cooperare con SMAD7 per inibire la segnalazione interna (Yan et al., 2009).

Come già menzionato sopra, a livello recettoriale, il dominio GS ha un ruolo cruciale per i recettori di tipo I. Questa regione interagisce fisicamente con la proteina citoplasmatica legante FK506 ("FK506 binding protein 12", FKBP12) che, legando il dominio GS dei recettori BMP di tipo I aiuta a mantenere il recettore inattivo, impedendo la segnalazione in assenza del ligando. La fosforilazione ligando dipendente del dominio GS, porta ad una modificazione conformazionale che causa il rilascio di FKBP12 e l'attivazione della cascata di segnalazione intracellulare. La riassociazione successiva riporta all'estinzione dell'attivazione recettoriale (Wang et al., 2004).

A livello intracellulare, la cascata di segnalazione può essere controllata grazie all'intervento delle SMAD con funzione inibitoria (i-SMAD), quali SMAD6 e SMAD7 o a proteine che legano le R-SMAD attivate e che portano alla terminazione del segnale ligando-dipendente (Nohe et al., 2003).

## **1.2 La FOP e *ACVRI***

Il gene *ACVRI* risulta mutato nella Fibrodisplasia Ossificante Progressiva (FOP, MIM 135100). La FOP, rara malattia genetica autosomica dominante, è una delle forme più gravi di ossificazione eterotopica progressiva. Gli individui affetti appaiono normali alla nascita se si fa eccezione per la presenza di una caratteristica malformazione dell'alluce. La malattia esordisce quindi nell'infanzia, con andamento episodico e progressivo, con fasi acute alternate a fasi di remissione che possono durare anche anni. Le fasi acute si caratterizzano per la comparsa di foci infiammatori dolenti che

progressivamente ossificano con un processo di tipo endocondrale e portano alla neoformazione di osso, con caratteristiche anatomo-istologiche normali, ma in sedi inappropriate. Si ha infatti una progressiva e permanente trasformazione dei tessuti connettivi molli (aponeurosi, fasci, legamenti, tendini e muscoli scheletrici) in osso ectopico, con congelamento progressivo delle articolazioni dello scheletro assiale e appendicolare e in definitiva un secondo scheletro che intrappola il paziente rendendo impossibile il movimento. Le fasi acute insorgono in seguito a stimoli diversi quali piccoli e grossi traumi, iniezioni intramuscolari, manovre medico-chirurgiche e odontoiatriche, affaticamento muscolare, e infezioni virali o, in certi casi senza alcuna causa apparente. L'età media di sopravvivenza è di circa 40 anni e la morte spesso si ha per complicazioni dovute alla progressiva insufficienza toracica.

Ad oggi, non esiste una cura per la FOP se non trattamenti sintomatici con anti-infiammatori e cortisonici. La diagnosi precoce e tempestiva permette di evitare tutte le manovre mediche e non, e di modificare quelle abitudini quotidiane che potrebbero aumentare il rischio di traumi o situazioni che potrebbero portare ad esacerbare la malattia e all'insorgenza di nuove fasi acute.

*ACVRI* risulta ad oggi mutato nella totalità dei pazienti con diagnosi clinica di FOP. La mutazione di gran lunga più frequente è la R206H che determina la sostituzione della Guanina con l'Adenina nel nucleotide 617 (c.617G>A, R206H). Il codone 206 si trova alla fine del dominio di attivazione GS.

La mutazione R206H in *ACVRI* altera la segnalazione delle BMP nella FOP, provocando un'iperattivazione del recettore che, da una parte segnala anche in assenza di ligando, dall'altra sembra essere ad esso iper-responsivo (Kaplan et al., 2007). Oltre alla mutazione classica ne sono state identificate altre a livello del dominio GS e di quello tirosin cinasico responsabili in entrambe i casi di un'attivazione incontrollata della cascata di segnalazione (Kaplan et al., 2008; Bocciardi et al., 2008). Un esempio tra queste mutazioni è la R258S (c.774G>C) che determina la sostituzione di un'Arginina con una Serina nel dominio tirosin cinasico del recettore *ACVRI* (Bocciardi et al., 2008); R258S è stata individuata per la prima volta in tre pazienti italiani dal gruppo in cui sono inserita.

### **1.3 Le regioni "trascritte ma non tradotte" all'estremità 3'.**

Gli RNA messaggeri (mRNA) occupano una posizione fondamentale nel flusso di informazioni che dal DNA arriva alle proteine.

Gli mRNA eucariotici sono costituiti da esoni, introni e regioni trascritte ma non tradotte ("Untranslated Transcribed Region", UTR) agli estremi 5' e al 3'. Vengono sintetizzati nel nucleo come pre-messaggeri e sottoposti ad un processo di maturazione che prevede l'aggiunta di una speciale base modificata all'estremità 5', chiamata CAP e di una coda di Adenine, detta coda di poli-A al 3', e quindi un processamento volto all'eliminazione delle sequenze introniche ("Splicing"). I messaggeri maturi vengono attivamente trasportati nel citoplasma dove possono essere tradotti in proteine.

La sintesi dei messaggeri, ovvero la trascrizione che porta all'espressione genica è un processo altamente regolato che prevede l'interazione specifica e dinamica tra il DNA delle sequenze regolatorie (regioni promotrici ma anche modulatrici negative e positive) e diversi tipi di fattori e cofattori trascrizionali. Alcuni di questi sono comuni perché appartenenti al macchinario trascrizionale di base, altri molto specifici, legati a stimoli, tessuti o stadi ben precisi.

Il controllo della trascrizione è quindi un evento molto complesso integrato in modo da consentire alla cellula di rispondere opportunamente ai segnali e agli stimoli ricevuti. Una volta sintetizzato e maturato, tuttavia, il messaggero può essere ancora l'oggetto di tutta una serie di processi di controllo e modulazione altrettanto complessi di quelli che avvengono a livello trascrizionale.

Durante la loro vita, i messaggeri sono "scortati" da tutta una serie di fattori che si associano ad essi sia in maniera stabile che in modo più dinamico. Questo complemento di fattori proteici e piccoli RNA non codificanti formano con il messaggero una particella ribonucleoproteica, ed è proprio la combinazione unica dei fattori che accompagnano il singolo messaggero e la posizione relativa che questi occupano lungo la sua molecola che poi ne andranno a determinare il destino nel citoplasma (per una revisione vedere Moore et al., 2005). Tra questi fattori solo un piccolo numero lega le strutture comuni a praticamente tutti i messaggeri, quali le proteine leganti il CAP al 5' o il poli-A, oppure le proteine di "impacchettamento" del messaggero che possono legarlo in modo piuttosto aspecifico in tutta la sua lunghezza. La

maggior parte dei fattori che legano gli mRNA, in realtà interagiscono solo con strutture e/o sequenze presenti in un messaggero e non in altri. Si tratta di interazioni assolutamente specifiche, mediate da elementi di riconoscimento che sono spesso localizzati nelle sequenze 5' e 3'UTR dei singoli messaggeri (Moore et al., 2005; Andreassi et al., 2009).

Queste componenti possono essere quindi viste come degli adattatori funzionali che permettono all'mRNA di interfacciarsi con i vari macchinari cellulari che ne determineranno la localizzazione subcellulare, l'efficienza di traduzione e la stabilità (Andreassi et al., 2009).

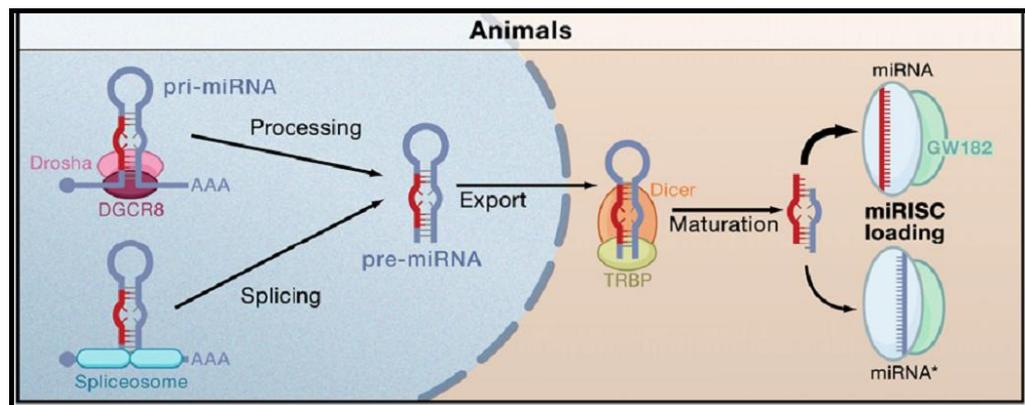
Ad esempio, il 5'UTR è maggiormente coinvolto nel controllo iniziale della traduzione dell'RNA messaggero, mentre il 3'UTR può essere oggetto di diversi meccanismi di regolazione coinvolti in numerosi aspetti del metabolismo dell' mRNA, incluso l'esportazione nucleare, la localizzazione citoplasmatica, l'efficienza di traduzione e la sua stabilità (Andreassi et al., 2009).

#### **1.4 I microRNA**

I microRNA (miRNA) sono un'ampia classe di piccoli RNA non codificanti (21-25 nucleotidi) che hanno un importante ruolo nelle regolazione positiva e negativa dell'espressione genica post-trascrizionale, legandosi a sequenze di riconoscimento specifiche contenute prevalentemente nelle regioni 3'UTR di molti messaggeri (per una revisione Carthew et al., 2009).

I geni per i miRNA sono generalmente trascritti nel nucleo dall'RNA Polimerasi II. Ai filamenti generati viene aggiunto il CAP all'estremità 5' e la coda di poliA all'estremità 3'. Sebbene alcuni miRNA siano prodotti da unità trascrizionali singole, la maggior parte vengono generati da unità trascrizionali che generano contemporaneamente più di un prodotto. Un trascritto può codificare infatti distinti miRNA o codificare miRNA e proteine (Carthew et al., 2009). Il miRNA maturo (Figura 2) deriva da due reazioni sequenziali di processamento del precursore o pri-miRNA. Un tipico pri-miRNA è formato da un filamento appaiato in modo imperfetto con un ansa terminale e due segmenti laterali, a ricordare una sorta di forcina (Bartel et al., 2004). Questi vengono processati a livello nucleare da un membro della famiglia delle RNAsi

di tipo III, DROSHA e dal suo cofattore DGCR-8 che costituiscono il "Microprocessor Complex" per formare un trascritto di circa 70 nucleotidi, il pre-miRNA che assume la forma di una forcina imperfetta (Denli et al., 2004). Il taglio mediato dal "Microprocessor Complex" non è il solo meccanismo di produzione dei pre-miRNA, un evento alternativo è lo "Splicing" del trascritto per liberare gli introni che caratterizzano anche i miRNA (Okamura et al., 2008).



**Figura 2** Meccanismo di sintesi e maturazione dei microRNA.

Il pre-miRNA viene attivamente trasportato dal nucleo al citoplasma trasportatore della famiglia delle RAN-GTPasi Esportina 5 (exp5) (Lund et al., 2004).

Il secondo taglio avviene ad opera dell'enzima Dicer che elimina l'ansa terminale e porta alla formazione di un piccolo RNA a doppio filamento di circa 22 bp di lunghezza costituito da quello che diventerà il miRNA maturo e il suo complementare (Bartel 2004).

L'ultima fase della formazione dei microRNA prevede l'intervento di RISC ("RNA induced-silencing complex"), un complesso multiproteico che favorisce il silenziamento dell'RNA. Sebbene i microRNA maturi possano risiedere su entrambi i filamenti dell'RNA duplex, è stato osservato che l'incorporazione preferenziale in RISC di uno rispetto all'altro, è principalmente dovuta all'instabilità relativa dell'estremità 5', ovvero viene "scelta" come microRNA attivo e maturo, la molecola dotata di minore stabilità alla sua estremità 5'. Sono dunque le proprietà termodinamiche del microRNA precursore a

determinare l'assemblaggio asimmetrico di RISC e di conseguenza, la specificità del mRNA bersaglio durante il meccanismo di inibizione post-trascrizionale. Nei rari casi in cui i due filamenti complementari hanno una stabilità simile al 5', entrambe i componenti del precursore possono assemblarsi al complesso RISC (He et al., 2004).

Con poche eccezioni i siti di legame per i miRNA si trovano al 3'UTR e sono generalmente presenti in copie multiple. Sebbene la caratteristica chiave del riconoscimento coinvolga l'appaiamento delle basi secondo Watson & Crick la maggior parte dei miRNA si lega con un appaiamento corretto per i primi 7-8 nucleotidi che si trovano al 5' del microRNA, tratto conosciuto come seme o core ("seed sequence") mentre per la restante sequenza si trovano numerosi disappaiamenti.

Il grado di complementarietà miRNA-mRNA è molto importante nel meccanismo regolatorio. La complementarietà perfetta permette alle proteine Argonata del complesso RISC di catalizzare il taglio del mRNA, mentre la presenza di disappaiamenti impedisce il taglio e promuove la repressione della traduzione. Probabilmente miRISC promuove un distacco prematuro dei ribosomi dal mRNA indirizzandolo verso la naturale degradazione del trascritto (Carthew et al., 2009).

Negli ultimi anni è stata dimostrata una funzione più ampia dei miRNA nel regolare l'espressione dei geni. Questi piccoli filamenti di RNA oscillano tra la repressione e l'attivazione in modo coordinato con il ciclo cellulare. Nelle cellule proliferanti prevale la repressione della traduzione, mentre nella fase G<sub>1</sub>/G<sub>0</sub>, che spesso precede il differenziamento mediano l'attivazione. Questo definisce un nuovo ruolo dei miRNA e dimostra un'ampia versatilità delle micro Ribonucleoproteine (microRNP) in risposta al ciclo cellulare, con importanti implicazioni per comprendere il contributo di questi RNA nello sviluppo, nella differenziazione cellulare e nella carcinogenesi (Vasudevan et al., 2007).

## 1.5 Le sequenze ricche in AU (ARE)

La regolazione della traduzione e della stabilità dell'RNA messaggero sono essenziali nel controllo dell'espressione genica. Numerosi studi hanno identificato specifici elementi in cis presenti all'interno del 3'UTR di molti trascritti, importanti nella regolazione della stabilità del messaggero (per una revisione Eberhart et al., 2007). Gli elementi *cis*-regolatori dell'mRNA più comuni e diffusi sono brevi sequenze ricche in Adenina e Uracile, chiamate elementi ricchi in AU (AU-rich Elements, ARE).

I motivi delle sequenze che formano le ARE, identificate all'interno del 3'UTR sono stati organizzati in tre classi, in base al numero e alla distribuzione dei pentameri AUUUA:

- la prima classe contiene copie multiple del motivo AUUUA disposte come ripetizioni in tandem;
- la classe II contiene solo alcune ripetizioni del motivo pentamerico AUUUA disperse nel 3'UTR e legate a sequenze ricche in uracile;
- la classe III non contiene motivi pentamerici AUUUA ma altri elementi *cis*-regolatori come ad esempio i motivi nonamerici UUAAUUUA(U/A) (U/A) o UUAUUUAUU (Eberhart et al., 2007).

Il meccanismo di degradazione nella maggior parte dei mRNA inizia con una deadenilazione esonucleolitica 3'-5' della coda di poliA seguito da una rapida degradazione del restante trascritto da parte di esonucleasi 3'-5' o in alcuni casi dall'azione di endonucleasi. Il taglio esonucleolitico è guidato dall'esosoma, un grosso complesso multiproteico che contiene differenti esonucleasi e che costituisce il principale meccanismo di degradazione per i mRNA contenenti le sequenze ARE. La degradazione di questi messaggeri può anche derivare dall'intervento di proteine che marciano il messaggero contenente le ARE e lo segnalano per la degradazione. Le "proteine che legano le sequenze ARE" (ARE binding protein, ARE-BP) sono numerose, ricordiamo ad esempio membri della famiglia HU, differenti proteine hnRNP, fattori di degradazione leganti ARE/poly-(U) (AUF), tristetraprolina (TTP), "proteina da shock termico 70" (heat shock protein 70, Hsp70) e molte altre (Eberart et al., 2007). La stabilità o meno del messaggero sarà quindi il risultato dell'interazione di questi differenti fattori che legano le ARE localizzate sul trascritto con altri

complessi proteici presenti all'interno della cellula; ad esempio TTP stimola la deadenilasi [poly(A)] ribonucleasi, promuovendo la deadenilazione e la successiva demolizione del filamento di RNA (Gherzi et al., 2004).

La regolazione della stabilità del messaggero attraverso le ARE permette una risposta rapida e precisa verso segnali extra e intracellulari in seguito al cambiamento delle condizioni ambientali; a questo fine i messaggeri che contengono le ARE codificano per una varietà di proteine strutturali e regolatorie con una miriade di funzioni differenti (Eberart et al., 2007).

Le sequenze ARE sono state identificate nel 3'UTR di numerosi mRNA labili per esempio quelli che codificano per citochine, proto-oncogeni, fattori di crescita, linfocine, attivatori trascrizionali, proteine regolatorie del ciclo cellulare e diversi mediatori infiammatori (Eberart et al., 2007).

Comune alla maggior parte dei geni che trascrivono messaggeri contenenti ARE sono il gran numero di stimoli che regolano la stabilità del messaggero. Questi possono essere classificati in differenti categorie: citochine fattori di crescita, mediatori cellulari dello stress cellulare, ormoni e agenti vasoattivi; oltre a questi fattori interni ci sono molti agenti ambientali che possono modulare la stabilità del trascritto come shock termico, ipossia ecc. Tutti questi stimoli convergono verso un numero limitato di segnali interni che vanno a regolare l'attività delle proteine che legano le ARE (Eberart et al., 2007).

Per concludere è importante notare che la presenza di ARE all'interno dei messaggeri è strettamente legata ad una precisa modulazione dei meccanismi intracellulari che se non finemente regolati possono alterare le condizioni della cellula; dimostrazione di questo è ad esempio la regolazione delle citochine pro-infiammatorie attraverso la stabilità del mRNA, mutazioni a carico delle ARE di questi trascritti impediscono il legame delle ARE-BP determinando una minor degradazione del messaggero, un suo accumulo a livello cellulare e l'istaurarsi di uno stato infiammatorio cronico (Eberart et al., 2007).

Pertanto il contesto in cui le ARE si inseriscono sarà caratterizzato dalla presenza di un forte stimolo esterno per un tempo necessario a mantenere una risposta intracellulare costante, nonostante la rapida degradazione del trascritto e un'interruzione molto rapida nel momento in cui lo stimolo viene meno.

## 2 SCOPO

Il lavoro di questa tesi si inserisce in un ampio progetto che vede coinvolto il nostro Laboratorio, riguardante lo studio di alcuni aspetti di base della biologia del gene *ACVRI* ancora poco indagati, quali la regolazione della sua espressione. Esistono infatti molti lavori focalizzati sul suo prodotto proteico e sulla via di segnalazione mediata dalle BMP, mentre poco o nulla si sa sui suoi trascritti, sull'espressione del gene e come questa sia regolata.

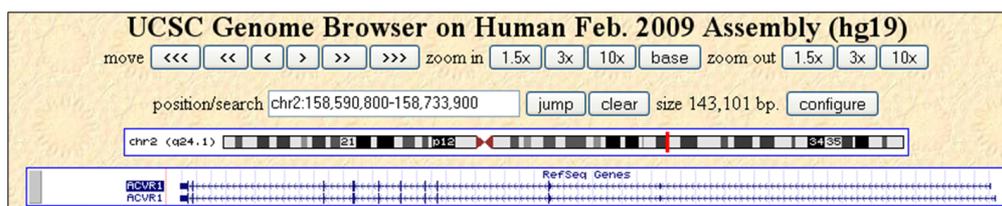
E' ormai ben noto che esistono vari livelli di regolazione dell'espressione genica, che può essere controllata sia a livello trascrizionale che a livello post trascrizionale e di traduzione.

Questa tesi si è focalizzata sul ruolo della sequenza 3'UTR del gene *ACVRI* nel modularne l'espressione. L'analisi bioinformatica iniziale ha permesso di evidenziare motivi strutturali e sequenze con un potenziale ruolo funzionale (siti consenso di legame per microRNA, elementi ARE), orientando la nostra indagine sperimentale verso la valutazione della stabilità del trascritto di *ACVRI* e il possibile ruolo degli elementi funzionali identificati *in silico*.

### 3 RISULTATI

## Struttura di *ACVR1*

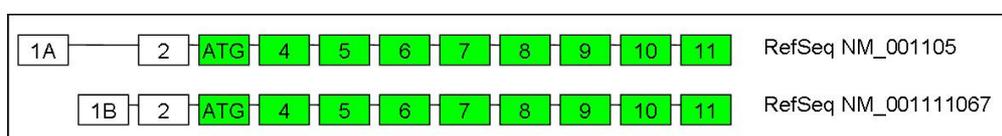
Il gene *ACVR1* è localizzato sul filamento *minus* del braccio lungo del cromosoma 2 in posizione 2q23-2q24, (chr2:158,592,95-158,732,374, ancoraggio al cromosoma in base a UCSC, versione Febbraio 2009) e copre una regione genomica di circa 139 417 bp (Figura 3).



**Figura 3** Localizzazione e isoforme di *ACVR1* riportate nella banca dati UCSC.

Quando è iniziato il lavoro di questa tesi, secondo quanto riportato in banca dati il gene sembrava essere trascritto in due isoforme principali, formate da undici esoni, caratterizzate da un primo esone trascritto ma non tradotto, un esone 2 comune e nove esoni codificanti, come schematizzato in Figura 4.

La prima isoforma riportata come RefSeq NM\_001105 rappresenta un trascritto di 3062 bp, la cui regione codificante inizia dal nucleotide 431 a livello dell'esone 3 e termina al nucleotide 1960 dell'esone 11. La seconda variante classificata come RefSeq NM\_00111067 è costituita da un trascritto di 2881 bp, la cui traduzione inizia dal nucleotide 250 dell'esone 3 e termina al nucleotide 1779 dell'esone 11.

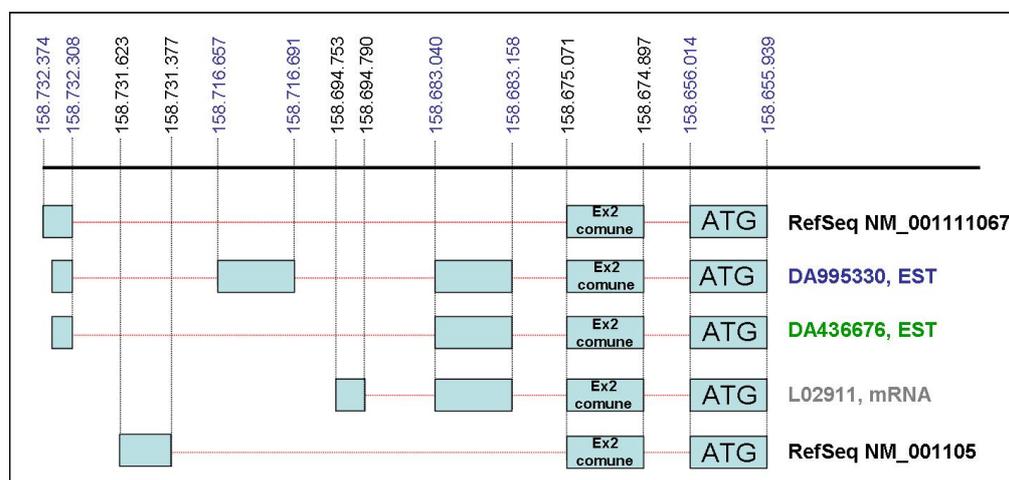


**Figura 4** Rappresentazione schematica della struttura delle isoforme principali riportate in banca dati.

I due trascritti hanno quindi in comune l'esone due non tradotto e la regione codificante per la proteina e mostrano un primo esone alternativo probabilmente dovuto alla presenza di due differenti siti di inizio trascrizione.

Recentemente i dati riguardanti il sequenziamento del genoma umano hanno

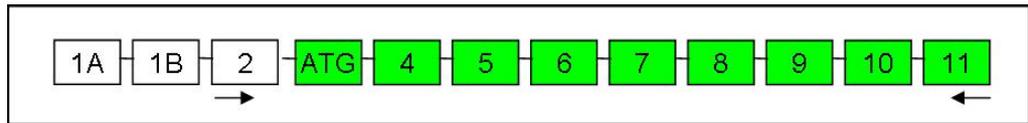
subito un processo di riannotazione e aggiornamento. Questo ha portato alla nostra attenzione la potenziale esistenza di messaggeri (mRNA) e sequenze EST (Expressed Sequence Tag) che potrebbero corrispondere a nuove isoforme trascrizionali del gene *ACVRI* che andrebbero ad affiancare le due sequenze di riferimento. Come schematizzato in Figura 5, ai due trascritti noti si aggiungono, almeno in via teorica, tre nuovi trascritti che hanno in comune un esone trascritto ma non tradotto (identificato da noi come esone 2 sulla base delle isoforme già riportate) e la parte codificante, e presentano una combinazione di diversi esoni alternativi, fiancheggiati da siti canonici di "splicing", nella regione 5'.



**Figura 5** Schema dell'estremità 5' dei messaggeri e delle EST di *ACVRI* depositate in banca dati. La numerazione relativa all'ancoraggio sul cromosoma 2 procede da destra verso sinistra perché il gene è localizzato sul filamento *minus*.

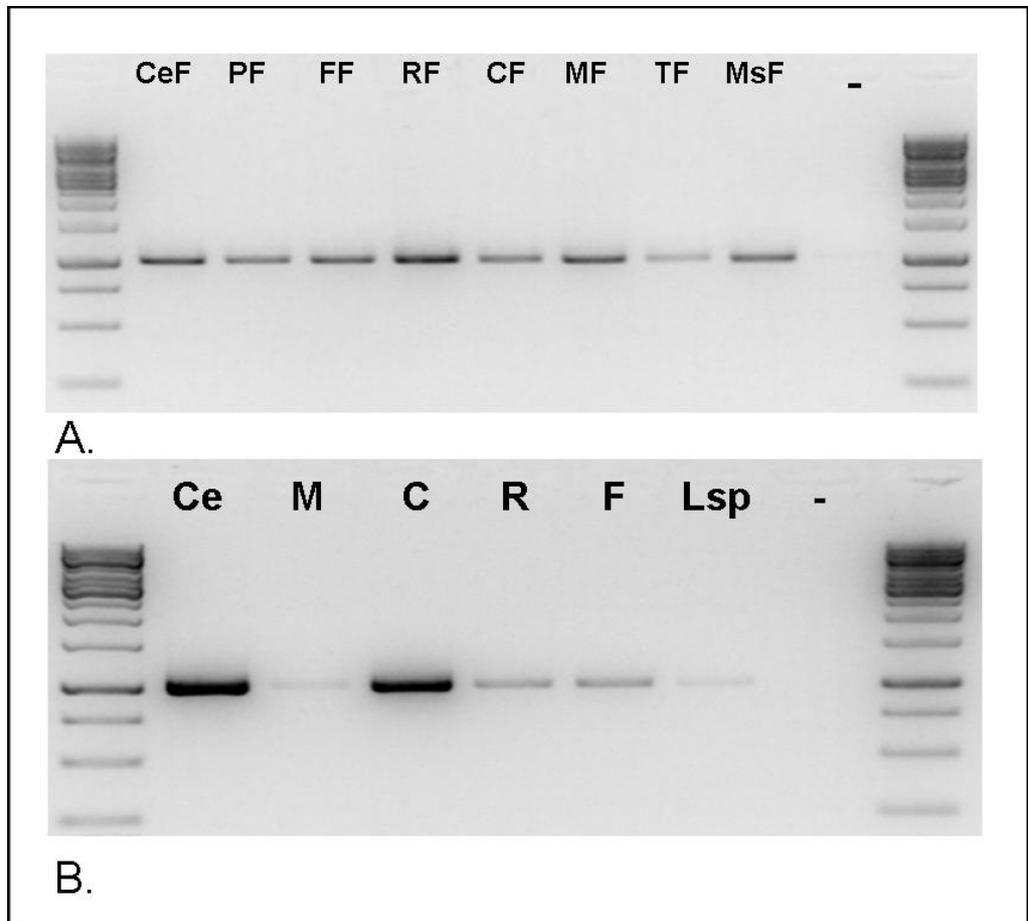
Abbiamo quindi iniziato il lavoro sperimentale analizzando il profilo di espressione di *ACVRI* e iniziato ad indagare le possibili isoforme utilizzando primer disegnati ad hoc per ciascuna EST

Per valutare l'espressione di *ACVRI* è stata fatta una serie di RT-PCR su cDNA umano proveniente da pannelli commerciali o ottenuto da linee cellulari normalmente usate nel nostro laboratorio. Inizialmente abbiamo deciso di utilizzare oligonucleotidi (vedi tabella 1) che permettessero di amplificare la parte del trascritto comune alle possibili isoforme (Figura 6).



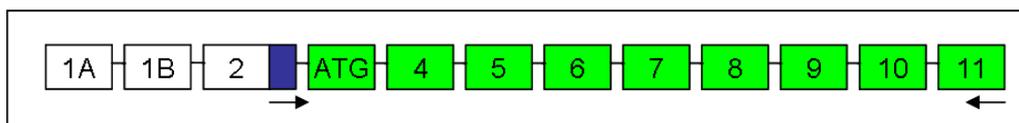
**Figura 6** Posizione degli oligonucleotidi (freccie) per l'amplificazione della regione di cDNA comune alle diverse isoforme .

Come si può vedere dalla Figura 7 A e B. *ACVR1* presenta un ampio profilo di espressione ed è presente in tutti i tessuti sia fetali (A) che adulti (B) analizzati, e, tra i tessuti adulti soprattutto a livello del cervello e del cuore.



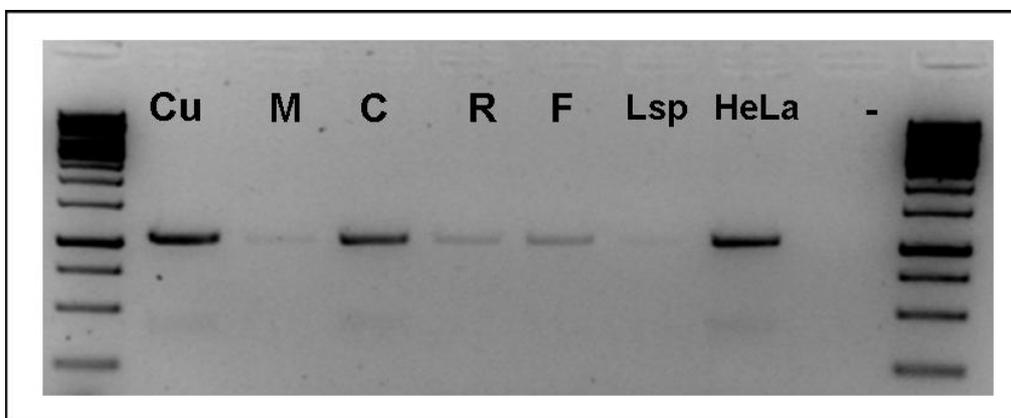
**Figura 7 A)** RT-PCR specifica per *ACVR1* su pannello di cDNA provenienti da diversi tessuti fetali: CeF cervello fetale, PF polmone fetale, FF fegato fetale, RF rene fetale, CF cuore fetale, MF milza fetale, TF timo fetale, MsF muscolo scheletrico fetale. **B)** RT-PCR specifica per *ACVR1* su pannello di cDNA di differenti tessuti adulti: Ce cervello, M milza, Cu cuore, R rene, F fegato, LSP Linfociti da sangue periferico.





**Figura 9** Posizione degli oligonucleotidi (freccie) per l'amplificazione della regione di cDNA contenente l'inserzione identificata.

In accordo con il fatto che probabilmente si tratta di un trascritto meno abbondante, a parità di quantità di cDNA di partenza e condizioni di reazione (i pannelli usati in questo studio sono tutti normalizzati), per ottenere questo prodotto abbiamo dovuto utilizzare un maggior numero di cicli di amplificazione (Figura 10).

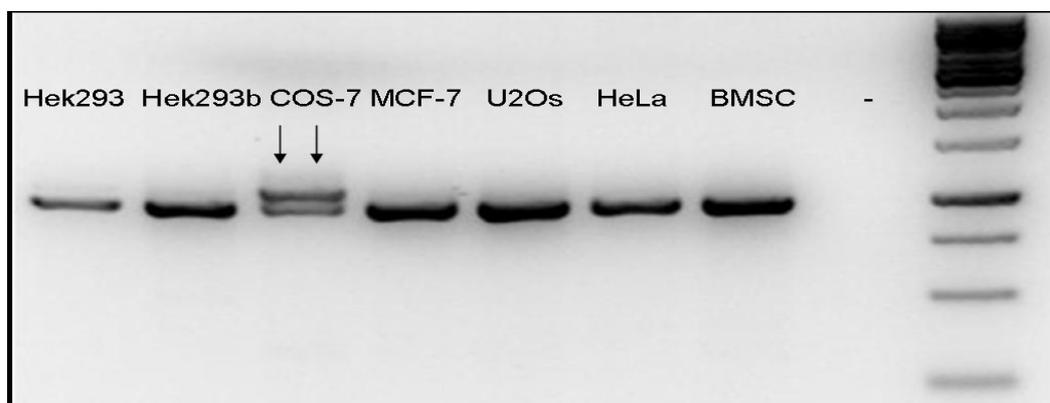


**Figura 10** RT-PCR specifica per l'isoforma di *ACVR1* contenente l'inserto di 116bp su pannello di cDNA proveniente da differenti tessuti adulti: Ce cervello, M milza, Cu cuore, R rene, F fegato, Lsp Linfociti da sangue periferico e HeLa (cellule tumorali da cancro alla cervice uterina).

*ACVR1* è risultato ben espresso anche in tutte le linee cellulari, da noi normalmente utilizzate, come dimostrato in Figura 11. E' da notare il fatto che in tutte è visualizzabile un'esile banda leggermente più alta di quella attesa, risultata corrispondente al trascritto con l'inserzione. La cosa interessante è che nelle cellule COS-7 entrambi i siti di splicing sembrano utilizzati con la stessa frequenza e gli amplificati del trascritto con e senza inserzione sono facilmente visualizzabili come bande (Figura 11, terzo pozzetto) di intensità simile, in esperimenti diversi e con controllo mediante sequenziamento.

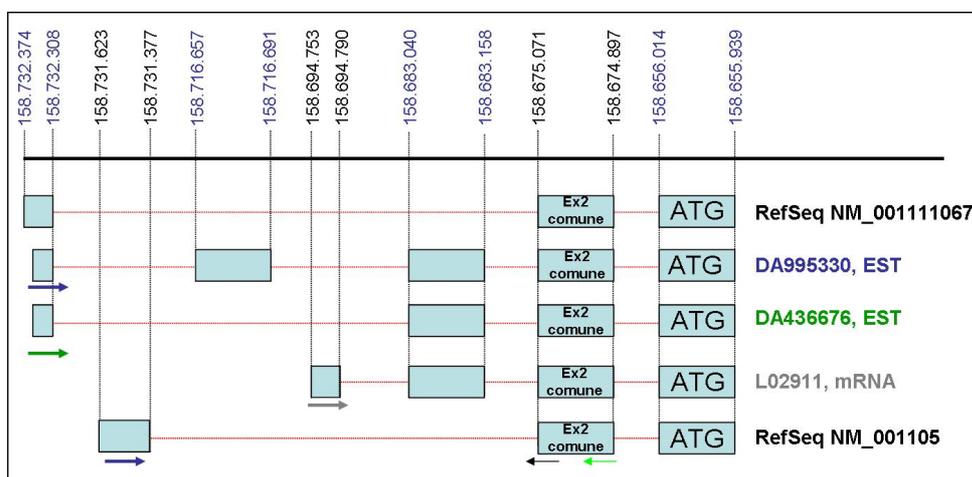
Purtroppo per mancanza di informazioni relative alle sequenze del genoma di *Cercopithecus aethiops* la scimmia verde africana, da cui provengono le cellule COS-7, non abbiamo potuto indagare ulteriormente le caratteristiche della

regione interessata dallo splicing alternativo sul DNA genomico, né fare un confronto con l'uomo.



**Figura 11** RT-PCR specifica per *ACVR1* su pannello di cDNA provenienti da linee cellulari: HEK-293 (cellule renali embrionali umane), COS-7 (cellule renali di scimmia), HeLa (cellule tumorali da cancro alla cervice uterina), MCF-7 (adenocarcinoma mammario umano), U2OS (osteosarcoma umano) e BMSC (cellule staminali mesenchimali midollari).

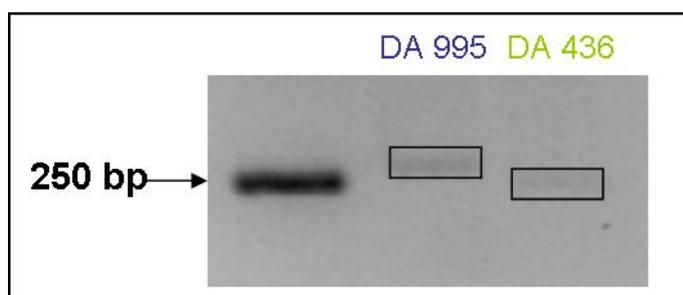
Abbiamo inoltre cominciato a caratterizzare la struttura e la composizione delle possibili isoforme con combinazioni diverse di esoni nella regione 5'UTR (Figura 12).



**Figura 12** Rappresentazione schematica dell'estremità 5' dei messaggeri e delle EST di *ACVR1* depositate in banca dati e posizione degli oligonucleotidi specifici (frecche colorate) per ognuna di esse. Come già indicato, la numerazione procede da destra verso sinistra perché il gene è localizzato sul filamento *minus* del cromosoma 2.

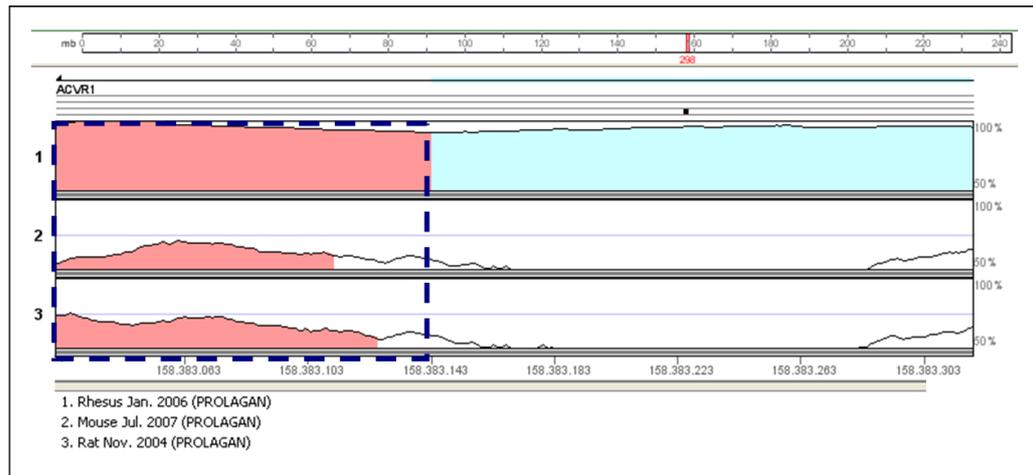
I risultati preliminari ottenuti dalle RT-PCR su cDNA umano a nostra disposizione (per oligo Tabella 1) hanno permesso di verificare l'esistenza

dell'isoforma corrispondente alla RefSeq NM\_001111067 e alle EST DA995330 e DA436676 (Figura 13) e la specificità del frammento ottenuto mediante sequenziamento diretto. Per il momento non siamo riusciti ad ottenere la regione 5' dei messaggeri L02911 e RefSeq NM\_001105 e stiamo ancora effettuando la messa a punto del protocollo di amplificazione. E' interessante notare il fatto che la RefSeq NM\_001105 corrisponde a EST di origine fetale (polmone e milza).



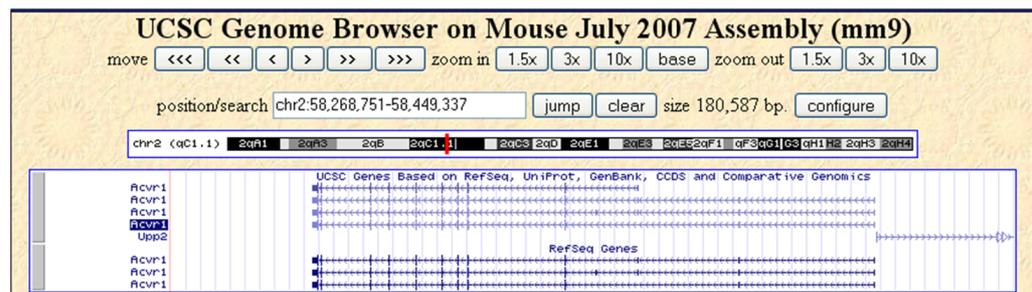
**Figura 13** RT-PCR specifica per le EST DA995330 e DA436676 ottenuta da cDNA di HeLa.

Per quanto riguarda l'inserzione da noi identificata, l'analisi comparativa mostra una significativa conservazione della sequenza tra le diverse specie. La Figura 14 mostra i risultati della comparazione ottenuta consultando il "VISTA Genome Browser". La regione riportata nel grafico corrisponde alla sequenza comprendente l'esone 2 trascritto ma non tradotto e la porzione intronica adiacente che si inserisce nel cDNA (finestra di 296 bp). Il gene è localizzato sul filamento minus del cromosoma 2, pertanto in figura è orientato da destra (5') verso sinistra (3'). I tre profili indicano il grado di conservazione nella comparazione uomo/scimpanzé, uomo/topo e uomo/ratto. In azzurro il programma indica le regioni codificanti, in rosso le regioni non codificanti ma con un livello di conservazione uguale o superiore al valore soglia del 70%. Come si può notare la regione che si inserisce presenta una conservazione superiore a quella dell'esone adiacente.



**Figura 14** Risultato dell'analisi di genomica comparativa ottenuta grazie al "VISTA Genome Browser" relativa alla sequenza dell'inserto di 116 bp (nel riquadro). 1, comparazione uomo/scimpanzé; 2, uomo/topo e 3, uomo/ratto. Le regioni codificanti sono in azzurro, le regioni non codificanti con un livello di conservazione uguale o superiore al valore soglia del 70% sono indicate in rosso.

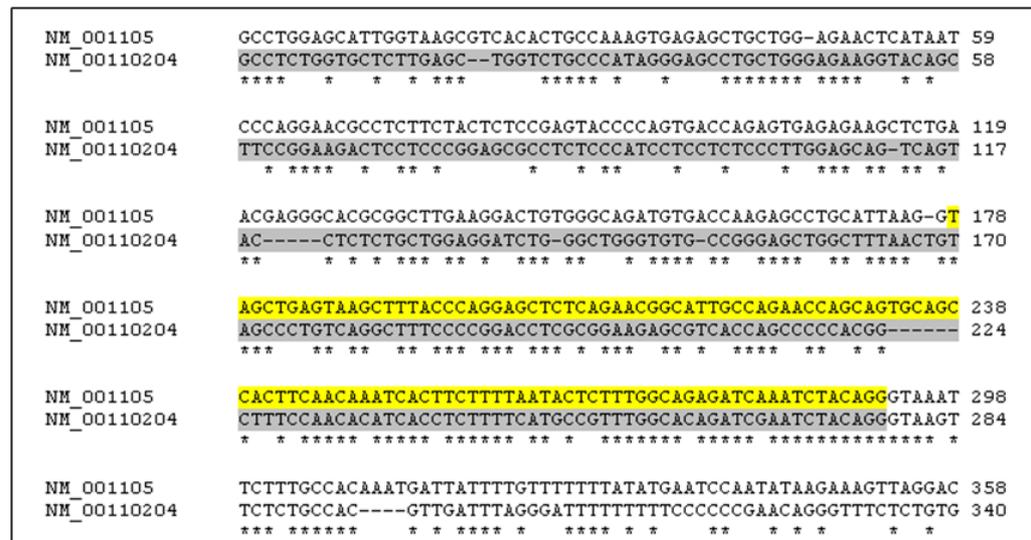
Da un'analisi più approfondita effettuata comparando i trascritti corrispondenti ad *ACVR1* umano e murino sono emerse due osservazioni: in primo luogo anche nel topo *ACVR1* è apparentemente espresso in diversi trascritti dovuti come nell'uomo sia a differenti siti di inizio trascrizione, che a processamenti alternativi che combinano esoni diversi nella regione 5'UTR ad una comune regione codificante (Figura 15).



**Figura 15** Localizzazione e isoforme di *acvr1* murino riportate nella banca dati UCSC.

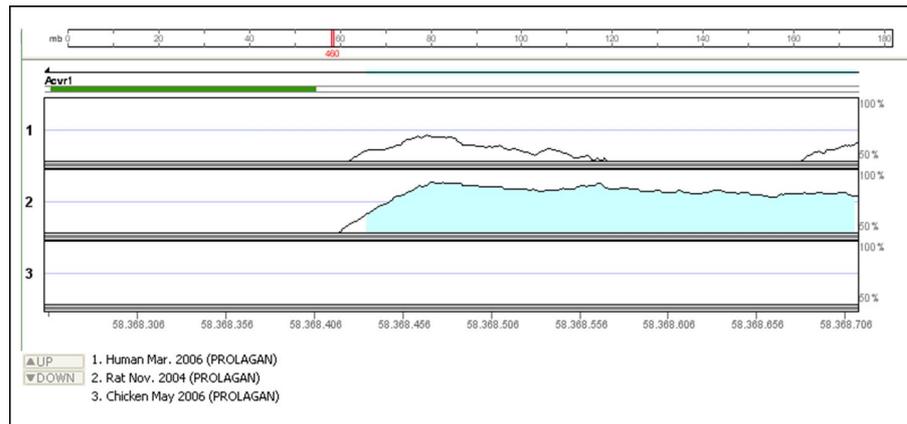
Inoltre, quello che nell'uomo si presenta come un esone che in alcuni casi può essere processato diversamente con l'inserzione alternativa delle 116 bp introniche adiacenti, nel topo è un esone unico, che può entrare o meno a far parte dei trascritti (lo contengono infatti la RefSeq NM\_00110204 e NM\_007394, ma non la RefSeq NM\_001110205), ma che apparentemente

viene processato come esone unico. In Figura 16 è riportato il confronto tra la sequenza umana (esone 2 + regione intronica adiacente evidenziata in giallo) e sequenza murina (evidenziata in grigio). Come si può osservare, il grado maggiore di identità di sequenza si ritrova nella regione corrispondente alle 116 bp dell'inserzione.



**Figura 16** Allineamento tra la sequenza umana (NM\_001105) e murina (NM\_00110204) corrispondente all'esone 2 trascritto ma non tradotto e all'inserzione di 116 bp (evidenziata in giallo). Gli asterischi indicano le posizioni conservate nelle due specie.

Analizzando nuovamente il profilo di conservazione (Figura 17), ma sfruttando come genoma di partenza quello di topo, si vede che in effetti nel topo e nel ratto questa sequenza forma un esone unico (indicato quindi in azzurro), rispetto all'uomo nel quale questa regione si è evoluta in modo indipendente favorendo la comparsa di uno "splicing" diverso. Quale possa essere il significato di questo processamento alternativo non è chiaro. Il mantenimento di un significativo livello di conservazione proprio della sequenza che in realtà diventa alternativa potrebbe indicare un qualche significato funzionale, magari in termini di regolazione dell'espressione del trascritto che la contiene, ma sono necessari approfondimenti sperimentali.



**Figura 17** Risultato dell'analisi di genomica comparativa ottenuta grazie al "VISTA Genome Browser" relativa alla sequenza dell'inserto di 116 bp e all'esone ad esso adiacente. In questo caso il genoma di riferimento per l'Analisi del profilo di conservazione è quello murino. 1, topo/uomo; 2, topo/ratto; 3, topo/pollo. Sia nel topo che nel ratto questa sequenza forma un esone unico (indicato in azzurro).

## **Il 3'UTR di *ACVRI***

Contrariamente a quanto osservato per la regione trascritta e non tradotta all'estremo 5' del gene, la regione 3'UTR è comune a tutte le isoforme identificate ed è rappresentato da una sequenza di 1081 bp (riportata in Figura 3). La regione 3'UTR regola numerosi aspetti del metabolismo del mRNA, incluso l'esportazione nucleare, la localizzazione citoplasmatica, l'efficienza di traduzione e la stabilità del mRNA (Andreassi et al., 2009), contribuendo a regolare l'espressione di un determinato gene a livello post trascrizionale.

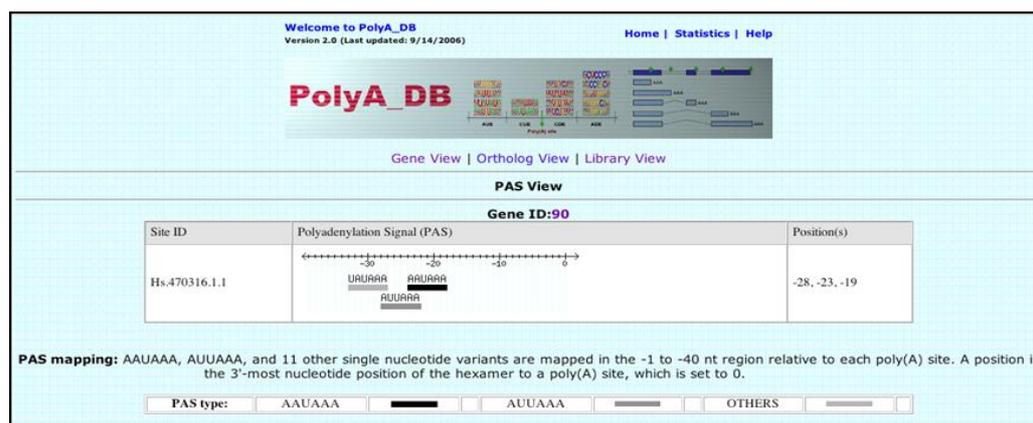
Il 3'UTR di *ACVRI* è stato sottoposto ad un'analisi bioinformatica per verificare la presenza di sequenze o motivi strutturali con potenziale ruolo nella regolazione post-trascrizionale; in particolare è stata approfondita la ricerca di siti alternativi di poliadenilazione, di siti putativi di legame con microRNA e la presenza di elementi ARE.

### **Siti di poliadenilazione**

L'analisi dei siti di poliadenilazione è stata effettuata consultando la banca dati PolyA\_DB disponibile in rete che consente di ottenere la localizzazione dei siti di poliadenilazione e la frequenza con cui possono essere riconosciuti.

In generale, esistono tre tipi di poliadenilazione alternativa, la poliadenilazione di tipo 1 può essere definita poliadenilazione costitutiva perché è presente solo un tipo di segnale al 3'UTR; la poliadenilazione di tipo 2 ha più di un segnale di poliadenilazione ma sono tutti presenti nell'esone al 3' e infine la poliadenilazione di tipo 3 coinvolge oltre alla poliadenilazione alternativa uno splicing alternativo (Lutz et al., 2008).

Come mostrato in Figura 18 e in Figura 21 il 3'UTR di *ACVRI* presenta una poliadenilazione costitutiva ossia un unico sito di poliadenilazione, identificato da un gruppo di tre sequenze parzialmente sovrapposte, tra cui una canonica AAUAAA che prevale sulle altre.



**Figura 18** Predizione siti di poliadenilazione presenti nel 3'UTR di *ACVRI* secondo il PolyA\_DB.

### Siti putativi di legame per i microRNA

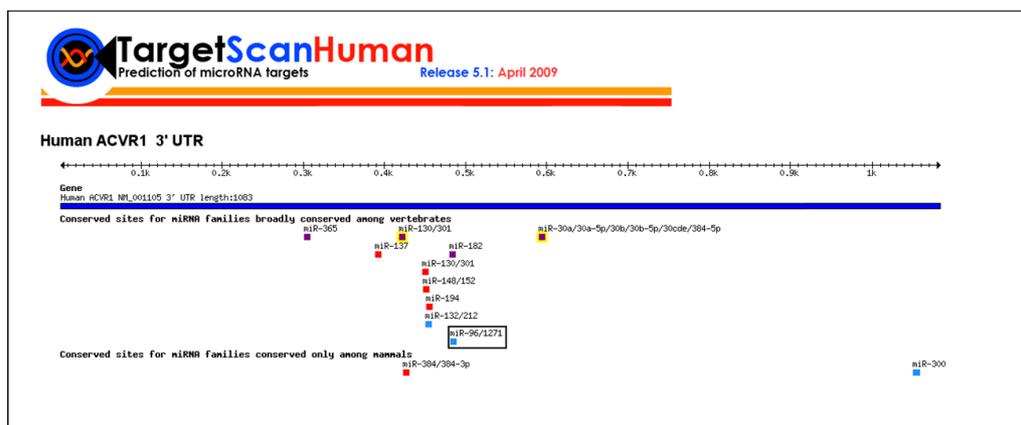
I miRNA hanno un ruolo ormai certo nella modulazione dell'espressione genica a livello post-trascrizionale sia durante processi fisiologici che patologici; essi agiscono legando le sequenze consenso presenti nel 3'UTR dei geni bersaglio dei quali regolano l'espressione sia in senso negativo che in senso positivo, nel primo caso avviando il trascritto alla degradazione o alterando la sua traducibilità, nel secondo promuovendo il re-inizio della traduzione e/o interagendo con altri elementi funzionali alla regolazione fine dei geni bersaglio (Carthew et al., 2009; Vasudevan et al., 2007).

Recentemente sono stati sviluppati numerosi programmi e banche dati disponibili in rete dedicati all'identificazione di putativi siti di legame per microRNA all'interno delle regioni 3'UTR dei potenziali geni bersaglio.

Per l'analisi del 3'UTR di *ACVRI* abbiamo utilizzato tre programmi disponibili in rete: Targetscan 5.1, PicTar e Mirbase.

Questi sono basati su algoritmi matematici in grado di identificare gli eventuali siti di legame tra i microRNA e gli RNA messaggeri bersaglio, tenendo in considerazione oltre alla complementarietà, alcuni parametri importanti come la conservazione evolutiva fra varie specie e la stabilità termodinamica dell'eteroduplex che si forma dall'interazione tra il microRNA e il 3'UTR.

In via teorica, si osserva che il 3'UTR di *ACVRI* contiene diversi siti di legame per microRNA, alcuni dei quali ben conservati tra le specie come dimostrano i risultati ottenuti con TargetScan 5.1 riportati in Figura 19 come esempio.



**Figura 19** Risultati dell'analisi del 3'UTR di *ACVR1* con il programma TargetScan 5.1. La figura mostra la posizione lungo la sequenza (schematizzata dalla barra blu) di siti di legame putativi per microRNA conservati nei mammiferi.

I tre programmi identificano numerosi potenziali siti di legame per miRNA nella sequenza 3'UTR di *ACVR1*, non completamente identici dal momento che gli algoritmi alla base di questi programmi presentano valori soglia e parametri di ricerca leggermente diversi.

La selezione dei microRNA da utilizzare per l'approfondimento funzionale è stata fatta tenendo conto di due aspetti che vengono normalmente considerati in letteratura per questo scopo, ovvero il grado di conservazione tra le specie di un determinato sito e il suo riconoscimento da parte di almeno due programmi differenti. Questo ci ha portato alla elaborazione di un elenco di 14 miRNA di potenziale interesse (Figura 20) al quale abbiamo anche aggiunto il miR-26a, il cui sito di legame era stato individuato da un unico programma ma che risultava comunque essere un candidato interessante da alcuni dati di letteratura.

		Micro RNA	PicTar	TargetScan	
MI0000748	MIMAT0000691	hsa-miR-130b: <a href="#">403</a> <a href="#">433</a>	P		
MI0000745	MIMAT0000688	hsa-miR-301a: <a href="#">405</a> <a href="#">432</a>	P	T	-14.46/-19.81
MI0000448	MIMAT0000425	hsa-miR-130a: <a href="#">406</a> <a href="#">435</a>	P	T	-16.35/-18.22
1- MI0000254 2- MI0000736	MIMAT0000244	hsa-miR-30c: <a href="#">576</a> <a href="#">605</a>	P		-15.92/-12.22
1- MI0000767 2- MI0000769	MIMAT0000710	hsa-miR-365: <a href="#">185</a> <a href="#">290</a>		T	-10.25/-14.17
MI0000454	MIMAT0000429	hsa-miR-137: <a href="#">372</a>		T	-13.49
MI0000253	MIMAT0000243	hsa-miR-148a: <a href="#">406</a> <a href="#">434</a>	P	T	-11.18/-12.05
MI0000255	MIMAT0000245	hsa-miR-30d: <a href="#">576</a>	P		-16.4
		hsa-miR-30e: <a href="#">576</a>	P		-14.28
MI0000272	MIMAT0000259	hsa-miR-182: <a href="#">458</a>	P	T	-12.34
MI0000462	MIMAT0000438	hsa-miR-152: <a href="#">436</a>	P	T	-12.21
MI0001145	MIMAT0001075	hsa-miR-384: <a href="#">412</a>	P	T	-11.47
MI0000811	MIMAT0000759	hsa-miR-148b: <a href="#">435</a>	P		-11.8
MI0000789	MIMAT0000736	hsa-miR-381: <a href="#">798</a>		T	-13.59

**Figura 20** Elenco dei miRNA selezionati in base ai criteri indicati nel testo.

### **Le sequenze ricche in AU**

Un ulteriore meccanismo di controllo dell'espressione genica è legato alla stabilità dei trascritti. Alcuni geni hanno messaggeri naturalmente instabili che vengono degradati rapidamente dalla cellula. Per mantenere alto il livello di espressione sono necessari segnali specifici che sostengano attivamente la trascrizione o che intervengano a stabilizzare i messaggeri. Le sequenze ricche in AU o ARE (AU rich elements) mediano spesso questa instabilità legando proteine specifiche e promuovendo la degradazione del trascritto (Eberart 2007).

Il 3' UTR è stato analizzato attraverso la consultazione della banca dati ARED disponibile in rete ed è emerso che *ACVRI* contiene sequenze ARE di gruppo I classe V ovvero diverse copie disperse del pentamero AUUUA (Figura 21)

10	20	30	40	50	60	70	80
* TGACATTTTC	* ATAGTGTCAA	* GAAGGAAGAT	* TTGACGTTGT	* TGTCATTGTC	* CAGCTGGGAC	* CTAATGCTGG	* CCTGACTGGT
90	100	110	120	130	140	150	160
* TGTCAGAATG	* GAATCCATCT	* GTCTCCCTCC	* CCAAATGGCT	* GCTTTGACAA	* GGCAGACGTC	* GTACCCAGCC	* ATGTGTTGGG
170	180	190	200	210	220	230	240
* GAGACATCAA	* AACCACCCTA	* ACCTCGCTCG	* ATGACTGTGA	* ACTGGGCATT	* TCACGAAGTC	* TTCACACTGC	* AGAGACTAAT
250	260	270	280	290	300	310	320
* GTTGGACAGA	* CACTGTTGCA	* AAGGTAGGGA	* CTGGAGGAAC	* ACAGAGAAAT	* CCTAAAAGAG	* ATCTGGGCAT	* TAAGTCAGTG
330	340	350	360	370	380	390	400
* GCTTTGCATA	* GCTTTCACAA	* GTCTCCTAGA	* CACTCCCCAC	* GGGAAACTCA	* AGGAGGTGGT	* GAATTTTTAA	* TCAGCAATAT
410	420	430	440	450	460	470	480
* TGCCGTGTCT	* TCTCTCTTT	* ATTGCACTAG	* GAATTCTTTG	* CATTCCTTAC	* TTGCACTGTT	* ACTCTTAAT	* TTAAAGACCC
490	500	510	520	530	540	550	560
* AACTTGCCAA	* AATGTTGGCT	* GCGTACTCCA	* CTGGTCTGTC	* TTTGGATAAT	* AGGAATTCAA	* TTTGGCAAAA	* CAAAATGTAA
570	580	590	600	610	620	630	640
* TGTCAGACTT	* TGCTGCATTT	* TACACATGTC	* CTGATGTTTA	* CAATGATGCC	* GAACATTAGG	* AATTGTTTAT	* ACACAACTTT
650	660	670	680	690	700	710	720
* GCAAATTATT	* TATTACTTGT	* GCACTTAGTA	* GTTTTTACAA	* AACTGCTTTG	* TGCATATGTT	* AAAGCTTATT	* TTTATGTGGT
730	740	750	760	770	780	790	800
* CTTATGATTT	* TATTACAGAA	* ATGTTTTTAA	* CACTATACTC	* TAAAATGGAC	* ATTTTCTTTT	* ATTATCAGTT	* AAAATCACAT
810	820	830	840	850	860	870	880
* TTTAAGTGCT	* TCACATTTGT	* ATGTGTGTAG	* ACTGTAACCT	* TTTTTCAGTT	* CATATGCAGA	* ACGTATTAG	* CCATTACCCA
890	900	910	920	930	940	950	960
* CGTGACACCA	* CCGAATATAT	* TACTGATTTA	* GAAGCAAAGA	* TTTCAGTAGA	* ATTTTAGTCC	* TGAACGCTAC	* GGGGAAAATG
970	980	990	1000	1010	1020	1030	1040
* CATTTTCTTC	* AGAATTATCC	* ATTACGTGCA	* TTTAAACTCT	* GCCAGAAAAA	* AATAACTATT	* TTGTTTTAAT	* CTACTTTTGG
1050	1060	1070	1080	1090	1100		
* TATTTAGTAG	* TTATTTGAT	* AAATTAATA	* AACTGTTTTC	* AAGTCAAAAA	* AAAAAAAAAA	* AAAAA	

**Figura 21** Sequenza corrispondente alla regione 3'UTR di *ACVRI* a partire dal codone di STOP (*tga*). Le sequenze consenso di legame per i microRNA sono indicate in blu e i relativi microRNA riportati sopra in rosa. In rosso sono evidenziati i pentameri ATTTA che definiscono gli elementi ARE ivi presenti. Infine, i rettangoli indicano la posizione dei siti di poli-adenilazione (vedere spiegazioni nel testo).

## Studi funzionali sul 3'UTR

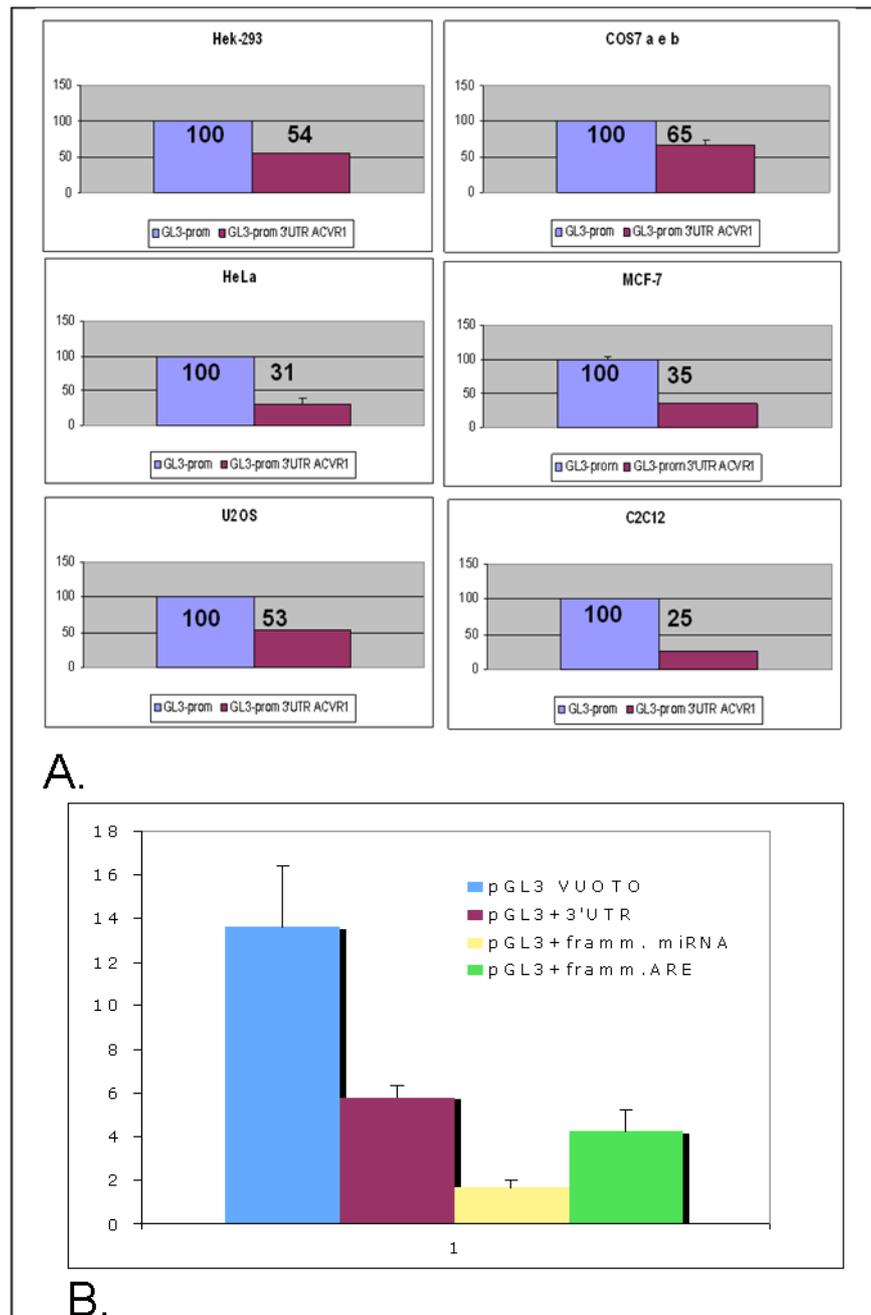
Per effettuare uno studio funzionale sul possibile ruolo della regione 3'UTR di *ACVRI* nella regolazione della sua espressione, l'intera sequenza di 1081 bp è stata subclonata nel vettore di espressione pGL3-Promoter (Promega) a valle della sequenza codificante per la Luciferasi. Il costrutto è stato trasferito in diverse linee cellulari ed è stata misurata l'attività del gene reporter. Questo tipo di esperimenti consente di effettuare un'analisi preliminare circa il possibile coinvolgimento a livello post-trascrizionale della regione in analisi. Se la sequenza inserita a valle del gene reporter ha un qualche ruolo questo verrà visualizzato in termini di variazione dell'attività luciferasica rispetto al

controllo costituito dal vettore vuoto.

La nostra analisi (Figura 22A) mostra una significativa diminuzione dell'attività del gene reporter in tutte le linee cellulari utilizzate, variabile dal 25% al 54% rispetto a quella misurata con il vettore vuoto. Quindi il 3'UTR sembra essere sede di una regolazione post trascrizionale.

L'analisi riguardante i siti di legame per microRNA e le sequenze ARE ha messo in evidenza una sorta di compartimentalizzazione con i siti per i microRNA aggregati nella posizione prossimale del 3'UTR e le sequenze ARE nella regione più distale. Per questo abbiamo preparato costrutti contenenti le due regioni individuate. Come mostrato in Figura 22B, entrambe i frammenti provocano una riduzione dell'attività luciferasica. Nel caso del frammento prossimale contenente i siti per miRNA la diminuzione è anche più significativa.

Questi risultati indicano che il 3'UTR del gene *ACVRI* può essere bersaglio di diversi meccanismi di regolazione post-trascrizionale. La riduzione dell'attività luciferasica può essere legata ad una destabilizzazione del messaggero del reporter, indotta dalla presenza del 3'UTR di *ACVRI* o in alternativa, ad un effetto sulla sua traducibilità. In entrambi i casi l'effetto viene visualizzato come riduzione di attività. Per cercare di capire il tipo di meccanismo che può intervenire siamo andati a verificare la stabilità del trascritto di *ACVRI*.



**Figura 22 A)** Valutazione dell'attività luciferasica mediante trasfezione del vettore di espressione pGL3 contenente la regione 3'UTR di *ACVR1* rispetto a trasfezione con il vettore vuoto in diverse linee cellulari. HEK-293, cellule renali embrionali umane; COS-7, cellule renali di scimmia; HeLa, cellule tumorali da cancro alla cervice uterina; MCF-7, adenocarcinoma mammario umano; U2OS, osteosarcoma umano; C2C12 mioblasti murini. **B)** Valutazione dell'attività luciferasica mediante trasfezione in cellule HeLa del vettore di espressione pGL3 vuoto (istogramma azzurro), o contenente l'intera regione 3'UTR (istogramma viola), il frammento recante i siti putativi di legame per i miRNA (istogramma giallo) o la regione con le sequenze ARE (istogramma verde).

## **Analisi della stabilità del trascritto di *ACVRI***

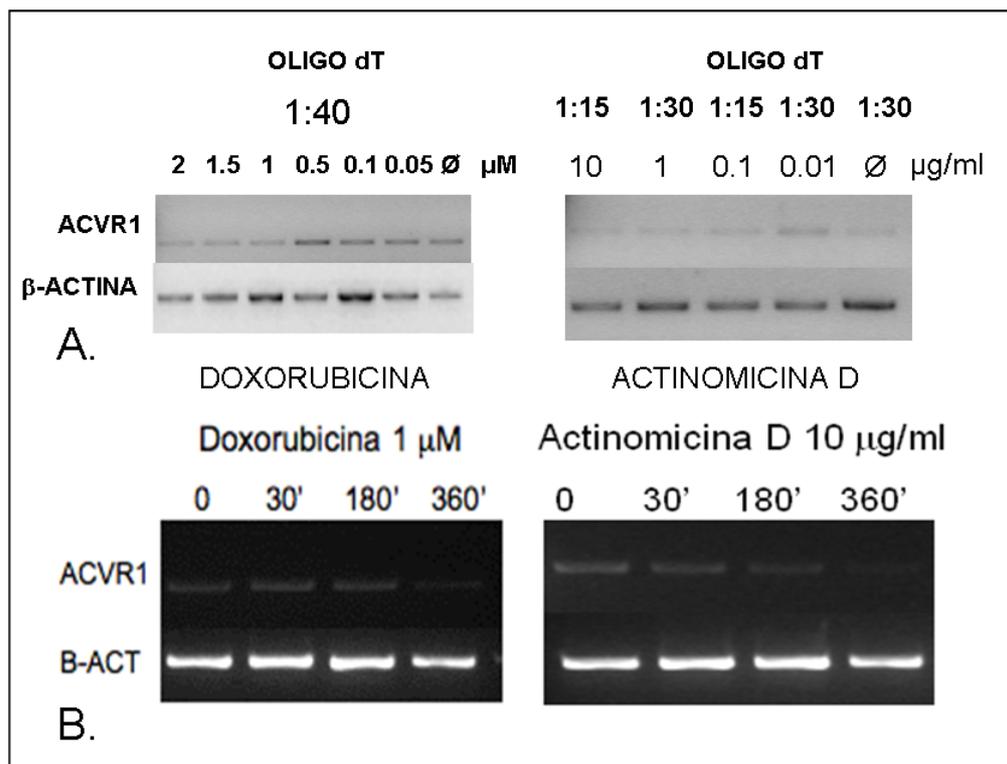
Lo studio della stabilità di un trascritto può essere valutata andando a misurare la sua espressione nel tempo in presenza di inibitori della trascrizione. Nel nostro studio sono stati utilizzati come inibitori: la doxorubicina e l'actinomicina D.

La Doxorubicina è un farmaco appartenente alla famiglia delle antracicline che viene comunemente utilizzato come antitumorale grazie alla sua elevata attività antineoplastica. Il suo meccanismo di azione è molto complesso e in parte sconosciuto; nota è la sua capacità di intercalare il DNA, impedendo così l'avanzamento della DNA topoisomerasi II la cui attività è quella di srotolare il DNA per la trascrizione (Kik et al., 2009).

L'Actinomicina D. appartiene ad una classe di antibiotici polipeptidici isolati dai batteri del genere *Streptomyces*. Questo farmaco è uno dei più vecchi chemioterapici utilizzati e viene anch'esso sfruttato in biologia cellulare per inibire la trascrizione. Questa caratteristica deriva dalla capacità di legare il DNA a livello del complesso di inizio della trascrizione e prevenire l'allungamento da parte dell'RNA polimerasi (Sobell 1985).

La messa a punto della concentrazione più efficace per valutare la stabilità del trascritto di *ACVRI* è avvenuta effettuando il trattamento a 3 ore con concentrazioni crescenti di Doxorubicina (2  $\mu$ M, 1,5  $\mu$ M, 1  $\mu$ M, 0,5  $\mu$ M, 0,1  $\mu$ M, 0,05  $\mu$ M) e Actinomicina D (10  $\mu$ g/ml 1  $\mu$ g/ml 0,1  $\mu$ g/ml 0,01  $\mu$ g/ml) per stabilire quella ottimale in cellule C2C12 (mioblasti murini) che esprimono costitutivamente *ACVRI*. L'espressione del gene è stata poi valutata mediante PCR semiquantitativa specifica.

La Figura 23A. mostra che a partire dalla dose 1  $\mu$ M di doxorubicina e 10  $\mu$ g/ml di Actinomicina D, l'espressione di *ACVRI* diminuisce significativamente rispetto al gene endogeno usato come riferimento ( $\beta$ Actina). L'effetto della doxorubicina e dell'Actinomicina D alle concentrazioni selezionate è stato valutato in cellule trattate per tempi differenti (30 minuti 180 minuti e 360 minuti ) e come mostrato nella Figura 23B. quello che si osserva dalla PCR semiquantitativa è una progressiva diminuzione dell'espressione del trascritto di *ACVRI*.



**Figura 23** Valutazione della stabilità del trascritto di *ACVR1* in presenza degli inibitori della trascrizione Doxorubicina e Actinomicina D mediante PCR semi-quantitativa. **A)** Il trascritto (*ACVR1*) mostra una progressiva diminuzione rispetto al gene di controllo ( $\beta$ -ACTINA) in presenza di concentrazioni crescenti dei due inibitori. **B)** Valutazione della stabilità del trascritto di *ACVR1* nel tempo in presenza delle concentrazioni di Doxorubicina e Actinomicina scelte.

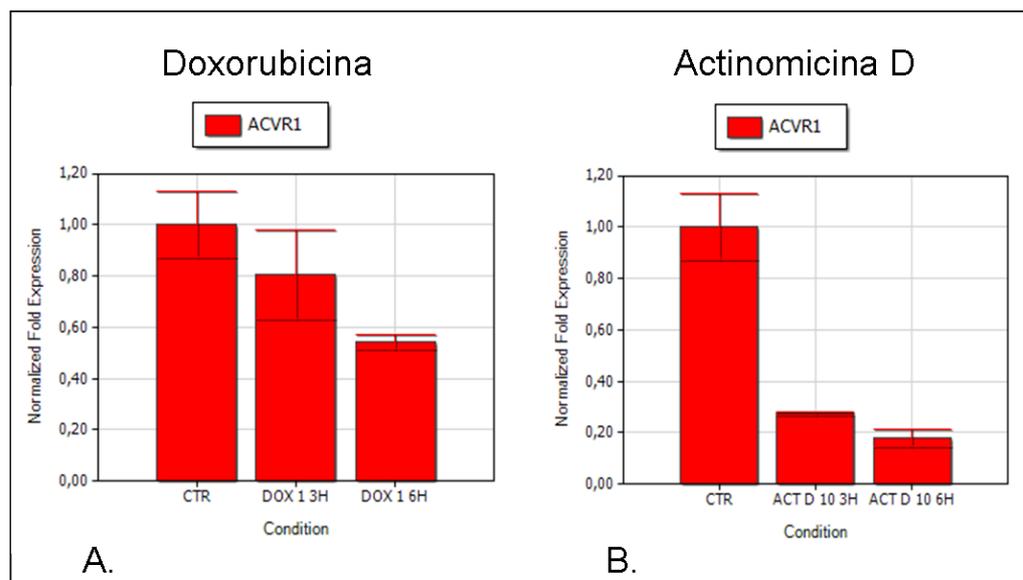
L'espressione del gene nei cDNA così ottenuti è stata anche misurata mediante PCR Real Time per avere una quantificazione reale del trascritto.

La Figura 22 mostra i grafici ottenuti dai valori medi dei triplicati in seguito ai trattamenti.

Per quanto riguarda la Doxorubicina (Figura 24A), dopo tre ore dal trattamento si osserva un calo dell'espressione di *ACVR1* del 20% che si riduce ulteriormente dopo sei ore (calo del 50%), rispetto al controllo rappresentato da cellule non trattate.

La Figura 24B mostra invece i grafici ottenuti dai valori medi dei triplicati in seguito a trattamento con Actinomicina D. Dopo tre ore si osserva un calo repentino del 75% dell'espressione di *ACVR1* che si riduce ulteriormente dopo sei ore, rispetto al controllo.

Questi dati indicano che il trascritto di *ACVR1* è instabile avvalorando l'ipotesi di una regolazione molto fine e precisa a livello cellulare.



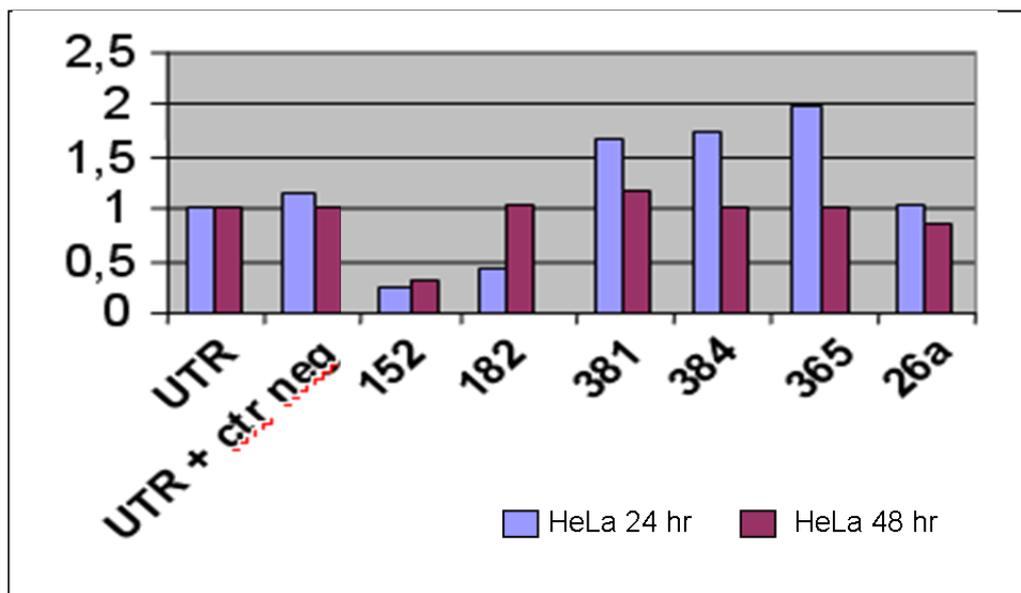
**Figura 24** Valutazione della stabilità del trascritto di *ACVR1* in presenza degli inibitori della trascrizione Doxorubicina e Actinomicina mediante PCR-quantitativa "real time". I grafici indicano i risultati ottenuto da un esperimento rappresentativo della quantificazione del trascritto di *ACVR1* in cellule di controllo (trattate con il solvente) e in cellule trattate con Doxorubicina (A) e Actinomicina (B) per 3 e 6 ore rispettivamente. La quantità normalizzata di trascritto nelle cellule di controllo è stata utilizzata come calibratore (valore 1,00) per determinare la variazione relativa della quantità di trascritto nelle cellule trattate.

### Azione dei miRNA sul 3'UTR di *ACVR1*

L'analisi *in silico* dei miRNA che potenzialmente legano il 3'UTR di *ACVR1*, ci ha portato a selezionarne alcuni secondo i criteri descritti in precedenza, come oggetto di approfondimento funzionale. Questa tesi riporta i risultati preliminari di alcuni di questi miRNA, ovvero miR-182, miR-152, miR-381, miR-384, miR-365 e miR-26a.

Questi, disponibili in commercio sotto forma di pre-miRNA (pre-MIR, Ambion; vedi Tabella 2) sono stati cotrasfettati in cellule HeLa assieme al costrutto pGL3-Prom- 3'UTR di *ACVR1*.

La Figura 25 mostra l'attività del gene reporter valutata dopo 24 e 48h dalla trasfezione; come si può osservare i miR-152 e -182 provocano una diminuzione dell'attività luciferasica, stabile per il primo e transiente per il secondo. I miR-381, -384 e -365 causano un aumento transiente dell'attività del gene reporter, mentre il miR-26a non sembra avere un effetto significativo.



**Figura 25** Effetto dei microRNA selezionati sul 3'UTR di *ACVR1*. I grafici riportano l'attività luciferasica relativa ottenuta cotrasfettando il vettore di espressione pGL3 contenente il 3'UTR di *ACVR1* con microRNA selezionati o un microRNA randomizzato come controllo negativo in cellule HeLa a 24 e 48 ore dalla trasfezione.

## 4 CONCLUSIONI

Lo studio delle isoforme di *ACVRI* ha permesso di individuare una notevole abbondanza a livello dell'estremità 5' trascritta ma non tradotta dell'RNA messaggero di *ACVRI*. Gli studi per approfondire queste osservazioni sono ancora in corso ma è già stato possibile confermare la presenza di alcune di queste isoforme (RefSeq NM\_001111067, EST DA995330 e DA436676) nei cDNA a nostra disposizione. Sarà interessante verificare l'esistenza delle altre isoforme e indagarne il profilo di espressione.

La regione 3'UTR di *ACVRI*, comune a tutte le isoforme, è stata sottoposta ad un'estesa analisi bioinformatica con diversi programmi e mediante la consultazione di specifiche banche dati disponibili in rete, ed è emersa la possibile esistenza di due moduli funzionali caratterizzati da una regione prossimale contenente diversi siti di legame per miRNA ed una più distale nella quale si concentrano elementi ARE.

L'analisi funzionale preliminare mostra che il trascritto di *ACVRI* è instabile e risulta significativamente diminuito in presenza degli inibitori della trascrizione e che il 3'UTR può essere sede di una regolazione post-trascrizionale.

I risultati ottenuti in questo lavoro costituiscono il punto di partenza per esperimenti mirati a valutare l'effetto dei singoli miRNA sul trascritto di *ACVRI*, il ruolo delle sequenze ARE e le possibili interazioni tra questi elementi funzionali nel modulare l'espressione del gene stesso.

## 5 MATERIALI E METODI

## **Analisi Bioinformatica**

I programmi utilizzati per le analisi bioinformatiche e le banche dati consultate in questo lavoro sono tutte liberamente disponibili in rete agli indirizzi elencati sotto.

- National Center for Biotechnology Information (NCBI), <http://www.ncbi.nlm.nih.gov/>
- UCSC, Genome Bioinformatics <http://genome.ucsc.edu/>
- VISTA "Tools for comparative Genomics", <http://pipeline.lbl.gov/cgi-bin/gateway2>

Analisi 3'UTR:

- PolyA\_DB, <http://polya.umdj.edu>
- TargetScan 5.1, <http://www.targetscan.org/>
- PicTar, <http://pictar.mdc-berlin.de/>
- miRBase, <http://www.mirbase.org/>.
- AU-rich Element- containing mRNA Database (ARED), <http://brp.kfshrc.edu.sa/ARED/>

## **Amplificazione di segmenti di DNA attraverso la Reazione a Catena della Polimerasi ("PolymeraseChain Reaction", PCR)**

La PCR è una tecnica che consente di amplificare una determinata sequenza di DNA. Si utilizzano due oligonucleotidi detti anche primer, appositamente disegnati per fiancheggiare il segmento di interesse, che fungono da innesco e un numero di cicli ripetuti di denaturazione, appaiamento dei primer e di allungamento/sintesi operata da una DNA polimerasi resistente alla temperatura (Taq). In condizioni standard le reazioni di PCR utilizzate in questo lavoro seguono il seguente schema generale: DNA genomico (100 ng), dNTPs, 2 mM per ciascuno ( dATP, dGTP, dTTP, CTP), Taq polimerasi (kit PCR extender system, 5PRIME), Tampone (kit PCR extender system, 5PRIME), Oligonucleotidi (10 pmoli/reazione). Nelle seguenti condizioni di amplificazione in termociclatore 2720 Thermal Cycler (Applied Biosystem): denaturazione iniziale (1 ciclo) 94°C x 5', ripetuto per 25/30 cicli: denaturazione 94°C x 45", appaiamento T°C oligo x 45", allungamento del

frammento 72°C x 1' ed estensione finale (1 ciclo) 72°C x 5'.

### **Purificazione e sequenziamento delle sequenze di PCR**

La purificazione del prodotto di PCR è stata effettuata per via enzimatica grazie all'uso di Exo-SAP-IT (USB), una miscela contenente Esonucleasi I e "Shrimp Alkaline Phosphatase" che consente di rimuovere dNTP e oligonucleotidi non utilizzati durante la reazione di PCR e che potrebbero interferire negativamente con le fasi successive.

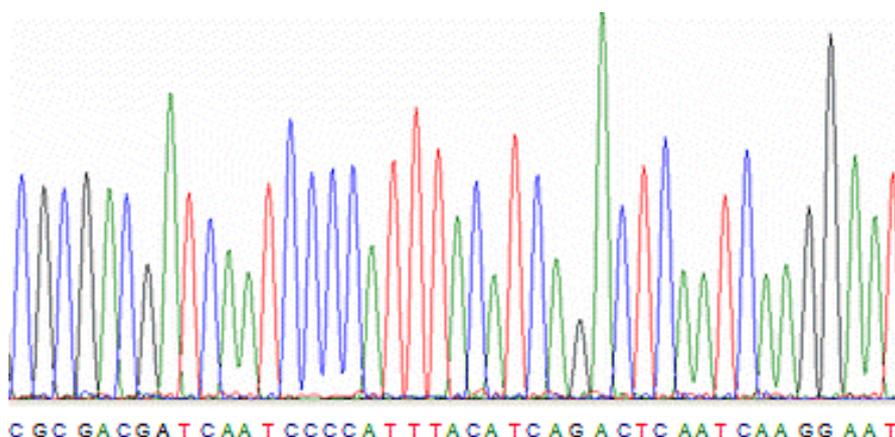
In condizioni standard l'azione dell' Exo-SAP-IT si ottiene mediante l'utilizzo di 2 µl di reagente ogni 6 µl di prodotto di PCR, nelle seguenti condizioni: 37°C x 30', 80°C x 15', 4°C x 5'.

Per il sequenziamento dei prodotti di PCR abbiamo utilizzato il BigDye Terminator Kit (Applied Biosystems) che consente di far avvenire le quattro reazioni del metodo di Sanger in un'unica reazione. I quattro dideossinucleotidi marcati con fluorocromi diversi vengono incorporati in corrispondenza della loro base complementare provocando la terminazione della reazione di sintesi di DNA.

In condizioni standard gli elementi della reazione sono: prodotto PCR purificato/plasmidi (500 ng per reazione), oligonucleotidi (5 pmoli a reazione), tampone fornito dal kit e miscela BigDye. Nelle seguenti condizioni di lavoro in termociclatore 2720 Thermal Cycler (Applied Biosystem): 96°C x 3', 94°C x 10", 56°C x 20" e 60°C x 4' ripetute per 25 cicli.

Le reazioni di sequenza sono state poi precipitate con 1/10 del volume di Sodio Acetato 3M pH 5,0 e 2,5 volumi di Etanolo assoluto. Dopo una centrifugata di 30' a 3900 RPM, si effettua un lavaggio con Etanolo 70% per 10' alla stessa velocità. Dopo essiccazione ottenuta mediante una breve centrifugazione a vuoto, le reazioni di sequenza sono pronte per la corsa su sequenziatore automatico ad elettroforesi capillare, nel quale la corsa dei frammenti, la rilevazione delle bande e l'identificazione delle basi è compiuta automaticamente (ABI PRISM, Applied Biosystems). Un computer collegato allo strumento elabora i segnali rilevati convertendoli in un profilo di intensità (elettroferogramma) nel quale ogni fluorocromo risulta diversamente colorato (vedi Figura 26). Le sequenze così ottenute vengono analizzate con il

programma Sequence Navigator che consente, tra l'altro, l'allineamento e il confronto con sequenze di riferimento.



**Figura 26** Esempio di elettroferogramma ottenuto dopo corsa su sequenziatore automatico ed elaborazione con il programma di visualizzazione Chromas.

### **Preparazione dei costrutti di espressione in pGL3 Promoter vector**

La regione 3'UTR e i suoi frammenti usati negli studi funzionali sono stati ottenuti tramite PCR da DNA genomico di controllo utilizzando oligonucleotidi contenenti una coda che permette l'introduzione del sito di restrizione XBA1 (Tabella 1). In tal modo il prodotto di PCR che si ottiene risulta fiancheggiato, all'estremo 5' dei due filamenti, dal sito di restrizione utile al subclonaggio successivo. Per verificare la specificità del frammento e che la Taq polimerasi non abbia introdotto errori nella sequenza, i prodotti di PCR ottenuti sono stati prima subclonati sfruttando il sistema commerciale TOPO-TA Cloning (Invitrogen). Molte Taq polimerasi normalmente utilizzate per PCR uniscono, in modo indipendente dallo stampo, alle estremità 3' del prodotto di amplificazione, una singola base protrudente, la A. Il kit della Invitrogen sfrutta questo principio fornendo un vettore plasmidico (pCR 2.1-TOPO) aperto e contenente una base T ad entrambe le estremità complementare alle "A" del frammento di PCR. Una ligasi fornita dal kit assieme ad un tampone specifico consente la ligazione di vettore e prodotto di PCR. Il protocollo è stato ottimizzato in modo da consentire il verificarsi della ligazione in pochi minuti a temperatura ambiente e in un totale di 6µl dei quali

0.5/4µl di prodotto di PCR, 1µl di soluzione salina ed 1µl di vettore TOPO/ligasi. Una parte della reazione viene poi trasformata in cellule competenti (TOP 10), sempre fornite dal kit, secondo una metodica standard, ovvero 20/30 minuti in ghiaccio seguiti da uno shock termico a 42°C per 30 secondi. Viene poi aggiunto il terreno SOC e le cellule vengono poste in coltura a 37°C per circa 1 ora. La coltura batterica viene quindi piastrata su petri contenenti LB/agar reso selettivo grazie all'aggiunta di Ampicillina ed X-gal (SIGMA). In particolare, la presenza di quest'ultimo, consente di selezionare i cloni ricombinanti mediante il sistema "blu/bianco". Infatti, il sito di clonaggio per il frammento di PCR è situato all'interno del gene LacZ. Se il vettore non è ricombinante, la chiusura su se stesso causerà il ripristino della sequenza codificante per la β-galattosidasi che, una volta espressa, permetterà ai batteri di metabolizzare l'X-Gal fornendo colorando la colonia di blu. Viceversa, se il vettore contiene il frammento di PCR, non si avrà il ripristino del LacZ e la colonia rimarrà bianca. Le colonie ricombinanti sono state caratterizzate per restrizione e sequenziamento. La regione 3'UTR e i suoi frammenti così ottenuti e verificati sono stati excisi mediante restrizione XBA1 e subclonati nel vettore di espressione pGL3-Promoter nel corrispondente sito XBA1 immediatamente a valle del codone di STOP della sequenza codificante per la Luciferasi. In tal modo si genera un trascritto ricombinante la cui espressione è diretta in modo costitutivo dal Promotore di SV40 contenuto nel vettore e nel quale la sequenza codificante per il gene reporter presenta il 3'UTR di *ACVRI* o un suo frammento.

### **Estrazione DNA plasmidico**

Le preparazioni di DNA plasmidico per il sequenziamento e per i subclonaggi sono state fatte con il kit "Fast plasmid mini" della 5 Prime, che consente di ottenere il DNA in modo veloce e semplice. Le colture batteriche sono state fatte crescere in terreno Luria-Bertani standard (LB) reso selettivo con Ampicillina in incubatore a 37°C, in agitazione per tutta la notte. Le cellule batteriche sono state raccolte per centrifugazione e direttamente lisate con una soluzione fornita dal kit. I lisati sono stati poi caricati su colonnine e dopo un lavaggio, il DNA si ottiene per eluizione con tampone fornito dal kit.

Per gli esperimenti di trasfezione il DNA plasmidico è stato preparato con il kit NucleoBond Xtra Midi EF (Macherey-Nagel) che consente di ottenere concentrazioni e livelli di purezza superiori. In breve, il kit prevede una procedura basata sul recupero del pellet batterico tramite centrifugazione, risospensione in una soluzione di lavaggio (soluzione MN1), lisi alcalina in presenza di NaOH/ SDS che provoca la rottura delle membrane cellulari, precipitazione di SDS e detriti cellulari tramite l'aggiunta di una soluzione (soluzione MN3) contenente Acetato di Sodio e successiva centrifugazione. Il sopranatante chiarificato viene quindi applicato su colonnine contenenti una resina capace di legare il DNA. Resina e DNA vengono quindi lavati con diversi tamponi forniti dal kit e il DNA viene eluito con Tris HCl 10 mM pH 8.5.

### **Culture cellulari**

Le linee cellulari utilizzate in questo lavoro erano già disponibili in Laboratorio, acquistate in precedenza dall'ATCC ("American Type Culture Collection") e conservate in azoto liquido in terreno di congelamento costituito per il 90% da siero e per il 10% dal dimetilsolfossido come agente crioprotettore. Allo scongelamento le cellule sono state poste nell'opportuno terreno di coltura in incubatore umidificato a 37°C e al 5% di CO<sub>2</sub>.

Hek-293, COS-7, MCF-7, U2OS e C2C12 sono state coltivate in terreno Dulbecco's Modified Eagle's Medium (DMEM) supplementato con 10% di siero bovino fetale, Penicillina/Streptomina ed L-Glutammina. Le cellule HeLa sono state coltivate in Minimum Essential Medium (MEM) supplementato come sopra al quale sono stati anche aggiunti amminoacidi non essenziali e Sodio piruvato.

### **Trasfezione del 3'UTR nelle cellule selezionate**

Il protocollo che segue è stato utilizzato per lo studio funzionale del 3'UTR e dei suoi due frammenti contenenti i siti di legame per i miRNA e le sequenze ARE rispettivamente.

Le trasfezioni effettuate con i soli costrutti plasmidici sono state effettuate con il FuGENE 6, reagente commerciale non-liposomale della Roche. Le trasfezioni con i precursori dei miR selezionati e le cotrasfezioni dei pre-MIR con il DNA plasmidico sono state invece effettuate con il siPORT*NeoFX*<sup>TM</sup>, reagente commerciale di tipo lipidico adatto anche alla trasfezione dei piccoli RNA.

Per la valutazione dell'attività Luciferasica le cellule sono state seminate in piastre da 24 pozzetti ad una densità variabile dalle 45 000 alle 90 000 cellule per punto, a seconda della linea cellulare, trasfettate al momento della piastratura con FuGENE 6 secondo il protocollo fornito e incubate per 24 e 48 ore. Oltre ai costrutti in analisi, è stato trasfettato il vettore vuoto come riferimento e un vettore contenente la Luciferasi di Renilla (pRL) come gene reporter, la cui attività è stata valutata come misura dell'efficienza di trasfezione per normalizzare l'attività Luciferasica dei costrutti in analisi. Per ogni costrutto gli esperimenti sono stati fatti in triplicato.

I miRNA utilizzati in questo studio sono stati acquistati dalla Ambion (Applied Biosystems) sotto forma di pre-miR, ovvero ancora nella forma a doppio filamento. Questi risultano più stabili e facilmente trasfettabili, una volta nella cellula vengono maturati a miR in modo molto efficiente e del tutto analoga a quello che succede per i miR endogeni. Come controllo negativo è stato utilizzato un pre-miR a sequenza randomizzata e coniugato a CY3, molecola fluorescente visualizzabile al microscopio a fluorescenza nelle cellule in coltura in modo da monitorare l'efficienza della trasfezione ad ogni esperimento. Per ogni miR è stato acquistato anche il relativo antisenso (anti-miR) per verificare la specificità degli effetti osservati. Il protocollo seguito per la trasfezione con i miRNA (15 pmoli di precursore miR o anti-miR per punto) nella cellule HeLa è uguale a quello appena descritto tranne che per l'utilizzazione del siPORT*NeoFX*<sup>TM</sup> come agente trasfettante.

Per la valutazione dell'attività luciferasica è stato utilizzato il Dual-Luciferase Reporter Assay System (Promega) con le cellule raccolte a 24 e 48 ore dalla trasfezione. In breve, le cellule sono state lisate in 100 µl di "Passive Lysis Buffer 1x" fornito dal kit per 15 minuti a temperatura ambiente e in agitazione. 20 µl di lisato così ottenuto sono stati cimentati con 50 µl di una soluzione

contenente un tampone (Luciferase Assay Buffer II) e il substrato (Luciferina) per la Luciferasi del pGL3-Promoter ed è stata effettuata una prima lettura al Luminometro (TD-20/20 Luminometer Turner Designs). Nella stessa provetta sono stati quindi aggiunti 50  $\mu$ l di una soluzione contenente un tampone che blocca la prima reazione Luciferina/ Luciferasi (Stop&Glo buffer) e il substrato per la Luciferasi di Renilla del vettore pRL. I dati ottenuti sono stati normalizzati per l'efficienza di trasfezione secondo la formula (valore Luciferasi x 1000)/valore Renilla, sono state quindi calcolate le medie e le deviazioni standard relative ai valori normalizzati dei triplicati per ogni singolo costrutto in analisi.

### **ANALISI DELLA STABILITA' DEL TRASCritto**

Lo studio della stabilità del trascritto è stato eseguito sottoponendo le cellule selezionate, C2C12 al trattamento con Doxorubicina e Actinomomicina D.

Per stabilire le concentrazioni ottimali sono state seminate circa 150000 cellule in piastre da 35 mm, lasciate in coltura per 24 ore con terreno completo e quindi trattate per 3 ore con concentrazioni crescenti dei due composti, 2  $\mu$ M, 1,5  $\mu$ M, 1  $\mu$ M, 0,5  $\mu$ M, 0,1  $\mu$ M, 0,05  $\mu$ M per la Doxorubicina e 10  $\mu$ g/ml 1  $\mu$ g/ml 0,1  $\mu$ g/ml 0,01  $\mu$ g/ml per l'Actinomomicina D, entrambe sciolte in DimetilSolfossido (DMSO).

L' RNA è stato estratto dalle cellule con il kit RNeasy Mini (Qiagen) seguendo le istruzioni riportate nel kit. In breve, la procedura prevede: l'utilizzo di un Tampone (RTL) a cui si aggiunge il  $\beta$ -Mercaptoetanolo per facilitare la rottura delle cellule, il trasferimento dei lisati cellulari nelle colonnine fornite dal kit, due successivi lavaggi con il tampone RPE fornito dal kit, eluizione con H<sub>2</sub>O RNasi free.

### **Retrotrascrizione**

La trascrittasi inversa è una polimerasi in grado di sintetizzare DNA a singolo filamento (DNA complementare o cDNA) a partire da RNA utilizzato come stampo. Come innesco per la reazione di retrotrascrizione vengono utilizzati

degli Oligo-dT che, appaiandosi alle sequenze di Poli-A presenti negli mRNA, rendono possibile la reazione e la sintesi di un filamento complementare ad essi. Le molecole di cDNA così ottenute possono quindi essere utilizzate, ad esempio, come stampo per reazioni di PCR con oligonucleotidi gene-specifici. Per la sintesi dei filamenti di interesse è stato utilizzato il protocollo previsto dal kit Advantage™ RT-for-PCR kit (Clontech) che viene raccomandato per ottenere una sintesi completa e ad alta sensibilità.

Esso prevede la preparazione di una miscela di reazione contenente, in un totale di 20µl: 1µg di RNA totale, Oligo(dT) (20 pmoli), tampone di reazione (50mM Tris-HCL, pH 8.3; 75 mM KCl; 3 mM MgCl<sub>2</sub>), i quattro dNTP (0.5 mM), un inibitore per l'Rnasi (1U/µl finali) e l'enzima M-MLV RNA retrotrascrittasi (≥200 U/µg RNA). Il programma, in Perkin Elmer Thermal Cycler PE 9700, prevede:

- un ciclo per l'RNA a 70° C per 2 minuti,
- aggiunta delle restanti componenti,
- sintesi a 42° C per 60 minuti,
- inattivazione dell'enzima a 94° C per 5 minuti.

I 20 µl della reazione vengono poi portati a 100µl finali con H<sub>2</sub>O RNAsi free.

### **PCR semiquantitativa**

I cDNA ottenuti sono stati utilizzati come stampo per PCR semiquantitative per quantificare l'espressione di *ACVRI* rispetto ad un gene di riferimento espresso costitutivamente dalle cellule, ovvero la β-actina di topo per i cDNA ottenuti da linee cellulari murine e la G3PDH per i cDNA di origine umana. Tali PCR sono state eseguite con diverse diluizioni dei campioni di partenza e a basso numero di cicli di amplificazione; quindi ci siamo basati sul confronto dell'intensità dell'amplificato finale del gene di interesse rispetto a quella dell'amplificato del gene di riferimento.

Dai dati ottenuti con la PCR semiquantitativa sono state stabilite le concentrazioni ottimali per il trattamento delle cellule con gli inibitori di trascrizione, che sono risultate essere 1 µM e 0,1 µM per la doxorubicina e 10 µg/ml per l'Actinomicina D, rispettivamente.

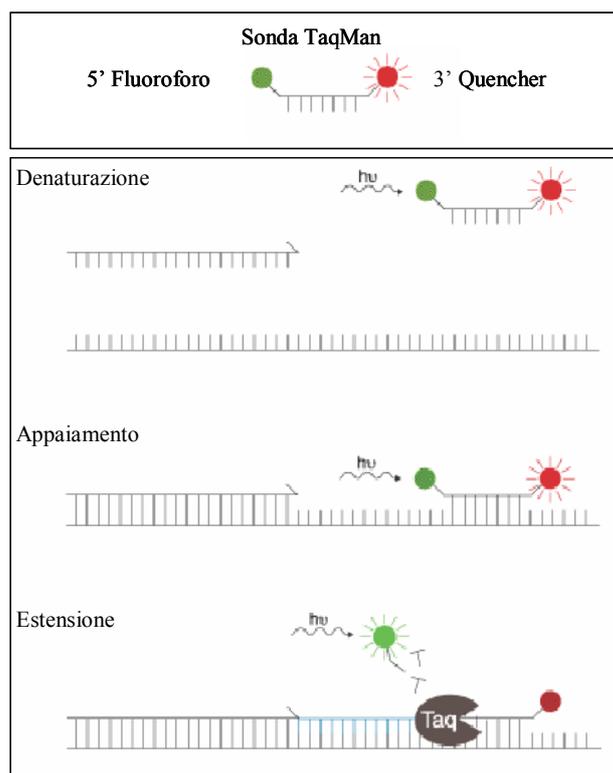
Le cellule, seminate in piastre da 35 mm ad una densità di circa 150000 cellule per punto, sono state trattate per tempi differenti (30 minuti, 180 minuti e 360 minuti) con i due composti. Il cDNA ottenuto dalla retrotrascrizione dell'RNA estratto dalle cellule trattate e di controllo (trattate con il solvente), è stato utilizzato come stampo per quantificare il trascritto di *ACVRI* mediante PCR in tempo reale (real-time PCR).

### **Real-Time PCR**

La "Real-Time PCR" prevede la quantificazione di un determinato campione di DNA o cDNA grazie alla misurazione in tempo reale dell'amplificato che si viene via via sintetizzato. La PCR porta ad una sintesi esponenziale del prodotto di interesse. Il grafico che consente di visualizzare l'accumulo di prodotto in funzione del tempo assume perciò un andamento sigmoideale con una fase di crescita assimilabile ad una retta, un punto di flesso e il raggiungimento di un valore massimo ("plateau"), che non può essere ulteriormente incrementato. In condizioni normali l'amplificato viene valutato solo al termine di questa curva quindi quando è stato raggiunto il valore massimo. Questo rende difficoltoso dedurre la quantità della specifica sequenza d'interesse presente nel campione di partenza usato come stampo. La PCR in tempo reale consente invece di valutare l'accumulo di prodotto opportunamente marcato, via via che viene sintetizzato, visualizzandolo nella fase di crescita proporzionale della curva. In questo contesto, maggiore è il numero di molecole stampo presenti all'inizio della reazione e minore sarà il numero di cicli necessari per raggiungere il punto di flesso della curva di amplificazione (valore soglia).

Esistono diversi protocolli per l'allestimento di PCR "real time". La chimica usata nei nostri esperimenti prevede l'impiego delle sonde commerciali TaqMan (Applied Biosystems). La sonda è complementare ad una regione contenuta all'interno del frammento di PCR che si ottiene dal DNA o cDNA che si desidera quantificare. Inoltre è marcata all'estremo 5' con un fluoroforo e al 3' con una molecola "estintore" o "quencher". Quando la sonda è intatta, sia libera che appaiata alla propria sequenza bersaglio non si ha emissione di fluorescenza per effetto dell'inibizione reciproca fluoroforo/"quencher".

Durante il processo di sintesi e allungamento dell'amplificato, l'attività esonucleasica della Taq polimerasi provoca la degradazione della sonda. Fluoroforo e "quencher" vengono quindi separati e si ha emissione di fluorescenza. Quest'ultima viene catturata in tempo reale da un sistema di rilevazione collegato al termociclatore, quantificata ed elaborata dal computer che gestisce tutto l'apparato (vedi Figura 27).



**Figura 27** PCR quantitativa "Real time" mediante utilizzazione di sonde TaqMan (Applied Biosystems). Nel Pannello in alto è indicata la struttura della sonda. Il pannello in basso, riporta invece una rappresentazione schematica del funzionamento di questo tipo di PCR.

Nei nostri esperimenti abbiamo utilizzato sonde TaqMan (Applied Biosystem) per *ACVR1* (ID Mm00431645\_m1 e Hs00153836\_m1) e per la  $\beta$ -actina (ID Mm00607939\_s1 e Hs99999903\_m1) e TaqMan Universal PCR Master Mix (Applied Biosystems). Dopo diverse prove per valutare la diluizione ottimale dei cDNA, i campioni sono stati utilizzati alla diluizione 1:27 e per ognuno sono state eseguite tre repliche.

Il programma impostato nello strumento IQ5 (BioRAD) è stato il seguente: 95 °C per 30'', 95 ° C per 1' (x2), 95 °C per 7' e ripetuto per 40 cicli: 95°C per 10'' e 60°C per 1'.

**Tabella n.°1**

<b>Oligo</b>	<b>Sequenza</b>	<b>T°C</b>
<b><i>Oligo utilizzati per la generazione del 3'UTR e frammenti derivati</i></b>		
ACVR1 3'XBA-F	TCTAGACCCTCGACAAATTGAAAACGAC	64°C
ACVR1 3'XBA-REV	TCTAGACTGGCAGAGTTTAAATGCACG	63°C
ACVR1 3'XBA-F	TCTAGACCCTCGACAAATTGAAAACGAC	64°C
XBA-ARE 1R	TCTAGAGCAAAGTTGTGTATAAACAATTC	59°C
XBA-ARE 1F	TCTAGAGAATTGTTTATACACAACCTTGC	59°C
XBA-ARE 2R	TCTAGAGACTTGAAAACAGTTTATTTAATTT	56.5°C
<b><i>Oligo disegnati sul cDNA di ACVR1 e possibili isoforme</i></b>		
cDNA-FOP4F	AGTGAGAGAAGCTCTGAACG	57°C
cDNA-FOP6R	CGAAGGCAGCTAACTGTATC	57°C
ACVR1-iso F	GAGTAAGCTTTACCCAGGAG	57°C
DA995	GTGGGCTCAGCCTTCTTTCC	61°C
DA436	ACCCGAGGGTCTTACTCTGTC	53°C
4F/Rev	GTCAGAGCTTCTCTCACTC	57°C

**Tabella n.º2**

Codici identificativi pre-miR Ambion utilizzati in questa tesi.

	<b>Micro RNA</b>	<b>miRBase Stem-Loop Accession ID</b>	<b>MiRBase Mature miRNA Accession ID</b>	<b>Mature miRNA Sequence</b>
<b>PM11133</b>	<b>hsa-miR-365</b>	hsa-miR-365 1- MI0000767 hsa-miR-365 2- MI0000769	MIMAT0000710	UAAUGCCCCUAAAAAUCCUUAU
<b>PM12369</b>	<b>hsa-miR-182</b>	MI0000272	MIMAT0000259	UUUGGCAAUGGUAGAACUCACACU
<b>PM12269</b>	<b>hsa-miR-152</b>	MI0000462	MIMAT0000438	UCAGUGCAUGACAGAACUUGG
<b>PM10588</b>	<b>hsa-miR-384</b>	MI0001145	MIMAT0001075	AUUCCUAGAAAUUGUUCAUA
<b>PM10242</b>	<b>hsa-miR-381</b>	MI0000789	MIMAT0000736	UAUACAAGGGCAAGCUCUCUGU
<b>PM10249</b>	<b>hsa-miR-26a</b>	hsa-mir-26a-1 MI0000083 hsa-mir-26a-2 MI0000750	MIMAT0000082 MIMAT0000082	UUCAAGUAAUCCAGGAUAGGCU

## 6 BIBLIOGRAFIA

Andreassi C & Riccio A. **To localize or not to localize: mRNA fate in 3'UTR ends** *Cell press* 2009; **19**: 465-474.

Barreau C, Paillard L & Osborne HB. **AU-rich elements and associated factors: are there unifying principles?** *Nucleic Acids Research* 2005; **33**: 7138-7150

Bartel DP, **MicroRNAs: Genomics, Biogenesis, Mechanism, and Function** *Cell* 2004; **116**: 281-297

Bocciardi R, Bordo D, Di Duca M, Di Rocco M & Ravazzolo R. **Mutational analysis of the ACVR1 gene in Italian patients affected with fibrodysplasia ossificans progressiva: confirmation and advancements.** *European Journal of Human Genetics* 2008; 1-8

Carthew R & Sontheimer EJ **Origins and mechanism of miRNA and siRNA** *Cell* 2009; **136** : 642-655.

Denli AM, Tops BB, Plasterk RH et al **Processing of primary microRNAs by the Microprocessor complex.** *Nature* 2004; **432**: 231-235.

Eberart W, Doller A, Akool ES, Pfeilschifter J. **Modulation of mRNA stability as a novel Therapeutic approach.** *Pharmacology & Therapeutic* 2007; **114**: 56-73.

Gherzi R, Lee KY, Briata P, Wegmüller D. **A KH Domain RNA Binding Protein, KSRP, Promotes ARE-Directed mRNA Turnover by Recruiting the Degradation Machinery** *Molecular Cell* 2004; **14**: 571-583.

He L & Hannon GJ. **MicroRNAs: small RNAs with a big role in gene regulation** *Nature* 2004; **5**: 522-531.

Kaplan FS, Groppe J, Pignolo RJ, Shore EM. **Morphogen receptor genes and metamorphosis: skeleton keys to the metamorphosis.** *Ann N Y Acad Sci*

2007; **1116**:113–133

Kaplan FS, Xu M, Seemann P et al. **Classic and Atypical Fibrodysplasia Ossificans Progressiva (FOP) Phenotypes Are Caused by Mutations in the Bone Morphogenetic Protein (BMP) Type I Receptor ACVR1** *Human mutation* 2008.

Kaplan FS, Shen Q, Lounev V et al. **Skeletal metamorphosis in fibrodysplasia ossificans progressiva (FOP)** *J Bone Metab* 2008; **26**: 521-530.

Lund E, Guttinger S, Calado A et al. **Nuclear export of microRNA precursor.***Science* 2004;**303**: 95-98.

Moore MJ. **From birth to death: the complex lives of eukaryotic mRNAs** *Science* 2005; **309** 1515-1518

Nohe A, Keating E, Knaus P, Peterson NO. **Signal transduction of bone morphogenic protein receptors.** *Cell Sign* 2003; **16**: 291-299

Okamura, K, Phillips MD, Tyler DM. **The regulatory activity of microRNA species has substantial influence on microRNA and 30 UTR evolution.** *Nat. Struct. Mol. Biol* 2008; **15**, 354–363.

Richard W. Carthew, and Erik J. Sontheimer **Origins and Mechanisms of miRNAs and siRNAs** *Cell* 2009; **136**: 642-655

Shi Y & Massague J **Mechanisms of TGF- $\beta$  Signaling from Cell Membrane to the Nucleus** *Cell* 2003; **113**: 685–700

Shore EM, Xu M, Feldman GJ et al. **A recurrent mutation in the BMP type I receptor ACVR1 causes inherited and sporadic fibrodysplasia ossificans progressive.** *Nature genetics* 2006; **38**: 525-527.

Sobell HM. **Actinomycin and DNA transcription** *Proc. Natl. Acad. Sci* 1985; **82**: 5328-5331

Vasudevan S, Tong Y, Steiz JA. **Swiching from repression to activation: microRNAs can up-regulate translation.** *Science* 2007; **318**:1931-1934

Wang T & Donahoe PK. **The immunophilin FKBP12: a molecular guardian of the TGF $\beta$  family type I receptors.***Front Biosci* 2004; **9**: 619-631

Yan X, Lin Z, Chenet F et al. **Human BAMBI Cooperates with Smad7 to Inhibit Transforming Growth Factor-Signaling** *Journal of Biological chemistry* 2009;•**284** 30097-30104.