



UNIVERSITÀ DEGLI STUDI DI UDINE

CORSO DI DOTTORATO DI RICERCA
IN SCIENZE E TECNOLOGIE CLINICHE
CICLO XXIV

Tesi di Dottorato di Ricerca

STUDI DI PROTEOMICA FUNZIONALE ATTI A DEFINIRE I
MECCANISMI REGOLATORI DEL FATTORE DI
TRASCRIZIONE Oct-4A

Dottoranda:
Renza SPELAT

Relatore:
Prof. Francesco CURCIO

ANNO ACCADEMICO 2011 - 2012

Tesi di Dottorato di Renza Spelat discussa presso l'Università degli Studi di Udine

INDICE

| | |
|--------------------------------------------------------------------------------------|------|
| 1- INTRODUZIONE | p.1 |
| 1.1- La cellula staminale | p.1 |
| 1.1.1- Definizione e proprietà | p.1 |
| 1.1.2- I meccanismi molecolari responsabili della staminalità | p.3 |
| 1.1.3- Classificazione delle cellule staminali in base al potenziale differenziativo | p.6 |
| 1.1.4- Classificazione delle cellule staminali in base all'origine | p.8 |
| 1.2- L'espressione genica | p.11 |
| 1.3- La regolazione trascrizionale | p.14 |
| 1.4- I fattori di trascrizione | p.15 |
| 1.4.1- Struttura dei fattori di trascrizione | p.16 |
| 1.4.2- L'omeodominio | p.17 |
| 1.4.3- I geni omeobox | p.19 |
| 1.5- Oct/4 e il "POU domain" | p.20 |
| 1.5.1- La famiglia dei fattori di trascrizione con "POU domain" | p.20 |
| 1.5.2- Caratteristiche del "POU domain" | p.22 |
| 1.5.3- Identificazione e caratterizzazione di Oct/4 | p.26 |
| 1.5.4- Oct/4 e le sue isoforme | p.28 |
| 1.6- Espressione di Oct4 | p.35 |
| 1.6.1- Espressione di Oct/4A | p.35 |
| 1.6.2- Regolazione dell'espressione di Oct/4A | p.37 |
| 1.6.3- Funzioni di Oct/4A | p.39 |
| 1.6.4- Espressione e funzione di Oct/4B | p.44 |
| 2- SCOPO | p.45 |
| 3- MATERIALI E METODI | p.46 |
| 3.1- Modelli cellulari e mantenimento in coltura | p.46 |
| 3.2- Trasfezioni | p.47 |
| 3.3- Clonaggio in vettore di espressione della proteina Oct-4A | p.47 |
| 3.3.1- Estrazione di RNA | p.48 |
| 3.3.2- Reazione di retrotrascrizione | p.48 |
| 3.3.3- PCR (<i>Polymerase Chain Reaction</i>) | p.49 |
| 3.3.4- Elettroforesi su gel di agarosio per acidi nucleici | p.52 |
| 3.3.5- Estrazione da gel | p.53 |
| 3.3.6- Quantificazione su gel d'Agarosio | p.55 |
| 3.3.7- Il vettore di espressione procariota pGEX4T1 | p.55 |
| 3.3.8- Digestione e defosforilazione del vettore | p.56 |
| 3.3.9- Ligasi vettore-inserito | p.57 |
| 3.3.10- Preparazione di cellule competenti | p.58 |
| 3.3.11- Trasformazione dei batteri E. coli BL21 | p.58 |
| 3.4- Miniprep | p.59 |
| 3.5- Miniprep eseguite mediante l'utilizzo del kit Eurogold (Euroclone) | p.60 |
| 3.5.1 Analisi di restrizione | p.61 |

| | |
|-------------------------------------------------------------------------------------------------------------|-------|
| 3.6- Midiprep | p.62 |
| 3.7- Maxiprep | p.64 |
| 3.8- Sequenziamento | p.65 |
| 3.9- Trasformazione del ceppo batterico KRX | p.66 |
| 3.10- Induzione dell'espressione proteica | p.66 |
| 3.11- Saggio di GST Pull-Down | p.67 |
| 3.11.1- Preparazione del <i>Bait</i> | p.67 |
| 3.11.2- Equilibratura della resina | p.68 |
| 3.11.3- Immobilizzazione del <i>Bait</i> | p.68 |
| 3.11.4- Preparazione del <i>Prey</i> | p.69 |
| 3.11.5- Cattura del <i>Prey</i> | p.70 |
| 3.11.6- Eluizione <i>Bait-Prey</i> | p.70 |
| 3.12- Frazionamento nucleo-citoplasmatico | p.72 |
| 3.13- Co-immunoprecipitazione | p.73 |
| 3.14- Immunoblotting | p.75 |
| 3.14.1- Elettroforesi su gel di poliacrilamide con Sodio Dodecil Solfato (SDS-PAGE) | p.75 |
| 3.14.2- Blotting | p.76 |
| 3.15- Immunofluorescenza | p.77 |
| 3.16- Analisi di Spettrometria di Massa | p.79 |
| 3.16.1- Dai dati dell'analisi spettrometrica all'identificazione del polipeptide | p.79 |
| 3.16.2- Digestione del campione | p.84 |
| 3.16.3- Identificazione con "MS peptide mapping" e analisi della sequenza | p.84 |
| 3.16.4- Arricchimento dei fosfo-peptidi tramite purificazione con biossido di titanio (TiO ₂) | p.85 |
| | |
| 4- RISULTATI | p.86 |
| | |
| 4.1- Verifica dell'espressione del fattore di trascrizione Oct-4A nelle cellule Ntera2 | p.86 |
| 4.1.1- Immunofluorescenza | p.86 |
| 4.1.2- Immunoblotting | p.87 |
| 4.2- Analisi della sequenza aminoacidica di Oct-4A | p.87 |
| 4.3- Erk interagisce <i>in vivo</i> e <i>in vitro</i> con Oct-4A | p.90 |
| 4.3.1- Saggio del GST Pull-Down: amplificazione del cDNA che codifica per la proteina Oct-4A | p.90 |
| 4.3.2- Saggio del GST Pull-Down: analisi di restrizione dei cloni trasformati con il vettore pGEX4T1-Oct-4A | p.92 |
| 4.3.3- Saggio del GST Pull-Down: sequenziamento del plasmide pGEX4T1-Oct-4A | p.93 |
| 4.3.4- Saggio del GST Pull-Down: induzione dell'espressione di Oct-4A | p.94 |
| 4.3.5- GST Pull-Down | p.97 |
| 4.3.6- Co-immunoprecipitazione | p.98 |
| 4.4- Erk media la down-regolazione di Oct-4A | p.98 |
| 4.5- L'attivazione di Erk induce l'esclusione nucleare di Oct-4A | p.101 |
| 4.6- L'attivazione di Erk incrementa i livelli di ubiquitinazione di Oct-4A | p.103 |

| | |
|--------------------------------------------|-------|
| 4.7- Oct-4A è fosforilata sulla Serina 111 | p.104 |
| 5- DISCUSSIONE | p.108 |
| 6- BIBLIOGRAFIA | p.113 |

1- INTRODUZIONE

Il fattore di trascrizione *Pou5f1*, meglio conosciuto con il nome *Oct-4* (*Oct-3*), appartiene alla famiglia di fattori di trascrizione che contengono il dominio *POU*. Il dominio *POU* è costituito da una sequenza che risulta altamente conservata tra i membri di questa famiglia e rende possibile il legame della proteina al DNA .

L'espressione della proteina *Oct-4* è stata dimostrata a livello delle cellule staminali embrionali e a livello delle cellule germinali, nonché delle cellule staminali adulte; questo fattore di trascrizione riveste un ruolo fondamentale come regolatore trascrizionale di geni coinvolti nel mantenimento di uno stato pluripotente indifferenziato e di *self-renewal* cellulare, impedendo l'espressione di geni attivati durante il differenziamento.

Per riuscire a comprendere a pieno le caratteristiche e la funzionalità del fattore di trascrizione *Oct-4*, è dunque opportuno porre le basi dei concetti cardine riguardanti la staminalità.

1.1- La cellula staminale

1.1.1- Definizione e proprietà

Sia da un punto di vista di ricerca scientifica che da un punto di vista etico, le cellule staminali risultano essere un argomento di grande attualità.

Negli anni passati solo le cellule embrionali venivano considerate staminali, in quanto dotate di capacità proliferative e differenziative tali da permettere la formazione di un nuovo organismo completo e funzionale; con il progredire degli studi scientifici invece, si è scoperto e dimostrato che è possibile ottenere cellule staminali da numerosi tessuti e organi del corpo umano.

Anche se nel corso degli anni sono state formulate svariate definizioni di cellula staminale, la comunità scientifica è attualmente concorde nel ritenere che una cellula staminale, che sia di origine embrionale, fetale o adulta, deve possedere, in certe condizioni fisiologiche o sperimentali, la capacità di riprodurre se stessa per lunghi periodi di tempo o,

nel caso delle cellule staminali adulte, per tutta la durata della vita dell'organismo (Smith, 2006).

Per definire una popolazione cellulare staminale, occorre che questa soddisfi, tramite saggi sperimentali *in vitro* e *in vivo*, una serie di criteri considerati assolutamente indispensabili.

Questi criteri sono (Verfaillie et al., 2002) :

- la **clonogenicità**: la capacità di una singola cellula staminale di dare origine ad una colonia di cellule *in vitro*;
- la **multipotenzialità**: la possibilità di una singola cellula staminale di dare origine a più di un tipo cellulare differenziato *in vitro*;
- l'**estensiva capacità di crescita**: la possibilità di amplificare una cellula staminale *in vitro*, per lungo tempo, mantenendone lo stato indifferenziato e senza che la popolazione da essa originata mostri segni evidenti di senescenza. Le cellule mantengono quindi uno stato indifferenziato, non essendo commissionate verso alcuna linea cellulare (Lovell-Badge R., 2001), ed esprimono fattori di trascrizione tipici quali Nanog, Sox2 e *Oct4* che hanno un ruolo cruciale nei primi stadi di sviluppo embrionale (Niwa et al., 2000)
- l'**autorinnovo**: la capacità di una cellula staminale di riprodurre se stessa per lunghi periodi di tempo, sia *in vitro* sia *in vivo*. Le cellule staminali adulte sono in grado di rinnovarsi per tutta la vita dell'organismo mediante un particolare tipo di proliferazione chiamata *divisone asimmetrica*. Questa divisone fa sì che, delle due cellule figlie, una si mantenga nello stato indifferenziato mentre l'altra si avvia a maturazione e specializzazione;
- la **riparazione tissutale**: definita come la proprietà, *in vivo*, di una cellula staminale di riportare nel tessuto danneggiato la *restitutio ad integrum*, ovvero il ripristino morfofunzionale dei tessuti danneggiati senza cicatrizzazione;
- il **turnover**: la capacità delle cellule staminali di garantire il normale ricambio cellulare, sostituendo le perdite fisiologiche che sono comunque presenti anche in assenza di danno tissutale. Sia la riparazione che il turnover sono processi che possono avvenire grazie alla capacità delle cellule staminali di proliferare e differenziarsi. Per differenziazione si intende quel processo che, partendo da una cellula indifferenziata, porta alla generazione di una cellula specializzata;
- la **plasticità**: intesa come la capacità di una cellula staminale di differenziarsi verso linee

cellulari distinte rispetto a quelle del tessuto da cui è stata prelevata (Weissman, 2001). Diversi studi hanno infatti dimostrato che le popolazioni di cellule staminali adulte sono in grado, dopo esposizione ad un diverso ambiente di coltura, di differenziare e ripopolare tessuti diversi rispetto a quello di origine (Wagers, 2004).

1.1.2- I meccanismi molecolari responsabili della staminalità

Molte delle caratteristiche peculiari che si riscontrano nei diversi tipi di cellule staminali sono determinate dall'attività di un gruppo di fattori di trascrizione tra cui Stat3, Oct-4 (Niwa et al., 2000), Nanog (Chambers et al., 2003), FoxD3 e Sox2 (Avilion et al., 2003).

Ad esempio, nelle cellule staminali embrionali e neurali, la presenza di questi fattori di trascrizione è indispensabile per mantenere elevata la capacità di auto-rinnovo; infatti è lo stesso complesso Sox2-Oct-4 che è in grado di legare le sequenze *enhancers* di alcuni geni target, come FGF-4 e UTF1 (Nishimoto et al., 2003), indispensabili per il mantenimento del *self-renewal* e della sopravvivenza delle cellule staminali embrionali e neurali (Fig. 1.1).

In vitro, le cellule staminali embrionali mostrano un'apparente illimitata capacità di auto-rinnovamento garantita da processi di divisione simmetrica, che consistono nella generazione di cellule figlie identiche alla cellula madre oppure che possono essere commissionate verso una determinata linea cellulare, regolati dai fattori di trascrizione Oct-4, Nanog e Sox2 e per questo considerati appunto specifici di uno stato di pluripotenza.

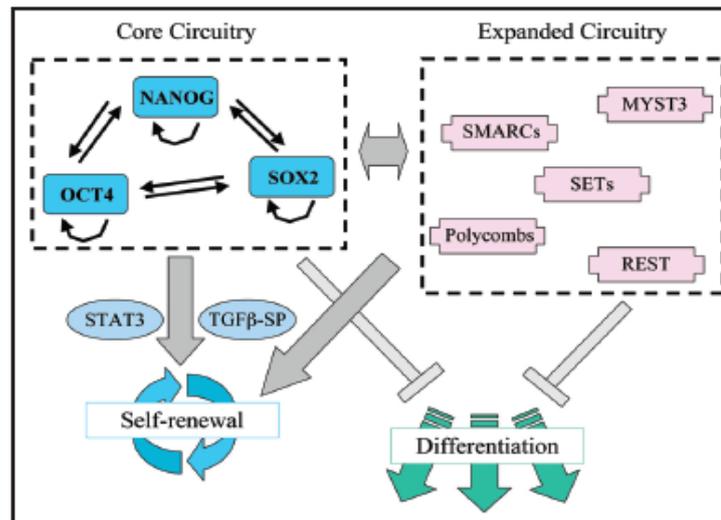


Fig. 1.1- Regolazione del self renewal e della differenziazione nelle ESC. I fattori di trascrizione Oct4, Nanog e Sox2 si attivano l'un l'altro e attivano la loro stessa trascrizione. Attraverso STAT3 e TGFβ-SP promuove self-renewal e blocca la differenziazione. Tutte le funzioni sono supportate dall'interazione con i regolatori della cromatina del "Expanded Circuitry". Schultz W.A. et al., epigenetics 2007.

Nanog sembra giocare un ruolo cruciale principalmente nella seconda fase dello sviluppo dello stadio di preimpianto: questa proteina è stata infatti trovata all'interno delle cellule allo stadio di *morula*, nell'*inner cell mass (IMS)*, e sembra che la sua espressione continui fino alla formazione dell'epiblasto, nella blastocisti tardiva, prevenendo così la sua differenziazione a endoderma primitivo. L'espressione di Nanog inizia ad essere regolata negativamente subito dopo l'impianto dell'embrione nell'endometrio uterino (Pan G. et al., 2007) (Fig. 1.2).

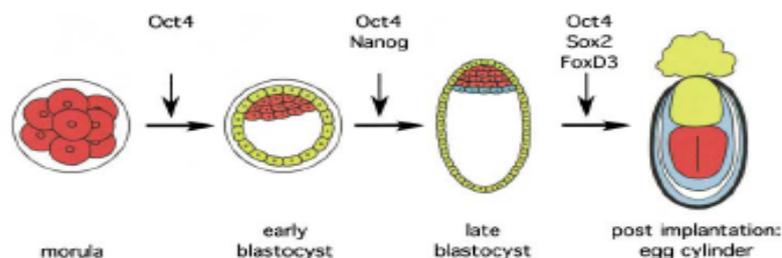


Fig.1.2- Modello del sistema di funzionamento dei fattori di trascrizione negli stadi precoci di sviluppo nel topo e nelle cellule staminali embrionali (ESC). Chambres at al., Cell 2003.

Un altro fattore di trascrizione molto importante nel mantenimento della staminalità cellulare è Sox2: anch'esso è espresso nelle cellule staminali embrionali dell'IMS e nell'epiblasto, che viene quindi mantenuto indifferenziato sia dall'azione di Nanog che da quella di Sox2. E' stato inoltre dimostrato che l'*enhancer* di Sox2 è controllato, almeno in parte, da un'altro fattore di trascrizione fondamentale per il mantenimento della staminalità, Oct-4, durante la transizione da epiblasto a ectoderma neurale (Catena et al., 2004) (Fig. 1.2).

L'espressione della proteina Oct-4 è necessaria sia per decidere il destino delle cellule allo stadio di morula sia nelle fasi successive, quando si ha la formazione dell'epiblasto e dell'ectoderma primitivo. Infatti, embrioni omozigoti per la delezione di Oct-4 (Oct-4^{-/-}), muoiono subito dopo l'impianto, prima che si abbia la formazione dell'epiblasto. Si è inoltre visto che l'espressione di questa proteina è essenziale anche, *in vitro*, per il mantenimento della pluripotenzialità delle cellule dell'ICM.

Tutte le proprietà delle cellule staminali sono inoltre garantite da un preciso controllo sui meccanismi di divisione cellulare. Infatti le cellule staminali si possono dividere secondo due diversi meccanismi, che sono:

- la *divisione asimmetrica*;
- la *divisione simmetrica* (Fig. 1.3);

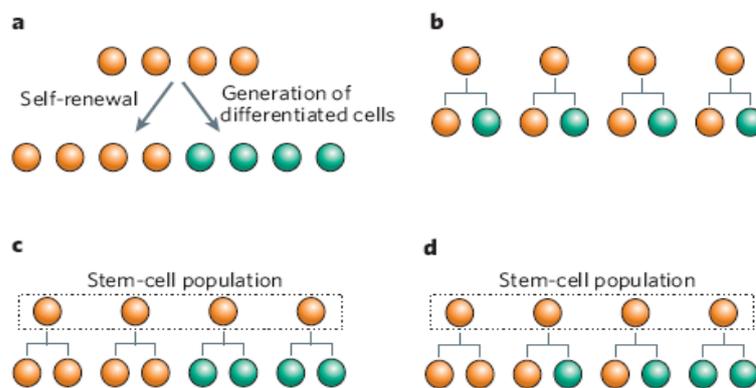


Fig. 1.3- Divisione simmetrica e asimmetrica delle cellule staminali. Morrison S.J. et al., 2006.

La *divisione asimmetrica* è un processo di divisione cellulare ristretta solamente alle cellule staminali più primitive ed è infatti fondamentale per garantire le proprietà di auto-rinnovo e di differenziazione caratteristiche delle cellule staminali. Essa porta alla formazione di due differenti tipi di cellule figlie: una identica alla cellula madre, e per tanto staminale, che garantisce l'auto-rinnovamento, l'altra, invece, destinata ad andare incontro a maturazione e quindi a differenziazione. La divisione asimmetrica è un processo molto importante perché garantisce il mantenimento di un *pool* numericamente costante di cellule staminali (Morrison J.S. et al., 2006).

La *divisione simmetrica*, invece, consiste nella generazione di due cellule figlie, che possono essere identiche alla cellula madre oppure possono essere commissionate verso una determinata linea cellulare. Questo meccanismo di divisione cellulare consente un'espansione del *pool* di cellule staminali e diversi studi hanno dimostrato che esso non interviene solo durante lo sviluppo degli organismi vertebrati ed invertebrati, ma anche nei processi di riparazione e rigenerazione che coinvolgono i tessuti adulti. La capacità di divisione simmetrica, che consente di modulare le dimensioni del pool di cellule staminali, garantisce plasticità nello sviluppo e aumenta la crescita rigenerativa dei tessuti (Morrison J.S. et al., 2006).

L'equilibrio tra questi due meccanismi di divisione cellulare è controllato dai segnali che presiedono allo sviluppo e che si trovano nel microambiente che ospita le cellule staminali; la perdita della capacità di regolare tali meccanismi di divisione delle cellule staminali potrebbe essere la causa che sta alla base di diversi processi patologici, tra cui quelli neoplastici.

1.1.3- Classificazione delle cellule staminali in base al potenziale differenziativo

In base al loro potenziale differenziativo, le cellule staminali possono essere distinte in:

- ***totipotenti***: capaci di dare origine a tutte le cellule e a tutti i tessuti di un organismo animale, comprese le cellule che non fanno prettamente parte dell'embrione, come quelle che costituiscono la placenta e il cordone ombelicale. Solamente lo zigote e le cellule che

derivano dalle prime divisioni mitotiche successive alla fecondazione dell'ovocita, sono considerate totipotenti.

- **pluripotenti:** sono cellule che hanno perduto la capacità di generare i tessuti degli annessi embrionali (Alison et al., 2002), ma che mantengono la capacità di originare cellule appartenenti ai tre foglietti germinativi (ectoderma, mesoderma ed endoderma). Vengono considerate pluripotenti le cellule staminali embrionali.
- **multipotenti:** appartengono a questa categoria le cellule che sono capaci di dare origine alle linee cellulari ristrette ad un determinato foglietto embrionale. Molte cellule staminali dell'adulto, come per esempio le cellule staminali somatiche, vengono definite multipotenti: non sono capaci di originare cellule germinali ma da esse derivano i progenitori commissionati per lo sviluppo di una sola linea cellulare.
- **unipotenti:** sono cellule staminali in grado di dare origine ad un unico tipo di cellula specializzata.

Non rientrano invece nel gruppo delle cellule staminali i progenitori: sono cellule parzialmente specializzate del feto e dell'adulto che sono in grado di proliferare estensivamente e di dare origine a cellule differenziate, ma che non sono capaci di auto-rinnovarsi per lunghi periodi. Queste cellule partecipano a meccanismi fisiologici di rimpiazzo di cellule morte o danneggiate, garantendo così l'integrità e la funzionalità del tessuto (Blau et al., 2001).

Da quanto detto sino ad ora emerge, quindi, che all'interno dei tessuti esiste un preciso ordine gerarchico: all'apice della scala gerarchica è situata la cellula maggiormente indifferenziata del sistema che, dividendosi, dà origine ad altre cellule staminali ed elementi progressivamente più maturi; da questi, a loro volta, deriveranno tutte le linee differenziate che si posizionano alla base della piramide gerarchica.

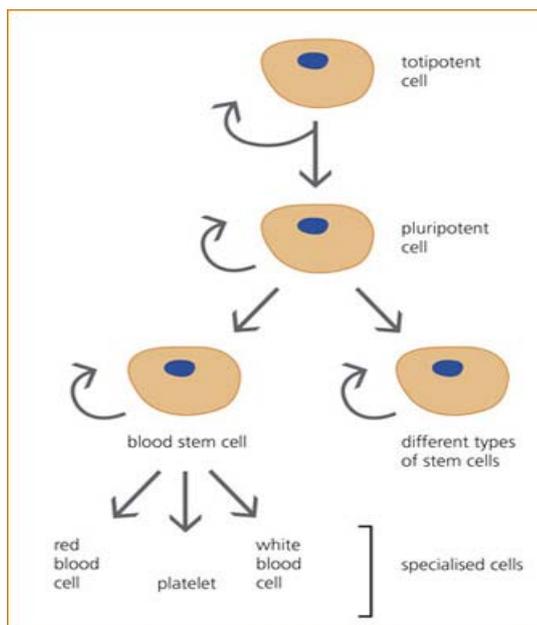


Fig. 1.4- Piramide gerarchica delle cellule staminali. www.xcell-center.it

1.1.4- Classificazione delle cellule staminali in base all'origine

Le cellule staminali possono essere classificate, in base alla loro origine, in:

- **cellule staminali embrionali**
- **cellule staminali adulte**

Cellule staminali embrionali. La cellula staminale indispensabile nello sviluppo embrionale umano è lo zigote, che consiste in una singola cellula dotata di tutte le istruzioni e capacità necessarie a differenziare in ogni tipo di cellula presente nell'organismo. Durante le prime ore trascorse dalla fecondazione dell'ovocita, lo zigote subisce la prima mitosi che porta alla formazione di due cellule figlie, le quali, a loro volta, si dividono ripetutamente a formare ammassi costituiti da 16 cellule, conosciuti con il nome di *morula*. Allo stadio di sedici-trentadue cellule, nella morula si forma una cavità, detta *blastocoele*, e a questo livello maturativo l'embrione viene definito *blastocisti*. Allo stadio di blastocisti, la distribuzione

delle cellule non è più uniforme: si può distinguere chiaramente un sottile strato esterno di cellule pavimentose, il *trofoectoderma*, che racchiude un piccolo gruppo eccentrico di cellule, la *massa cellulare interna (ICM)*. Il trofoblasto darà origine alla placenta, mentre la massa di cellule situate internamente alla blastocisti darà luogo a tutti i tessuti del corpo, in un processo chiamato *gastrulazione* (Young H.E. et al., 2004).

Comunemente, quando si parla di cellule staminali embrionali, ci si riferisce alle cellule ottenute dalla disgregazione della massa cellulare interna della blastocisti (ICM) dell'embrione dei mammiferi, che acquisiscono *in vitro* estensiva capacità di auto-rinnovo. La blastocisti è lo stadio embrionale che precede l'impianto nell'utero materno, e può contenere fino a 100 cellule non specializzate, pluripotenti e fondatrici dell'intera struttura dell'organismo, nel caso in cui l'embrione riesca ad impiantarsi e a svilupparsi nell'utero materno. Dopo l'impianto, le cellule pluripotenti non sono più in grado di auto-mantenersi e iniziano a differenziarsi in cellule specializzate, che daranno origine a tutti i tessuti dell'organismo adulto (Keller, 2005).

Cellule staminali dell'adulto. Tradizionalmente il concetto di staminalità era sempre stato associato all'embrione, costituito da cellule ancora indifferenziate, destinate a proliferare ampiamente e a generare molteplici tipi cellulari distinti. Molti tessuti degli organismi adulti possiedono una o più popolazioni di cellule non differenziate per molti aspetti simili alle cellule staminali embrionali, che pertanto vengono chiamate cellule staminali dell'adulto.

Le cellule staminali adulte o somatiche possiedono due caratteristiche principali: sono capaci di duplicarsi per lunghi periodi e possono dare origine a cellule mature con morfologia e funzionalità specifiche. La presenza delle cellule staminali adulte si verifica a partire dal VI° mese di gestazione e la loro funzione riguarda principalmente il mantenimento dell'omeostasi tissutale, la sostituzione delle cellule danneggiate e morte a causa di invecchiamento tissutale, disfunzioni o danni (Leblond, 1964). Esse possono essere residenti, integrate nell'architettura dell'organo, oppure circolanti nel flusso ematico.

A differenza di quelle embrionali, le cellule staminali dell'adulto hanno una capacità di differenziamento più ristretta, dato che spesso sono limitate ai soli tipi cellulari del tessuto di origine, e *in vitro* tendono a differenziare molto più rapidamente o ad andare incontro a senescenza. Le cellule staminali dell'adulto consentono, a differenza di quelle embrionali, trapianti autologhi che non necessitano quindi di immunosoppressione: infatti possono essere isolate da un paziente, espanse in coltura e poi re-iniettate nello stesso paziente senza dare

problemi di rigetto.

E' ormai noto che le cellule staminali adulte sono presenti in quasi tutti i tipi di tessuti, inclusi la pelle (Alonso L. et al., 2003), il sistema nervoso (Gage F.H., 2000), il muscolo scheletrico, il midollo osseo, il cuore, il fegato, il pancreas, le ghiandole mammarie, il follicolo pilifero, l'intestino (Brittan M. et al., 2004) e la polpa dentale (Kerkis I. et al., 2006).

L'identificazione di cellule staminali non solo in organi ad alto *turnover*, quali la cute o l'intestino, ma anche in organi a basso *turnover*, quali il cuore o il cervello, ha esteso in pratica a tutti i distretti tissutali il modello gerarchico del midollo osseo, che prevede l'esistenza di rare cellule staminali lentamente ciclanti in grado di dividersi mantenendo il *pool* staminale e dando altresì origine a progenitori altamente ciclanti e commissionati allo sviluppo di una sola linea cellulare (Young H.E. et al., 2004b).

Il comportamento delle cellule staminali è fortemente influenzato dal microambiente tissutale da cui originano e in cui risiedono (Alonso L. et al., 2003) e quindi dall'esistenza di nicchie, strutture tissutali specializzate in grado di regolare proliferazione e differenziazione delle cellule staminali in esse situate.

1.2- L'espressione genica

L'espressione genica rappresenta quel processo grazie al quale viene creato il prodotto di un gene (un RNA o un polipeptide). La produzione di un polipeptide, a partire dalle istruzioni contenute nel DNA di un gene, prevede due passaggi: la trascrizione e la traduzione.

Durante la trascrizione un enzima chiamato *RNA polimerasi*, realizza una copia di uno dei due filamenti di DNA: la copia ottenuta è una molecola di RNA. In questo modo le informazioni contenute all'interno del DNA vengono "convertite" enzimaticamente in RNA complementare. Nel corso della traduzione, il filamento di RNA messaggero (mRNA) porta le istruzioni genetiche ai siti di sintesi delle proteine nelle cellule, i ribosomi, ed è usato come stampo per la produzione di una specifica proteina. I ribosomi leggono la sequenza di un mRNA in triplette, ovvero tre basi alla volta, e portano il corrispondente amminoacido a legarsi alla catena della proteina nascente. Quando i ribosomi raggiungono un segnale di stop rilasciano la proteina completata.

La trascrizione è il primo passo nell'espressione di un gene e consiste di tre fasi: inizio,
 Tesi di Dottorato di Renza Spelat discussa presso l'Università degli Studi di Udine

allungamento e terminazione. Nella fase di inizio della trascrizione, l'RNA polimerasi si lega al promotore del gene e causa una localizzata denaturazione di un breve tratto di DNA ; successivamente comincia a sintetizzare la catena di RNA unendo i primi nucleotidi all'RNA tramite la formazioni di legami fosfodiesteri. Il primo nucleotide conserva il gruppo trifosfato. Durante la fase di allungamento, la bolla di denaturazione del DNA si muove con la polimerasi permettendo all'enzima di leggere le basi dello stampo di DNA e di produrre un RNA complementare ed esso in direzione 5' → 3'. Infine la terminazione ha luogo quando la polimerasi raggiunge un segnale di terminazione, che determina la dissociazione della polimerasi e del prodotto di RNA dallo stampo di DNA.

Il processo di traduzione, chiamato anche sintesi proteica, costituisce la seconda fase del processo di espressione genica, ovvero il processo in cui l'informazione contenuta nel DNA dei geni viene convertita in proteine che svolgono nelle cellule un'ampia gamma di funzioni. Nella sintesi proteica, il filamento di RNA, ottenuto per trascrizione del DNA, è utilizzato come stampo per la produzione di una specifica proteina. Anche il processo della traduzione consta di tre fasi: inizio, crescita e termine del polipeptide. L'inizio della sintesi vede i ribosomi legarsi al codone di avvio dell'mRNA, che indica il punto in cui l'mRNA comincia a codificare la proteina. Il tRNA iniziatore accoppia le sue basi con quelle del codone di avvio e si lega al sito P del ribosoma formando un ponte tra la subunità minore e quella maggiore del ribosoma, che creano così un complesso. Durante la fase di allungamento, un nuovo aminoacil-tRNA, complessato con il fattore di allungamento TU, entra nel sito A del ribosoma e accoppia le sue basi con quelle dell'mRNA. Grazie all'azione peptidil-transferasica della subunità maggiore, viene creato un legame peptidico tra gli amminoacidi vicini. Non appena si forma il legame, l'amminoacido sul sito P si stacca dal suo tRNA e la catena polipeptidica in crescita si lega al tRNA sul sito A. Il ribosoma quindi si muove lungo l'mRNA spostando il peptidil-tRNA dal sito A al sito P, liberando nel contempo il tRNA vuoto. Questo processo è noto con il nome di *traslocazione*. La terminazione avviene nel momento in cui il ribosoma incontra uno dei tre possibile codoni di arresto (codoni di stop), UAG, UAA e UGA. La crescita della proteina si interrompe ed i fattori di rilascio, proteine che simulano l'azione del tRNA, si legano al sito A e liberano così la proteina nel citoplasma della cellula.

Ogni organismo contiene lo stesso codice genetico ma questa serie uguale di istruzioni genetiche può portare a differenti tipi di cellule. L'espressione differenziale di gruppi di geni a partire da un uguale repertorio è proprio quel meccanismo molecolare che sta alla base della

diversità cellulare. La regolazione dell'espressione genica è di fondamentale importanza perché permette ad una cellula di controllare le proprie funzioni, come la differenziazione cellulare, la morfogenesi o i vari processi di adattamento.

Alcuni geni vengono espressi costitutivamente all'interno della cellula; questi sono geni ubiquitari, *housekeeping*, che nella maggior parte dei casi sono espressi in tutte le cellule che costituiscono un organismo. Questo gruppo di geni determina le funzioni cellulari di base, come la sintesi degli acidi nucleici e delle proteine, il trasporto e il metabolismo dei nutrienti, la biosintesi del citoscheletro e degli organelli. Altri geni sono definiti inducibili perché sono espressi solo in certe condizioni ambientali o in determinati stadi dello sviluppo in risposta alle necessità della cellula e dell'organismo. I geni specializzati variano da cellula a cellula e definiscono le caratteristiche peculiari di ciascun tipo cellulare.

Comunque tutti i geni sono costantemente regolati. Il controllo dell'espressione genica rappresenta un aspetto essenziale nello sviluppo ed esso può avvenire a diversi livelli:

1. *Trascrizione differenziale dei geni*, regola quali dei geni nucleari vengono trascritti in RNA e la quantità di trascritti prodotta.
2. *Controllo del processamento dell'RNA*, regola la produzione di mRNA maturo attraverso la scelta del sito di poliadenilazione e la scelta del sito di splicing.
3. *Controllo del trasporto dell'RNA* selezionando quali mRNA debbano essere esportati dal nucleo al citoplasma.
4. *Controllo della stabilità dell'mRNA* destabilizzando selettivamente certe molecole di mRNA nel citoplasma.
5. *Controllo traduzionale della biosintesi proteica* selezionando quali mRNA debbano venire tradotti dai ribosomi.
6. *Controllo post-traduzionale*, attivando o inattivando selettivamente molecole proteiche.
7. *Compartimentalizzazione delle proteine*.

Per la maggior parte dei geni un punto di controllo importante è la regolazione trascrizionale.

1.3- La regolazione trascrizionale

Il controllo trascrizionale è la principale modalità di regolazione dell'espressione genica negli eucarioti così come nei batteri. Esso contribuisce a delineare l'insieme delle proteine espresse da una cellula e, più in generale, le funzioni caratteristiche di una cellula. Ciò consente, in particolare negli organismi pluricellulari, di esprimere specifici geni nelle cellule appropriate durante la crescita, lo sviluppo e il differenziamento cellulare.

A partire dalla scoperta dell'operone del lattosio in *E. coli* (Jacob and Monod, 1961), sono stati descritti molti meccanismi di regolazione trascrizionale dell'espressione genica sia in organismi eucarioti che in quelli procarioti.

Malgrado inizialmente fosse diffuso un modello che metteva in relazione le variazioni di attività trascrizionale a singoli fattori di trascrizione, attualmente è accettata l'idea che la regolazione trascrizionale si avvalga di meccanismi integrati in cui cooperano diversi fattori di trascrizione.

Negli eucarioti il sistema di regolazione della trascrizione si compone di molteplici regioni di controllo costituite da sequenze in *cis*, così definite poiché sono parte della stessa molecola di DNA in cui è localizzato il gene. All'interno del complesso promotore-*enhancer* del DNA genomico, alcuni elementi di controllo sono collocati vicino al sito d'inizio della trascrizione mentre altri sono molto distanti; entrambi possono generalmente stimolare o inibire la trascrizione anche se posizionati invertiti. Le proteine regolatrici quindi possono agire a distanza, legandosi alle sequenze di controllo poste anche a migliaia di nucleotidi dal promotore ed influenzando la trascrizione. Il DNA compreso tra l'*enhancer* e il promotore si ripiega in un'ansa, per consentire alla proteina regolatrice di entrare in contatto con la RNA polimerasi o con uno dei fattori di trascrizione legati al promotore. Se la proteina è un attivatore, favorisce la formazione del complesso d'inizio e aumenta l'attività della RNA polimerasi; se invece è un repressore rende difficile o impossibile la formazione del complesso, con conseguente riduzione o blocco dell'attività della RNA polimerasi (Weaver,

Tesi di Dottorato di Renza Spelat discussa presso l'Università degli Studi di Udine

2005).

I fattori di trascrizione vengono dunque visti come elementi di complessi multiproteici che regolano l'attività trascrizionale in maniera combinata, ossia più proteine regolano un solo gene in modo da integrare molti e differenti segnali cellulari. La regolazione della trascrizione è quindi stabilita da un complesso multiproteico chiamato *enhanceosoma* (Merika and Thanos, 2001).

1.4- I fattori di trascrizione

I fattori di trascrizione sono proteine in grado di riconoscere sequenze bersaglio presenti nelle regioni *enhancer* o nei promotori e agiscono per attivare o reprimere la trascrizione di un particolare gene. I fattori di trascrizione possono essere suddivisi in tre gruppi:

- I fattori di trascrizione *generali* sono implicati nel reclutamento e nel corretto posizionamento dell'RNA polimerasi sul sito d'inizio della trascrizione e sono necessari per l'inizio della trascrizione a livello di tutti i promotori. Partecipano alla formazione del complesso basale della trascrizione.
- I fattori di trascrizione *promotore-specifici*, riconoscono siti a monte del sito d'iniziazione e vi si legano attraverso un legame sequenza-specifico al DNA. Agiscono solo su un sottogruppo di promotori, quelli che contengono la sequenza consenso, accrescendo l'efficienza di trascrizione e garantendo un adeguato livello funzionale del promotore.
- I fattori di trascrizione *tessuto-specifici*, sono presenti solo in pochi tipi cellulari e, dunque, partecipano alla regolazione di geni espressi esclusivamente in quei tessuti. Grazie ai fattori di trascrizione tessuto-specifici ogni cellula esprime uno specifico sottoinsieme dei suoi geni, diverso da quello espresso in altre cellule, a seconda del proprio stato proliferativo e della propria posizione spazio-temporale.

Questi fattori hanno quindi un ruolo fondamentale per l'espressione del fenotipo differenziato di ogni cellula.

I fattori *generali* di trascrizione di per sé indicano il punto di inizio e la direzione della trascrizione, ma sono in grado di sostenere solo un livello basale di trascrizione. Poiché la

trascrizione dei geni attivi nelle cellule è più alta del livello basale, le cellule eucariotiche possiedono appunto fattori di trascrizione gene-specifici (attivatori) che si legano agli *enhancers* fornendo i livelli elevati di trascrizione necessari.

Alcuni attivatori sono in grado di mettere in moto la macchina trascrizionale per interazione diretta sui promotori, molti altri invece esplicano la loro funzione grazie alla mediazione di coattivatori trascrizionali. Tali coattivatori sono proteine che intervengono nel contatto diretto dei fattori di trascrizione sia con il promotore che con il complesso della RNA polimerasi. Essi agiscono nell'ambito di complessi multiproteici costituiti da diverse unità; l'eterogeneità del complesso, e quindi la combinazione dei coattivatori, risulta selettiva non solo per tipo cellulare ma anche per complesso trascrizionale.

Pertanto, i fattori di trascrizione rappresentano le fondamenta di una struttura caratterizzata da molecole accessorie che fanno da ponte tra la struttura cromatinica aperta ed il complesso di trascrizione basale. Le varie combinazioni di fattori e cofattori danno origine a complessi multiproteici differenti che, in presenza di numerose sequenze *cis*-regolatorie a monte dei geni in un dato tipo cellulare, conferiscono un controllo spaziale e temporale sull'espressione genica. E' ciò che si verifica, per esempio, nel corso del differenziamento cellulare: vengono assemblati diversi complessi che regolano direttamente la costruzione dei complessi d'inizio della trascrizione sui promotori dei geni *target* e la frequenza con cui la trascrizione stessa è avviata. Alla base del differenziamento cellulare c'è quindi la differente combinazione di proteine specificamente espresse in ciascun tipo cellulare.

1.4.1- Struttura dei fattori di trascrizione

I fattori di trascrizione possono sia stimolare che inibire la trascrizione da parte dell'RNA polimerasi e possiedono strutture composte da almeno due domini funzionali:

- Il dominio di *legame* al DNA, fondamentale per consentire il legame specifico del fattore di trascrizione alle sue sequenze bersaglio.
- Il dominio di *attivazione* che stimola la trascrizione interagendo con altri fattori e cooperando all'assemblaggio del complesso di trascrizione sul promotore.

Alcuni fattori di trascrizione hanno anche un dominio di dimerizzazione che permette loro di legarsi gli uni agli altri, formando omodimeri, eterodimeri o persino multimeri. Altri

presentano anche siti di legame per molecole effettrici come gli ormoni steroidei.

Diversi gruppi di proteine che regolano la trascrizione condividono motivi (configurazioni ricorrenti) di legame al DNA. I motivi più comuni sono:

- *Moduli di zinco* (dita di zinco, Zn finger), servono da piattaforma strutturale per legare il DNA.
- *Motivo elica-giro-elica* (helix-turn-helix: HTH), legano l'ansa maggiore del DNA.
- *Motivi a cerniera di leucine* (bZIP) e
- *Motivi elica-ansa-elica* (bHLH), sono entrambi domini combinati di dimerizzazione e legame al DNA.

A seconda del dominio di legame al DNA posseduto, i fattori di trascrizione vengono classificati in varie famiglie. Residui aminoacidici differenti presenti su fattori trascrizionali appartenenti alla stessa famiglia permettono il legame specifico a sequenze nucleotidiche differenti.

Una grande famiglia di fattori trascrizionali è caratterizzata da un dominio di legame al DNA definito *omeodominio*.

1.4.2- L'omeodominio

Le proteine contenenti un *omeodominio* costituiscono una super-famiglia di fattori di trascrizione caratterizzati da un dominio molecolare di 60 amminoacidi (omeodominio, HD) che si lega a determinate regioni del DNA. In questo modo vengono attivati precisi set di geni che specificano le proprietà di ogni segmento del corpo. Questi fattori di trascrizione sono infatti fondamentali per il corretto sviluppo di un asse corporeo antero-posteriore sia nei vertebrati che negli invertebrati.

L'omeodominio è codificato da una sequenza di DNA di 180 bp altamente conservata e conosciuta come "*homeobox*". Le *homeobox* furono scoperte per la prima volta nel 1984 all'interno dei geni omeotici di *Drosophila Melanogaster* e successivamente in molti altri organismi eucarioti, inclusi vertebrati e uomo (Quian Y.Q. et al. 1989, Gehring W.J. et al. 1994). Attraverso questa regione, la proteina è in grado di legarsi a specifiche sequenze di DNA, regolando l'espressione genica (Laughon A., 1991). La funzione di questi fattori di trascrizione è quindi essenziale per specificare, nel corso dello sviluppo embrionale, il destino

differenziativo di ogni cellula.

Dati strutturali indicano che gli omeodomoni si ripiegano in tre α -eliche: la prima e la seconda elica sono connesse da un esapeptide e orientate in modo anti-parallelo, mentre la terza elica è perpendicolare alle prime due (Fig.1.5). L'elica 3, che come l'elica 2 genera una conformazione a elica-giro-elica (*helix-turn-helix*), è l'elica di riconoscimento ed entra specificamente in contatto con il solco maggiore del DNA (Laughon A., 1991). Le altre due eliche, invece, fanno parte di un *core* idrofobico e sono necessarie all'elica di riconoscimento per raggiungere la corretta posizione rispetto all'intero omeodominio. La regione N-terminale dell'omeodominio è flessibile ed in grado di legare il solco minore del DNA (Quian Y.Q. et al., 1992). Le eliche 1 e 2 rimangono fuori dalla doppia elica, solo la porzione aminoterminale dell'elica 1 entra in contatto con il DNA nel solco minore. Si possono inoltre formare legami cooperativi tra proteine contenenti omeodominio che si legano al DNA su siti multipli; questo fenomeno si pensa conferisca ulteriore specificità nella regolazione dei geni target (Beachy P.A. et al., 1993).

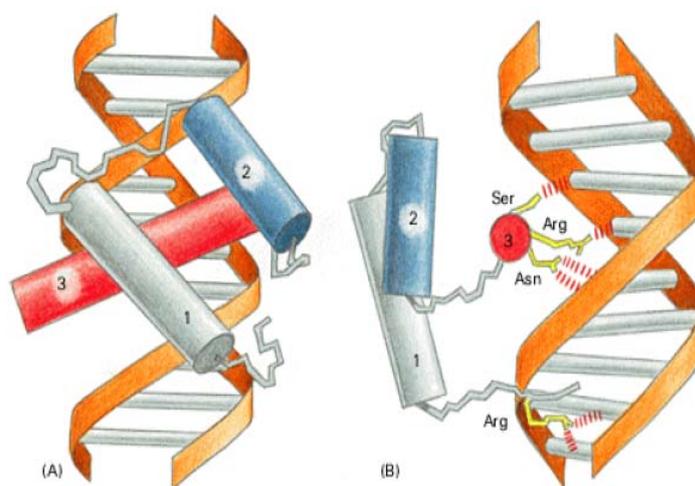


Fig. 1.5- Schema di un omeodominio e legame al DNA

1.4.3- I geni omeobox

L'omeodominio deriva il proprio nome dal fatto di essere stato identificato per la prima

Tesi di Dottorato di Renza Spelat discussa presso l'Università degli Studi di Udine

volta nei geni *omeotici* o geni *HOM* (così detti per la natura delle loro mutazioni) del complesso di *antennapedia* e *bithorax* (*HOM-C*) di *Drosophila*. Nei mammiferi i geni omologhi degli HOM di *Drosophila* sono chiamati geni *Hox* (Furukubo-Tokunaba K. et al., 1993). Il genoma umano e di topo contengono geni che possiedono una omeobox simile a quella del gene *antennapedia* di *Drosophila* e che codificano per dei fattori di trascrizione sequenza-specifici che funzionano da regolatori nella determinazione delle identità regionali lungo gli assi del corpo durante lo sviluppo embrionale (Taylor H.S., 2000).

I geni *Hox* dei vertebrati sono organizzati su quattro cluster (A, B, C, D) localizzati su cromosomi differenti. Come in *Drosophila* anche nei Vertebrati questi geni sono espressi lungo l'asse antero-posteriore (AP) dell'embrione in maniera colineare con la posizione che occupano lungo il cromosoma (Banerjee-Basu S. et al., 2001). Questo vuol dire che i geni posti all'estremità del 3' sono espressi più anteriormente e gli altri via via più posteriormente (Fig.1.6). Mentre il ruolo ancestrale dei geni *Hox* può essere attribuito alla specificazione dell'identità regionale lungo l'asse corporeo AP, nei vertebrati i geni *Hox* sono coinvolti in numerosi altri processi. Ad esempio nel processo di ematopoiesi, geni situati al 3' dei cluster A e B sono espressi in cellule a stadi precoci di differenziazione, mentre quelli presenti all'estremità 5' di questi loci sono attivi in stadi più avanzati della differenziazione (Sauvageau G. et al., 1994). Gli omeogeni sono stati suddivisi in varie classi e sottoclassi in base all'organizzazione all'interno del genoma, alle capacità di legare il DNA e all'omologia di sequenza aminoacidica degli omeodomini da essi codificati. I raggruppamenti più comunemente noti sono: la classe homeobox ANTP, PRD, LIM, POU, HNF, SINE, TALE, CUT, PROS e ZF (Banerjee-Basu S. et al., 2001; Holland P.W.H. et al., 2007).

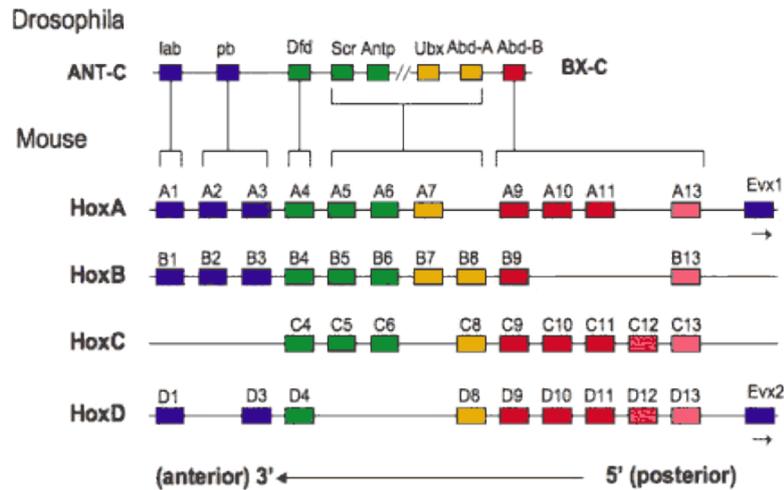


Fig. 1.6- Schema del raggruppamento dei clusters dei geni Hox nel genoma di Drosophila e di topo. I geni Hox nei Vertebrati sono omologhi dei geni homeobox clusterizzati nei complessi antenapedia (Ant) e bithorax (Bx) di Drosophila (HOM-C). www.scienceblogs.com

1.5- Oct-4 e il “POU domain”

E' già stato rimarcato il ruolo chiave del fattore di trascrizione Oct-4 (*Pou5f1*) nel mantenimento di uno stato di pluripotenza delle cellule staminali embrionali e delle cellule germinali e nell'impedimento dell'espressione di alcuni geni attivati durante il processo di differenziamento.

Oct-4 appartiene alla famiglia dei fattori di trascrizione con *dominio POU*, un particolare dominio legante il DNA che risulta altamente conservato tra i membri di questa famiglia di fattori di trascrizione eucarioti.

1.5.1- La famiglia dei fattori di trascrizione con “POU domain”

Le proteine presentanti l'omeodominio (“*Homeodomain proteins*”), che comprendono la famiglia dei fattori di trascrizione con *POU domain*, possono essere suddivise in sei differenti classi sulla base delle loro caratteristiche funzionali e strutturali (Banerjee-Basu S. et al.,

2001). Le omeoproteine regolano diversi programmi di sviluppo modulando pattern di espressione di geni *target* secondo uno schema temporale, spaziale e tessuto-specifico.

La classe di proteine “*POU domain*” include diversi fattori di trascrizione ubiquitari, come le proteine di mammifero *Pit1*, *Oct1*, *Oct2*, *Oct3* e una proteina codificata dal gene *Unc-86* del nematode *Caenorhabditis Elegans*.

Le proteine di mammifero sono implicate nel controllo trascrizionale, mentre *Unc-86* è coinvolto nel processo di differenziamento cellulare.

Pit-1 è un fattore di trascrizione pituitario che regola l'espressione dei geni per la prolattina e per la somatotropina normalmente espressi in due diversi tipi cellulari, le cellule lactotrofe e le cellule somatotrofe, situate nella ghiandola ipofisaria anteriore (Herr W. et al., 1988).

Oct-1 e *Oct-2* sono una coppia di proteine leganti il DNA differenzialmente espresse che riconoscono la stessa sequenza ottamerica, *ATGCAAAT*. Questo motivo ottamerico è un importante elemento cis-regolatore nei promotori di geni espressi in modo ubiquitario e nei promotori cellula-specifici (es. i promotori delle immunoglobuline) (Sturm R.A. et al., 1988). *Oct-1* è espresso in modo ubiquitario e probabilmente rappresenta la proteina legante la sequenza ottamerica già identificata con i nomi di NF-A1 e OTF-1. *Oct-2* è un fattore di trascrizione specifico dei linfociti B e può attivare, *in vitro*, la trascrizione dei geni per le immunoglobuline (Herr W. et al., 1988).

Recentemente Holland e Colleghi (2007) hanno identificato un totale di 16 geni *homeobox* della “classe POU” all'interno del genoma umano (Tab.1.1); questa classe raggruppa sei famiglie di geni ampiamente conosciuti (*Pou1* fino a *Pou6*) la cui nomenclatura, tempo addietro, è stata revisionata in modo tale da assegnare ciascun gene alla propria famiglia (Ryan AK et al., 1997).

| Family | Gene symbol | Gene name | Location | Entrez gene ID | Previous ymbols |
|--------|-------------|-------------------------------------|----------|----------------|-------------------------------------|
| Hdx | HDX | highly divergent homeobox | Xq21.1 | 139324 | CXorf43 |
| Pou1 | POU1F1 | POU class 1 homeobox 1 | 3p11.2 | 5449 | PIT1, GHF1 |
| Pou2 | POU2F1 | POU class 2 homeobox 1 | 1q24.2 | 5451 | OCT1, OTF1 |
| | POU2F2 | POU class 2 homeobox 2 | 19q13.2 | 5452 | OCT2, OTF2 |
| Pou3 | POU2F3 | POU class 2 homeobox 3 | 11q23.3 | 25833 | OCT11, PLA1, EPOC1, SKN1A |
| | POU3F1 | POU class 3 homeobox 1 | 1p34.3 | 5453 | OCT6, OTF6, SCIP |
| | POU3F2 | POU class 3 homeobox 2 | 6q16.2 | 5454 | OCT7, OTF7, BRN2, POUF3 |
| Pou4 | POU3F3 | POU class 3 homeobox 3 | 2q12.1 | 5455 | OTF8, BRN1 |
| | POU3F4 | POU class 3 homeobox 4 | Xq21.1 | 5456 | OTF9, BRN4 |
| | POU4F1 | POU class 4 homeobox 1 | 13q31.1 | 5457 | BRN3A, RDC1, Oct-T1 |
| Pou5 | POU4F2 | POU class 4 homeobox 2 | 4q31.22 | 5458 | BRN3B, BRN3.2 |
| | POU4F3 | POU class 4 homeobox 3 | 5q32 | 5459 | BRN3C |
| | POU5F1 | POU class 5 homeobox 1 | 6p21.33 | 5460 | OCT3, OTF3, OCT4, OTF4 |
| Pou6 | POU5F1 P1 | POU class 5 homeobox 1 pseudogene 1 | 8q24.21 | 5462 | OTF3C, OTF3P1, POU5FLC8 |
| | POU5F1 P2 | POU class 5 homeobox 1 pseudogene 2 | 8q22.3 | 100009665 | |
| | POU5F1 P3 | POU class 5 homeobox 1 pseudogene 3 | 12p13.31 | 642559 | OTF3L, POU5FIL, POU5FLC12, POU5FLC1 |
| | POU5F1 P4 | POU class 5 homeobox 1 pseudogene 4 | 1q22 | 645682 | |
| | POU5F1 P5 | POU class 5 homeobox 1 pseudogene 5 | 10q21.3 | 100009667 | |
| | POU5F1 P6 | POU class 5 homeobox 1 pseudogene 6 | 3q21.3 | 100009668 | |
| | POU5F1 P7 | POU class 5 homeobox 1 pseudogene 7 | 3q12.1 | 100009669 | |
| | POU5F1 P8 | POU class 5 homeobox 1 pseudogene 8 | 17q25.3 | 100009670 | |
| Pou6 | POU5F2 | POU class 5 homeobox 2 | 5q15 | 134187 | SPRM1 |
| | POU6F1 | POU class 6 homeobox 1 | 12q13.13 | 5463 | BRN5, MPOU, TCFB1 |
| | POU6F2 | POU class 6 homeobox 2 | 7p14.1 | 11281 | WTS, WTSL, RPF1 |

Tabella 1.1- Geni homeobox e pseudogeni della “classe POU” nell’uomo e loro suddivisione in famiglie; sono indicati il nome intero, la localizzazione cromosomica, l’”Entrez Gene ID” e il simbolo precedentemente assegnato. (Holland PWH, Booth HAF and Bruford EA. Classification and nomenclature of all human homeobox genes. BMC Biology. 2007; 5(47)).

1.5.2- Caratteristiche del “POU domain”

Il dominio POU è costituito da una sequenza di circa 160 aminoacidi, all’interno della quale si possono distinguere due sotto-domini strutturalmente indipendenti: una regione POU specifica amino-terminale lunga circa 80 aminoacidi (POU_S) e un omeodominio carbossi-terminale lungo 60 aminoacidi (POU_H), interconnessi da una regione spaziatrice ipervariabile di lunghezza compresa tra i 15 e i 50 aminoacidi (Kristie and Sharp, 1990; Verrijzer et al., 1992) (Fig. 1.7). Le regioni esterne al dominio POU non risultano essere critiche per il legame al DNA e mostrano brevi sequenze ripetute.



Fig.1.7- Rappresentazione schematica del dominio POU e le relative posizioni di POU_S, POU_H e del segmento linker. Winship H. et al., 1995.

Per legare il DNA con alta affinità e in modo sito-specifico, i fattori di trascrizione della classe POU richiedono entrambi i domini. La presenza di due sotto-domini di legame al DNA strutturalmente indipendenti rende i membri della classe POU in grado di assumere diverse configurazioni monomeriche sul DNA, ampliando la loro capacità di riconoscere una varietà considerevole di target. Queste proteine possono anche formare omodimeri o eterodimeri in modo da accrescere ulteriormente il loro repertorio funzionale (Ryan A.K. et al., 1997; Holland P.W.H. et al., 2007).

Entrambi i domini POU_S e POU_H instaurano interazioni specifiche con il DNA attraverso il motivo di legame al DNA *helix-turn-helix* (HTH), tipico degli omeodomini di molti fattori di trascrizione (Brehm A. et al., 1997; Pabo C.O. and Sauer R.T., 1992). Questa struttura è stata il primo motivo proteico, appartenente a fattori di trascrizione, ad essere riconosciuta ed è costituita da due α -eliche connesse da una breve catena di aminoacidi, *giro*, che forma una curvatura e fa sì che le due catene non giacciano sullo stesso piano. Le due eliche sono posizionate ad angolo fisso l'una rispetto all'altra grazie alle interazioni presenti tra di esse. L'elica N-terminale contiene residui basici che legano il DNA, l'elica C-terminale è costituita da aminoacidi idrofobici spazati ad intervalli regolari. L'elica più carbossi-terminale si chiama *elica di riconoscimento*, perché si adatta alla scanalatura principale del DNA. Le catene laterali dei suoi aminoacidi, che sono diverse da proteina a proteina, svolgono un ruolo importante nel riconoscimento della sequenza specifica di DNA a cui si lega la proteina. Il motivo *helix-turn-helix* non è stabile di per sé, ma guadagna stabilità attraverso le interazioni con altri segmenti proteici che possono variare ampiamente nella struttura. Nel dominio POU_S sono due le α -eliche che stabilizzano il motivo HTH ($\alpha 1$ e $\alpha 4$), mentre nel dominio POU_H, come si verifica in altri omeodomini, è una singola elica ($\alpha 1$) che stabilizza il motivo HTH (Herr W. et al., 1995) (Fig. 1.8).

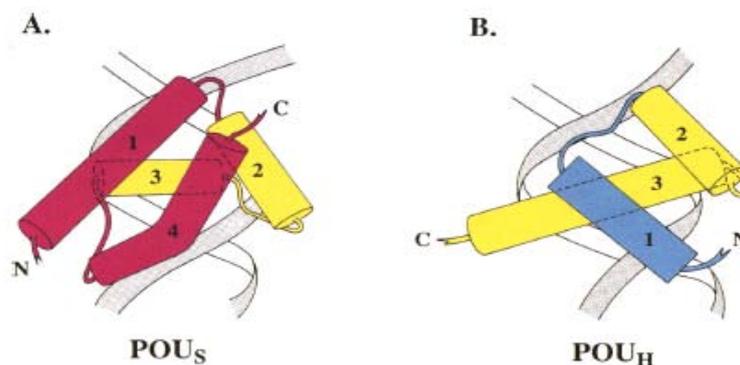


Fig. 1.8- Domini POU_S e POU_H in Oct-1, leganti la sequenza ottamerica dell'istone H2B (Klemm et al., 1994). I motivi HTH sono colorati in giallo. L'immagine A è mostrata con prospettiva dal basso mentre l'immagine B con prospettiva dall'alto. Winship et al., 1995.

Studi biochimici e strutturali hanno dimostrato che i due sottodomini POU possiedono strutture indipendenti (Botfield M.C. et al., 1992) e che instaurano interazioni sequenza-specifiche con il DNA (Aurora R. and Herr W., 1992; Verrijzer C.P. et al., 1992). Queste conclusioni sono state dimostrate tramite l'analisi, con metodi di cristallografia a raggi-X, del dominio POU di OCT-1 che lega naturalmente e con alta affinità la sequenza ottamerica *ATGCAAAT* nel promotore dell'istone H2B.

Entrambi i domini POU_S e POU_H di Oct-1 giacciono in corrispondenza del solco maggiore, ma si trovano posizionati ai lati opposti rispetto alla doppia elica di DNA: il dominio POU_S prende contatto con la sequenza 5'-ATGC tramite legami ad idrogeno instaurati dai residui esposti sulla superficie dell' α -elica 3, mentre il dominio POU_H effettua interazioni, specialmente di tipo idrofobico ed elettrostatico, con il tratto 3'-AAAT della sequenza ottamerica.

Un aspetto importante è rappresentato dal fatto che, nell'interazione "dominio-sequenza ottamerica", i due subdomini POU non effettuano interazioni reciproche ma appaiono tenuti insieme solamente tramite la sequenza spaziatrice (*linker*) (Herr W. et al., 1995). Il braccio amino terminale del dominio POU_H , che costituisce parte della sequenza spaziatrice, si estende lungo il solco minore e, a questo livello, viene instaurato un legame ad idrogeno tra l'Arg5 del dominio POU_H e la base A in posizione 5 della sequenza ottamerica, mentre in corrispondenza della base complementare T, si verifica l'interazione con la Leu55 del dominio POU_S . Questo set di interazioni fa in modo che la posizione 5 della sequenza ottamerica

rappresenti l'unica coppia di basi riconosciuta sia dal dominio POU_H che dal dominio POU_S (Phillips K. et al., 2000) (Fig. 1.9).

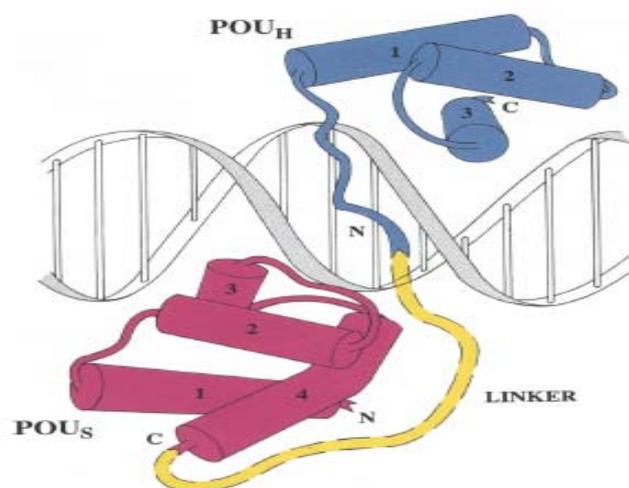


Fig. 1.9- Schema del legame tra il dominio POU di OCT-1 e la sequenza ottamerica dell'istone H2B. Il dominio POU_S è colorato in rosso, il dominio POU_H in blu e il linker in giallo; le linee tratteggiate nella regione spaziatrice indicano che non tutta la sua struttura è nota. Nei due domini POU le eliche sono numerate. (N) amino terminale, (C) carbossi terminale. Klemm J.D. et al., 1994.

Sulla base dei risultati ottenuti da svariati studi, sembra che un potenziale ruolo attribuibile alla sequenza spaziatrice ipervariabile sia quello di mimare l'effetto cooperativo della dimerizzazione (Perutz M.F., 1990); infatti legando insieme i due subdomini, il *linker* aumenta la concentrazione locale di entrambi i domini POU nelle vicinanze di un promotore così che il dominio si comporta come un eterodimero (Herr W. et al., 1995). La lunghezza della regione spaziatrice, che può variare nel range di 15-50 aminoacidi, sembra giocare un ruolo cruciale anche nell'assunzione di un corretto orientamento da parte del dominio POU_S rispetto al dominio POU_H e, conseguentemente, nella determinazione della versatilità tramite la quale i domini POU possono riconoscere diversi frammenti genici (Ryan A.K. et al., 1997).

I fattori di trascrizione della famiglia POU, il cui nome deriva da tre proteine di mammifero Pit1, Oct1 e Oct2, e Unc-86, possono attivare l'espressione dei loro geni *target* mediante il legame ad un motivo di sequenza ottamerica (*octameric sequence motif*). Si tratta

della sequenza-consenso *ATGCAAAT* che originariamente fu identificata in un elemento di controllo del promotore della catena pesante dell'immunoglobulina (Herr W. et al., 1995).

Il *POU domain* è dunque un elemento che accomuna una serie di fattori di trascrizione (*POU class*) che si sono rivelati essere importanti regolatori dell'espressione genica tessuto-specifica in un'estesa trama di processi di sviluppo, inclusi quello linfoide, endocrino e la neurogenesi precoce dei mammiferi (Banerjee-Basu S. et al., 2001).

1.5.3- Identificazione e caratterizzazione di Oct-4

Il fattore di trascrizione Oct-4 gioca un ruolo fondamentale nel mantenimento dello stato di pluripotenza delle cellule staminali embrionali e, con tutta probabilità, potrebbe anche prevenire l'espressione di geni attivati durante il processo di differenziamento (Lee J. et al., 2006). Oct-4 appartiene alla famiglia dei fattori di trascrizione POU e rientra nella classe V delle proteine *POU domain*, per questo è denominato spesso *Pou5f1* (POU class homeobox 1).

Oct-4 fu identificato per la prima volta nel topo, grazie alla scoperta di un'attività di interazione presente negli estratti di cellule staminali embrionali indifferenziate e di cellule di carcinoma embrionale (Schöler H.R. et al., 1989; Okamoto K. et al., 1990).

Nel topo il gene Oct-4 codifica per una proteina di 352 amminoacidi (Okamoto K. et al., 1990; Rosner M.H. et al., 1990) costituita da una regione conservata di circa 150 amminoacidi, nota come dominio POU. All'interno di questa regione possono essere distinti un dominio POU-specifico e un omeodominio POU che mediano il legame sequenza-specifico al DNA (Herr W. and Cleary M.A., 1995). La proteina Oct-4 attiva la trascrizione attraverso un motivo ottamerico, di 8 bp, localizzato prossimalmente o distalmente rispetto al sito d'inizio della trascrizione (Lamb K.A. et al., 1998) e ritrovato a livello di promotori ed *enhancers* di molti geni ubiquitari e tessuto-specifici (Ruvkin G. and Finney M., 1991). Inoltre si è visto che questo fattore di trascrizione viene espresso durante lo sviluppo embrionale precoce nelle cellule embrionali totipotenti.

Nel topo, il gene Oct-4 è localizzato nel cromosoma 17 e mappa all'interno o nelle immediate vicinanze del complesso maggiore di istocompatibilità (MHC), (Schöler H.R. et

al., 1990b). L'RNA messaggero è costituito da 5 esoni (Okazawa H. et al., 1991; Yeom Y.I. et al., 1991) che codificano per un trascritto di circa 1.5 kb (Okamoto K. et al., 1990; Rosner M.H. et al., 1990).

La regione del promotore non presenta la TATA box e la trascrizione viene avviata in siti multipli (Okazawa H. et al., 1991).

I geni ortologi a Oct4 mostrano un grado elevato di conservazione nell'organizzazione strutturale del genoma e nella sequenza; questi sono stati identificati in altre specie di mammifero, come uomo, bovino e ratto (Pesce M. and Schöler H.R., 2000). Confrontando le sequenze di promotori ed *enhancer* dei geni Oct4 di mammifero con quelle dei geni murini ortologi, è infatti emersa un'organizzazione comune degli elementi *cis*-regolatori. Questo ha fatto supporre una modalità di funzionamento simile durante il processo di sviluppo embrionale (Nordhoff V. et al., 2001).

Takeda J. et al., in un lavoro pubblicato nel 1992, sono riusciti ad identificare per la prima volta Oct-4 nell'uomo, mediante l'isolamento del cDNA con la tecnica 5'-RACE e la sua successiva amplificazione con la metodica della PCR (*Polymerase Chain Reaction*).

Questo studio ha inoltre permesso di determinare la localizzazione cromosomica del gene che codifica per Oct-4: il gene *OTF3* è stato mappato all'interno del cromosoma 6, in stretta congiunzione con i *loci* HLA, e consta di 6 esoni per una lunghezza di 7 kb (Takeda J. et al., 1992).

Ricorrendo alla tecnica dell'ibridazione *in situ*, Guillaudeux e colleghi (1993) localizzarono il gene *OTF3* in posizione 6p22-p21.3, confermando quindi la sua vicinanza alla regione in cui è situato il complesso maggiore di istocompatibilità di classe I (MHC class I), come era stato sostenuto nel precedente studio di Takeda et al..

L'anno successivo, Crouau-Roy e i suoi Colleghi (1994) collocarono il gene *OTF3* all'interno della regione MHC di classe I, approssimativamente a 100 kb di distanza dal gene HLA-C.

Il gene del fattore di trascrizione ottamerico, *POU5F1* è quindi stato localizzato in una regione interposta tra i geni HLA-C e S, circa 30 kb dal centro del gene S (Krishnan B.R. et al., 1995) (Fig. 1.10).

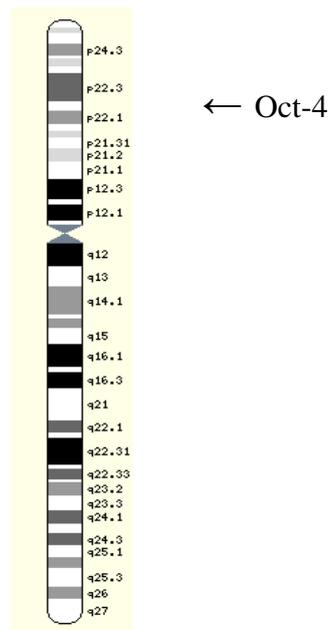


Fig. 1.10- Illustrazione schematica del cromosoma 6 e della localizzazione del gene che codifica per la proteina Oct-4.

1.5.4- Oct-4 e le sue isoforme

Gli studi effettuati da Takeda e i suoi collaboratori nel 1992 hanno portato all'amplificazione di sequenze codificanti le proteine che legano l'ottamero, Oct-1 e Oct-4 (Oct3): le sequenze dei due prodotti di PCR erano identiche alla proteina umana Oct-1 e altre due erano omologhe alla proteina murina Oct-4. L'isolamento dei cDNA codificanti Oct-4 ha rivelato l'espressione di due isoforme di questo fattore di trascrizione, denominate Oct-4A e Oct-4B, che rappresentano due differenti varianti di *splicing* (Fig. 1.11).

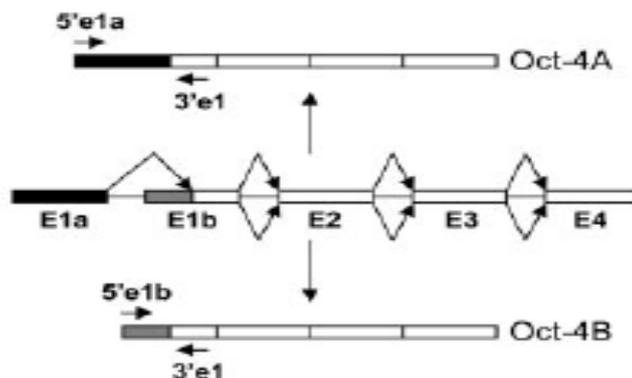


Fig. 1.11- Rappresentazione schematica dei trascritti del fattore di trascrizione Oct-4 umano: Oct-4A e Oct-4B originano dallo stesso gene mediante splicing alternativo. L'estremità amino-terminale nell'isoforma Oct-4A è indicata in nero mentre l'estremità amino-terminale dell'isoforma Oct-4B è mostrata in grigio. L' mRNA di Oct-4A contiene l'esone 1a (E1a), la metà 3' dell'esone 1b (E1b) e gli esoni 2, 3, 4 (E2, E3, E4). L'mRNA dell'isoforma Oct4/B comprende tutti gli esoni dall'1b all'4 (E1b, E2, E3, E4) e all'estremità 5' presenta un introne. Lee et al., 2006.

A livello nucleotidico, entrambe le sequenze delle isoforme Oct4/A e Oct4/B possiedono gli esoni 2, 3, 4 e 5; l'isoforma Oct-4B non contiene l'esone 1 ma presenta invece 202 bp relative ad una regione intronica (Liedtke S. et al., 2008).

Oct-4A e Oct-4B sono costituiti da 360 e 265 amminoacidi rispettivamente. Entrambe le isoforme possono essere suddivise sia strutturalmente sia funzionalmente in tre domini:

- Il dominio N-terminale (*NTD*), che comprende la regione di attivazione trascrizionale, consta di 133 amminoacidi nell'isoforma A e di 40 amminoacidi nell'isoforma B;
- Il dominio POU centrale, che consente di legare il DNA in modo sequenza-specifico, è costituito da 156 amminoacidi in Oct-4A e da 154 amminoacidi in Oct-4B;
- il dominio C-terminale (*CTD*), che controlla la funzione transattivante, comprende 71 amminoacidi nell'isoforma A e 71 nell'isoforma B (Fig. 1.12)

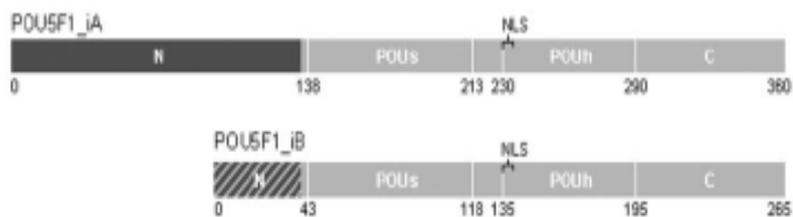


Fig. 1.12- Rappresentazione schematica di Oct4/A e Oct4/B: sono distinguibili i 3 domini NTD, Pou domain, CTD e la loro rispettiva lunghezza (in amminoacidi). Entrambe le isoforme contengono sequenze di localizzazione nucleare (NLS) all'inizio dell'omeodominio POUh. Cauffman et al., 2006.

Mentre il dominio carbossi-terminale risulta essere identico nelle due isoforme, il dominio POU dell'isoforma B contiene due amminoacidi in meno, in corrispondenza della sequenza amino-terminale, rispetto al dominio POU dell'isoforma Oct-4A. Infatti il dominio POU di Oct4/B (POU^B) consta di 154 amminoacidi mentre il dominio POU di Oct4/A (POU^A) contiene un totale di 156 amminoacidi (Cauffman G. et al., 2006) (Fig. 1.12). Ad eccezione di queste regioni, le due isoforme non presentano molta similarità: ad esempio, Oct4/B non possiede un dominio di transattivazione riconoscibile, come si riscontra invece all'amino-terminale dell'isoforma A, e per tanto non sembra possedere la capacità di attivatore trascrizionale (Lee J. et al, 2006). E' stato inoltre dimostrato che il dominio N-terminale di Oct-4B inibisce il legame sequenza-specifico al DNA: infatti sono presenti due regioni aminoacidiche, dall'amminoacido 1 al 20 e dall'amminoacido 21 al 40, che impediscono l'interazione specifica con il DNA e sembrano quindi mantenere Oct-4B in uno stato latente caratterizzato da scarsa affinità per le sequenze bersaglio (Lee J. et al., 2006).

Dall'analisi della sequenza nucleotidica è emerso che entrambe le isoforme del fattore di trascrizione Oct4 possiedono sequenze di localizzazione nucleare (Fig. 1.12), ma ciò nonostante è stato dimostrato, fino ad oggi, che solamente Oct-4A viene traslocata nel nucleo mentre l'isoforma B rimane prevalentemente localizzata a livello citoplasmatico. Lee e i suoi collaboratori hanno eseguito diversi esperimenti per tentare di chiarire il meccanismo di mancata traslocazione nucleare dell'isoforma B: in uno di questi esperimenti, è stata creata una proteina di fusione contenente l'estremità N-terminale di Oct-4B e il relativo POU domain e, tramite analisi di immunofluorescenza, si è potuto osservare che questa forma “tronca” di

Oct4/B continuava a rimanere localizzata a livello di citoplasma cellulare. Per spiegare i dati ottenuti, Lee et al. hanno ipotizzato due modelli principali che potrebbero essere responsabili della mancata traslocazione nucleare: il primo modello sostiene che il dominio N-terminale di Oct-4B potrebbe contenere un segnale di esportazione nucleare, riconosciuto da alcuni recettori di trasporto, che causerebbe la traslocazione di Oct-4B dal nucleo al citoplasma; nel secondo modello invece, la sequenza di localizzazione nucleare (*NLS*) interna al dominio POU potrebbe venire bloccata oppure nascosta dall'estremità N-terminale di Oct4/B, impedendo così la traslocazione della proteina a livello nucleare (Lee et al., 2006).

Analisi di sequenza della proteina Oct-4B hanno rilevato la presenza di potenziali siti di fosforilazione, a livello di residui di serina, treonina e tirosina; è plausibile dunque pensare che l'esclusione nucleare possa essere anche indotta dalla fosforilazione dell'isoforma B ad opera di alcune molecole segnalatorie (Lee et al., 2006).

Lee et al., mediante l'utilizzo di saggi EMSA, hanno inoltre dimostrato che l'isoforma B di Oct-4 sembrerebbe non essere in grado di legare sequenze target di DNA. Questi risultati potrebbero indicare o che il dominio N-terminale di Oct-4A rappresenta un requisito indispensabile per un efficiente legame al DNA, oppure che Oct-4B presenta, a livello dell'estremità amino-terminale, sequenze aminoacidiche che inibiscono l'interazione con sequenze specifiche di DNA.

Per chiarire quali delle due cause sopracitate potesse essere quella responsabile della mancata interazione di Oct-4B con il DNA, Lee e i suoi collaboratori hanno creato una serie di mutanti caratterizzati da delezioni a livello dell'estremità N-terminale o di quella C-terminale. L'interazione dei mutanti appositamente creati con sequenze *consensus*, è stata poi testata mediante saggio EMSA. I saggi eseguiti con mutanti che presentavano solo i domini POU rispettivamente di Oct-4A e di Oct-4B, si sono dimostrati interagenti con la sonda di DNA; i mutanti che contenevano, oltre al dominio POU, anche l'estremità C-terminale risultavano anch'essi capaci di formare complessi con il DNA sonda, mentre i mutanti costituiti dal dominio POU e dall'estremità N-terminale di Oct-4B non mostravano capacità di interazione con la sonda specifica di DNA. I risultati ottenuti suggerivano quindi che fosse proprio il dominio amino-terminale di Oct-4B ad inibire l'interazione sequenza-specifica con il DNA.

Per approfondire ulteriormente i risultati ottenuti dai saggi EMSA eseguiti sui vari mutati, Lee et al. hanno creato una serie di proteine chimeriche ottenute scambiando i domini

tra le due isoforme della proteina Oct-4. Mutanti che presentavano il dominio POU e l'estremità C-terminale dell'isoforma A e l'estremità N-terminale dell'isoforma B, non erano capaci di interagire con la sonda di DNA, mentre mutanti costituiti dal dominio POU e dall'estremità carbossi-terminale di Oct-4B e dall'estremità amino-terminale di Oct-4A si dimostravano interagenti con la sonda di DNA specifica. Questi risultati sembrano quindi confermare l'ipotesi dell'esistenza di una o più sequenze a livello dell'estremità amino-terminale dell'isoforma B che sono in grado di inibire il legame al DNA e che agiscono sia sul dominio POU dell'isoforma A sia sul dominio POU dell'isoforma B.

Ulteriori esperimenti eseguiti da Lee ed i suoi collaboratori su mutanti che presentavano delezioni a livello dell'estremità N-terminale, hanno permesso di identificare principalmente due regioni, la sequenza che comprende gli aminoacidi 1-20 e la sequenza che comprende gli aminoacidi 21-40, come probabili responsabili dell'inibizione del legame al DNA di Oct-4B (Lee j. et al., 2006).

Il dominio N-terminale di Oct4/A, a differenza del dominio amino-terminale dell'isoforma B, è ricco in prolina e residui acidi, mentre il dominio C-terminale è ricco in prolina, serina e treonina (Schöler H.R. et al., 1991a).

In Oct-4A, il dominio N-terminale è stato accettato come effettore della transattivazione (Brehm A. et al., 1997), anche se non è escluso un coinvolgimento, in tale processo, del dominio C-terminale (Vigano M.A. et al., 1996). E' stato dimostrato, inoltre, che l'attività del dominio C-terminale di Oct4/A è specifica per il tipo cellulare ed è regolata attraverso la fosforilazione, cosa che non si verifica per il dominio N-terminale (Brehm A. et al., 1997; Brehm A. et al., 1999). L'attività tipo cellulare-specifica viene riscontrata solo quando il dominio C è legato al dominio POU di Oct4; ciò è in accordo con l'osservazione che il dominio POU è funzionale per l'interazione con fattori secondari (Schöler H.R., 1991a; Brehm A. et al., 1997). E' fondamentale sottolineare che la regolazione da parte del *POU domain* non è emersa dallo studio di altri domini di transattivazione, ma appare specifica per il dominio C di Oct4/A (Brehm A. et al., 1997; Ovitt C.E. et al., 1998).

Il fatto che il dominio N e il dominio C differiscano in termini di attività e specificità per il tipo cellulare, può spiegare la diversità funzionale del fattore Oct4/A. Il dominio C-terminale, infatti, è in grado di attivare determinati geni bersaglio che non rispondono al dominio N-terminale durante lo sviluppo (Brehm A. et al., 1997).

Le differenze a livello strutturale evidenziate nelle due isoforme del fattore di trascrizione Oct-4, sono ritenute responsabili delle differenti funzionalità e della diversa localizzazione che, le due proteine, hanno mostrato sia *in vivo* sia *in vitro*: infatti, mentre Oct-4A risulta localizzata a livello nucleare ed è sicuramente coinvolta nella trascrizione di diversi geni che sostengono il *self-renewal* e la pluripotenza delle cellule staminali, Oct4/B è situata principalmente a livello citoplasmatico e, anche se il suo ruolo non risulta ancora chiaro, sembrerebbe non essere coinvolta nel mantenimento della pluripotenza e del *self renewal* delle cellule staminali (Atlasi Y. et al., 2008).

Recentemente, in uno studio pubblicato da Atlasi et al. nel 2008, è stata riscontrata la probabile esistenza di una terza isoforma della proteina Oct-4, denominata Oct-4B1.

La potenziale nuova isoforma di Oct-4 è stata identificata in saggi di RT-PCR allestiti per ottenere il cDNA dell'isoforma Oct-4B, nei quali è stato utilizzato un *primer forward*, localizzato nell'esone 1b, e un *primer reverse*, situato all'interno dell'esone 3. Infatti, analizzando i prodotti di PCR ottenuti tramite elettroforesi su gel di agarosio, Atlasi e i suoi colleghi hanno notato la presenza di un banda di altezza anomala (~500 bp), rispetto alle bande normalmente ottenute dall'amplificazione dell'isoforma Oct-4B (~267 bp). Potendo escludere la possibilità che si trattasse di un prodotto aspecifico, visto che i campioni di RNA erano stati trattati con DNase prima dell'utilizzo, Atlasi e i suoi collaboratori hanno ipotizzato potesse trattarsi di una nuova isoforma del fattore di trascrizione Oct-4 (Atlasi Y. et al., 2008).

Questa potenziale nuova isoforma consta di 492 bp e contiene tutti e 5 gli esoni presenti in Oct-4B con l'aggiunta di un nuovo esone, l'esone 2B, situato tra l'esone 2 e l'esone 3. Dall'analisi della sequenza nucleotidica, Atlasi e i suoi colleghi hanno osservato che, in seguito all'inserzione dell'esone 2B, la sequenza di Oct4/B1 ha acquisito un nuovo codone di stop *TGA in frame*, interno all'esone 2B, che crea un *open reading frame (ORF)* di 348 nucleotidi; questo potrebbe portare alla formazione di un potenziale corto peptide della lunghezza di 115 amminoacidi (Fig. 1.13) (Atlasi Y. et al., 2008), ipotesi che però attualmente non è stata ancora confermata.

Allineando la sequenza aminoacidica di Oct-4B con la sequenza di Oct-4B1, si è visto che la regione che si estende dall'amminoacido 1 all'amminoacido 80 è identica nelle due isoforme; si è riscontrata, infatti, similarità tra i domini N-terminale delle due isoforme e anche, in parte, tra i due domini POU₅: infatti i primi 37 amminoacidi del dominio POU₅

sono identici nelle due isoforme ma in Oct-4B1 manca la restante parte del dominio, così come il dominio POU_H e il dominio C-terminale transattivatore (Atlasi Y. et al., 2008).

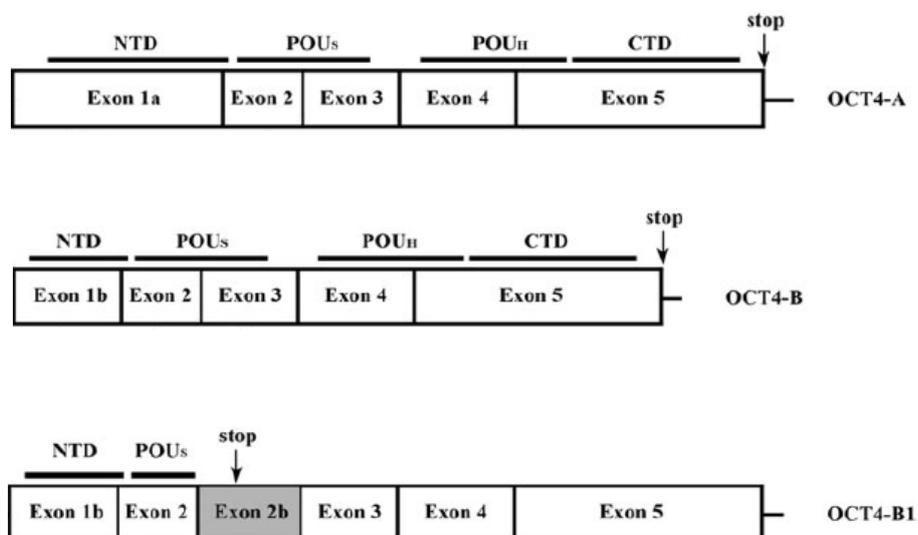


Fig. 1.13- Rappresentazione schematica dell'organizzazione in esoni e dei domini proteici delle tre isoforme di Oct-4.

Anche l'isoforma B1 sembrerebbe essere coinvolta nel mantenimento di uno stato cellulare indifferenziato sia in cellule staminali embrionali sia in cellule staminali cancerose. Infatti, tramite saggi di Q-PCR e RT-PCR, sono stati evidenziati, proprio in questi tipi cellulari, alti livelli di espressione di Oct-4B1 mentre, in molti tipi di cellule staminali adulte, quali le cellule staminali mesenchimali o le cellule tumorali immortalizzate derivanti da cancro della cervice uterina (*HeLa*), Oct-4B1 mostra bassissimi livelli di espressione. Inoltre si è notato che l'espressione di Oct-4B1 inizia ad essere down-regolata durante il corso della differenziazione cellulare. Questi risultati sembrerebbero attribuire alla nuova potenziale isoforma di Oct-4, un ruolo da marcatore specifico di cellule staminali embrionali e cancerose (Atlasi Y. et al., 2008).

1.6- Espressione di Oct-4

1.6.1- Espressione di Oct-4A

Il ruolo di Oct-4A nel mantenimento della pluripotenzialità delle cellule staminali, nel corso degli anni, è stato indubbiamente accertato: molti studi su diverse linee cellulari staminali hanno dimostrato che, in seguito a differenziazione, i livelli di espressione di Oct-4A diminuiscono sensibilmente (Reubinoff B.E. et al., 2000) e la sua ridotta espressione, ottenuta tramite il metodo dell' *RNA interference (RNAi)*, induce differenziazione cellulare (Hay D.C. Et al., 2004).

Data la localizzazione nucleare di Oct-4A, sono state assegnate a questa isoforma le proprietà di mantenimento della staminalità e del *self-renewal* cellulare; tuttavia è stato dimostrato che la sola espressione di Oct-4A non è sufficiente a garantire il mantenimento della staminalità, ma è necessaria un'attività cooperativa con altri fattori di trascrizione, quali Sox2 e Nanog (Cauffman G. et al., 2006). Questi tre fattori di trascrizione insieme promuovono la pluripotenzialità e il *self-renewal* delle cellule staminali embrionali grazie alla regolazione positiva dei loro geni e dei geni codificanti componenti essenziali nelle *pathways* segnalatorie e grazie anche alla regolazione negativa di geni importanti nei processi di sviluppo (Boyer L.A. et al., 2005; Kuroda T. et al., 2005).

Lo schema di espressione della forma umana di Oct-4A può essere paragonato a quello che si riscontra per la forma murina della proteina, suggerendo quindi un funzione simile nel prevenire il differenziamento delle cellule staminali embrionali nell'uomo.

Durante lo sviluppo embrionale Oct-4A viene inizialmente espresso solo a livello di blastomeri. Successivamente, la sua espressione si focalizza nelle cellule della *inner cell mass* mentre viene downregolato nel trofoectoderma (TE) e nell'endoderma primitivo. Nello sviluppo embrionale più tardivo, l'espressione di Oct-4A è mantenuta unicamente nelle cellule germinali primordiali (PGCs), i precursori dei gameti (Pesce M. et al., 1998a; Pesce M. et al., 2001). Studi su cellule staminali murine, in cui il pattern di espressione della proteina Oct-4A è paragonabile a quello umano (Hansis C. et al., 2000), hanno dimostrato che cellule mancanti di Oct-4A durante lo sviluppo embrionale differenziano in linee somatiche, mentre cellule che mantengono l'espressione della proteina per tutta la durata dello sviluppo embrionale, non

solo conservano la propria pluripotenza e mantengono la capacità di svilupparsi in cellule germinali, ma possono anche convertirsi in cellule staminali cancerose (Pesce M. et al., 2001).

Quando cellule staminali embrionali vengono indotte al differenziamento, l'espressione dell'mRNA di Oct-4A è downregolato (Leonardo M.J. Et al., 1989; Schöler H.R. Et al., 1989). L'espressione del fattore di trascrizione è quindi associata alle cellule indifferenziate sia nello stadio del preimpianto dell'embrione che in coltura.

Singolare risulta la correlazione tra l'espressione di Oct-4A e il potenziale differenziativo: tutte le cellule staminali totipotenti e pluripotenti dell'embrione esprimono Oct-4A.

Recentemente, si è potuto notare che embrioni omozigoti per la delezione di Oct-4A (Oct-4^{-/-}), ottenuti tramite silenziamento, muoiono al momento dell'impianto a causa della mancata formazione dell'*ICM*: studi di tali embrioni *in vitro* hanno infatti dimostrato che la mancanza di Oct-4A cambia il fato delle cellule destinate a costituire la *ICM*, che differenziano in cellule del trofoblasto. E' interessante notare che la soppressione della funzione di Oct-4A non influenza il mantenimento del fenotipo pluripotente dei blastomeri prima della formazione della blastocisti; questo suggerisce che la transattivazione mediata da Oct-4, negli embrioni precoci di topo, inizia ed essere essenziale per il mantenimento dello stato di pluripotenza quando dall' *ICM* si origina la prima linea cellulare somatica, il trofoectoderma (Pesce M. et al., 2001). Perciò, la decisione di mantenere espresso Oct-4 nell'*inner mass* e di downregolarlo nelle cellule più esterne, rappresenta uno degli eventi cruciali che permette il corretto sviluppo dell'embrione preimpianto (Pesce M. et al., 1999).

Hansis e i suoi collaboratori, hanno esaminato l'espressione di Oct-4A nella blastocisti umana: l'*ICM* e il trofoectoderma di 17 blastocisti umane sono stati isolati e i livelli di mRNA di Oct-4A sono stati esaminati tramite RT-PCR (Hansis C. et al., 2000; Hansis C. et al., 2001). I risultati ottenuti dimostrano che il livello medio di Oct-4A misurato a livello di *ICM* risultava trenta volte più elevato rispetto ai livelli registrati nelle cellule differenziate del trofoectoderma (Hansis C. et al., 2000; Hansis C. et al., 2001). Tali risultati suggeriscono che il pattern di espressione di Oct-4 rilevato nelle cellule di topo risulta molto simile a quello evidenziato nelle cellule umane.

1.6.2- Regolazione dell'espressione di Oct-4A

Dato il ruolo critico nel mantenimento della pluripotenzialità, l'attività di Oct-4 deve essere strettamente regolata per garantire la continuità della linea germinale e un'adeguata differenziazione dei vari tessuti ed organi.

L'espressione di Oct-4A è regolata a livello trascrizionale da elementi che agiscono in *cis*, situati a monte del gene Oct-4, e dalla metilazione della struttura cromatinica (Ben-Shushan E. et al., 1993).

A partire dall'analisi dell'espressione del gene reporter LacZ, posto sotto il controllo di un frammento di 18 kb del *locus* genomico di Oct-4, Yeom Y.I. Et al. (1996) hanno identificato due elementi denominati *proximal enhancer* (PE) e *distal enhancer* (DE), responsabili della regolazione dell'espressione cellulo-specifica di Oct-4 (Yeom Y.I. et al., 1996), (Fig. 1.13)

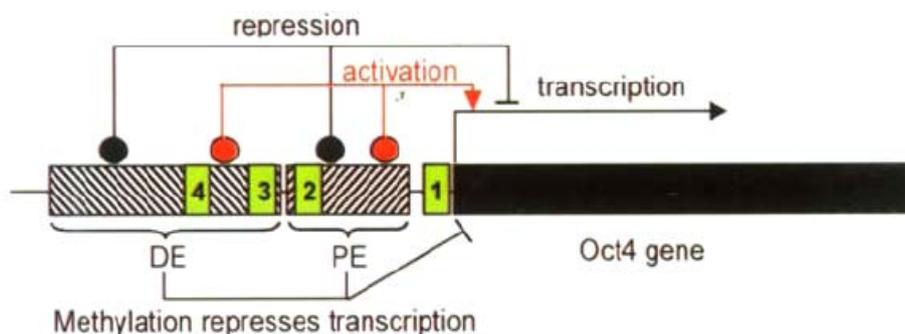


Fig. 1.13- Gli elementi a monte del gene Oct4 (DE e PE) ne regolano l'espressione. Ci sono 4 regioni altamente conservate corrispondenti agli elementi promotore/enhancer di Oct4 nell'uomo, nel bovino e nel topo(1, 2, 3, 4). Ogni enhancer contiene molti siti potenziali di legame per fattori di trascrizione che possono sia attivare (rosso) che inibire l'espressione di Oct4. La metilazione a livello di queste regioni inibisce l'espressione di Oct4 nelle cellule differenziate.

Utilizzando la tecnica del *footprinting in vivo*, Yeom Y.I. E i suoi colleghi hanno definito precisamente i siti di legame per i fattori di trascrizione all'interno di questi due *enhancer*; sono stati identificati due siti: uno chiamato 1A, localizzato lungo la sequenza di PE, e un altro, chiamato 2A, situato lungo la sequenza DE. Entrambi questi siti mostrano sequenze identiche, omologhe al GC box e sono cruciali per l'attività di PE e DE,

rispettivamente (Yeom Y.I. et al., 1996).

Inoltre Nordhoff V et al. (2001), eseguendo un'analisi comparata tra le sequenze dei promotori a monte di Oct4 nei bovini, nell'uomo e nei topi, hanno rilevato quattro regioni di omologia conservate (da *CR1* a *CR4*) tra queste specie (conservazione pari al 66-94%) (Nordhoff V. et al., 2001). Essi hanno scoperto che il sito 1A di PE è situato approssimativamente tra *CR2* e *CR3* (Nordhoff V. et al., 2001). Unitamente a ciò è stato identificato un gran numero di motivi *CCC(A/T)CCC*, i quali esibiscono vari livelli di omologia all'interno delle regioni a monte del gene (Nordhoff V. et al., 2001). Tali sequenze potrebbero essere essenziali per l'espressione di Oct4/A.

Hummelke e Cooney riportano che il fattore nucleare delle cellule germinali (*GCNF*), un recettore nucleare orfano, è in grado di reprimere l'attività del gene Oct4, legandosi all'interno della sequenza PE (Hummelke G.C. et al., 2001; Fuhrmann G. et al., 2001). L'espressione di *GCNF* risulta infatti inversamente correlata all'espressione di Oct4 nelle cellule embrionali in differenziamento (Hummelke G.C. et al., 2001; Fuhrmann G. et al., 2001). Negli embrioni murini *knock-out* per *GCNF*, l'espressione di Oct4 dopo la gastrulazione non viene più limitata alla linea germinale (Fuhrmann G. et al., 2001), suggerendo che il gene *GCNF* è responsabile del silenziamento del gene Oct4 durante la differenziazione delle cellule staminali.

In aggiunta a tali elementi in *cis*, sono stati riscontrati meccanismi addizionali che regolano l'attività di Oct4/A. Un'ipotesi consiste nel fatto che il livello basale di Oct4 nelle cellule totipotenti possa essere visto come una conseguenza della presenza di cromatina attiva, piuttosto che una condizione dovuta all'azione dei fattori attivanti la trascrizione. Ben-Shushan E. et al. (1993) hanno provato che la mancanza dell'attività di Oct4 nelle cellule ibride staminali-fibroblasti è accompagnata da una rapida metilazione delle sequenze regolatrici come PE e DE nelle regioni promotrici (promotore/*enhancer*) di Oct4 (Ben-Shushan E. et al., 1993). Jaenisch R. ha proposto il verificarsi di una fase di metilazione che si svolge nelle cellule somatiche dell'embrione (Jaenisch R., 1993). Questi studi concordano sull'ipotesi che la metilazione delle sequenze regolatrici di Oct4 (come PE e DE) inibisca l'espressione di Oct4. Le cellule germinali pluripotenti (PGC), che originano dal mesoderma extraembrionario, mantengono un livello basale dell'espressione di Oct4, suggerendo la presenza di un meccanismo in grado di prevenire la metilazione del loro materiale genetico, almeno per quanto concerne la sequenza regolatrice di Oct4 (Goto T. et al., 1999). D'altra

parte il mantenimento dell'espressione di Oct4 nelle PGC e negli oociti potrebbe essere dovuto ad una mancata riprogrammazione della cromatina tramite metilazione, che avviene solitamente nelle cellule epiblastiche al momento della gastrulazione (Pesce M. et al., 1998b).

1.6.3- Funzioni di Oct-4A

Poiché l'attività tipo cellulare-specifica dei fattori regolatori assicura l'espressione di geni *target* seguendo un preciso ordine durante il processo di sviluppo, Oct4 e i suoi partners funzionali devono essere regolati in modo specifico nel corso dell'embriogenesi dei mammiferi. Infatti, Oct4 viene espresso nelle cellule germinali a partire dallo stadio di zigote totipotente fino a quello di oocita altamente specializzato (Palmieri S. et al., 1994; Pesce M. et al., 1998a; Pesce M. et al., 2001). Sembra che Oct4 debba funzionare in sinergia con altri regolatori al fine di attivare specifici geni target in particolari tipi cellulari e in punti precisi del processo di sviluppo.

Con lo scopo di definire la relazione che intercorre tra l'espressione di Oct4 e la pluripotenzialità delle cellule staminali, Niwa et al. misurarono il grado di espressione di Oct4 in cellule staminali embrionali a varie condizioni (Niwa H. et al., 2000; Niwa H., 2001). I risultati ottenuti indicano che il controllo esercitato da Oct4 sulla pluripotenzialità dipende dal livello stesso di Oct4, considerato in termini quantitativi. Più precisamente, essi hanno determinato che un alto livello di espressione di Oct4 indirizza le cellule staminali a differenziarsi in cellule dell'endoderma e del mesoderma, mentre cellule con un basso livello di Oct4/A differenziano in cellule del trofoectoderma (Niwa H. et al., 2000; Niwa H., 2001). Solamente un "livello normale" di Oct4 può mantenere le cellule staminali in uno stato pluripotente (Niwa H. et al., 2000; Niwa H., 2001). Tali osservazioni mettono in luce la diversità di Oct4 rispetto a molti fattori di trascrizione noti, che invece svolgono le loro funzioni secondo una modalità "on-off".

In alcuni casi, Oct4 può agire come un repressore di geni target, mentre in altri casi assume il ruolo di attivatore. Per esempio, le sequenze consenso per il legame da parte di Oct4 sono state riportate all'interno dei promotori prossimali che regolano la trascrizione dei geni

per la subunità α e la subunità β della gonadotropina corionica umana (hCG) (Liu L. et al., 1996; Liu L. et al., 1997). Nelle cellule staminali Oct4 funge da repressore per entrambi questi geni attraverso il legame al motivo ottamerico (Liu L. et al., 1996; Liu L. et al., 1997) (Fig.1.14B). Diversamente, nelle cellule differenziate del trofoectoderma, Oct4 è *downregolato* perciò non è in grado di trans-reprimere l'espressione di hCG; questo è uno degli eventi iniziali che stabiliscono un nuovo *pattern* di espressione genica volto appunto ad indirizzare il differenziamento verso la linea cellulare del trofoectoderma.

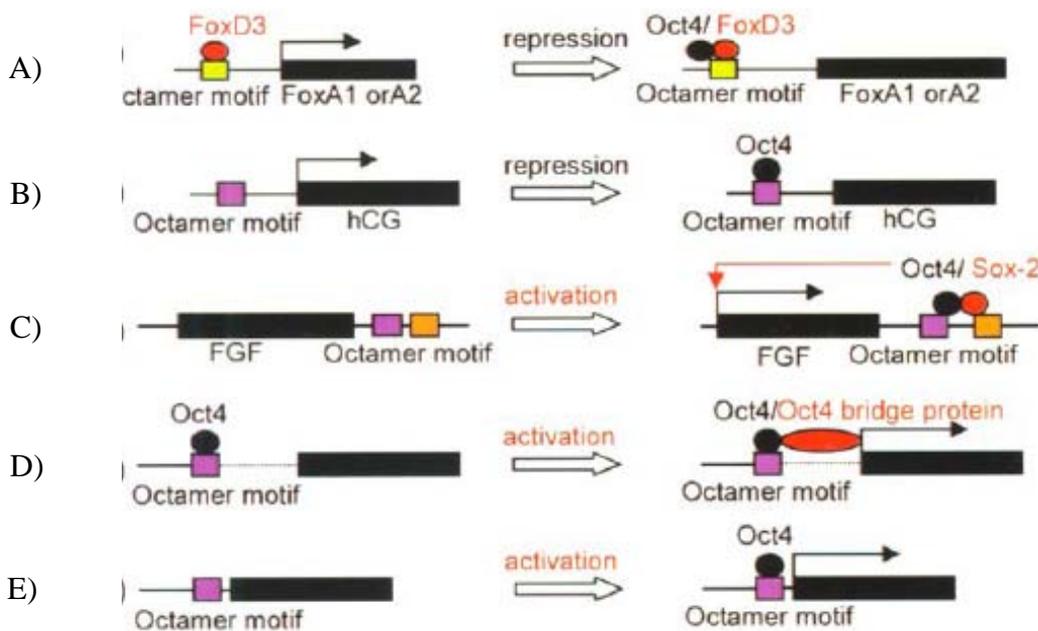


Fig. 1.14- Modalità d'azione di Oct4 su differenti geni target. Oct4 reprime l'espressione genica sia indirettamente neutralizzando attivatori come FOXD3 (esempio A), sia direttamente tramite il legame dei promotori (esempio B). Oct4 agisce anche come attivatore della trascrizione genica legando le sequenze ottameriche situate a monte (esempio D ed E) o a valle (esempio C) dei geni bersaglio. In sintesi, Oct4 si lega ai motivi ottamerici posti immediatamente a monte del promotore per attivare direttamente l'espressione del gene (esempio E). Alternativamente, Oct4 può entrare in sinergia con altri fattori come Sox2 attivando la trascrizione genica (esempio C). Quando localizzato a distanza considerevole, come nell'esempio D, proteine adattatrici vengono reclutate per collegare Oct4 al complesso basale della trascrizione promuovendo l'attivazione trascrizionale.

Il ruolo di attivatore o repressore di geni bersaglio svolto da Oct4 può essere mediato dall'azione di altri fattori di trascrizione. I membri della famiglia *Forkhead Box* (Fox)

possiedono una struttura di legame per il DNA e sono fortemente implicati nelle prime fasi del commissionamento delle linee cellulari embrionali, specialmente nello sviluppo dell'endoderma e conseguentemente nell'organogenesi di derivazione endodermica (Hromas R. et al., 1995). FoxD3, un membro di tale famiglia, può legarsi e attivare i promotori degli altri componenti della famiglia, come ad esempio FoxA1 e FoxA2, coinvolti nello sviluppo embrionale degli organi della parte anteriore dell'intestino di derivazione endodermica (Guo Y. et al., 2002). Guo et al. affermano che Oct4 è in grado di reprimere l'espressione di FoxA1 e FoxA2 mediante l'interazione con il dominio di legame al DNA del loro attivatore FoxD3 (Guo Y. et al., 2002) (Fig.1.14A).

Oct4 può esercitare la sua funzione di regolatore attraverso un meccanismo diretto che comprende la transattivazione dei bersagli genici nelle cellule staminali. Oct4, infatti, nelle cellule staminali può transattivare direttamente i geni *target* sia che esso si trovi in prossimità o a una certa distanza, a seconda della localizzazione dei suoi siti di legame sui promotori dei *target*.

Quando agisce a lunga distanza, Oct4 può reclutare in suo aiuto una serie di coattivatori specifici per le cellule staminali capaci di collegare il fattore al complesso basale della trascrizione (Fig.1.14D). In modo relativamente inaspettato, si è trovato che l'oncoproteina adenovirale E1A è in grado di mimare questa funzione dei coattivatori cellula staminale-specifici (Schöler H.R. et al., 1991b; Butteroni C. et al., 2000). È stato inoltre riportato che un'altra oncoproteina, HPV-E7, svolge un ruolo simile nel processo di attivazione genica mediata da Oct4 (Bhrem A et al., 1999). Entrambe queste proteine funzionano come fattori di connessione tra Oct4 e il macchinario trascrizionale.

Da numerosi studi, è emerso che Oct4 agirebbe in sinergia con altri fattori trascrizionali che si legano nelle vicinanze degli elementi *cis*-regolatori dei promotori *target*. Un esempio riguarda la regolazione dell'espressione di FGF4. FGF4 è un fattore di crescita specifico per le cellule staminali ed è caratterizzato dalla presenza di un *enhancer*, posizionato all'interno della regione 3'-non tradotta (UTR) del gene, responsabile dell'espressione specifica nelle cellule staminali (Curatola A.M. et al., 1990). L'*enhancer* contiene una sequenza ottamerica adiacente al sito in cui Oct4 e il fattore di trascrizione Sox2 si legano sinergicamente e cooperativamente per attivare la trascrizione (Yuan H. et al., 1995) (Fig.1.14C). Tale sinergismo è mediato da interazioni proteina-proteina (Ambrosetti D.C. et al., 1997). In assenza di Sox2, Oct4 non è sufficiente per stimolare l'*enhancer* di FGF4, anche se è presente

il fattore di collegamento E1A (Ambrosetti D.C. et al., 1997). Oltre a ciò, la formazione del complesso Oct4/Sox2 sembrerebbe essere un evento reciproco: l'organizzazione del complesso può smascherare domini di attivazione su entrambe le proteine, fino a quel momento rimasti nascosti, e promuovere così l'attivazione della trascrizione (Ambrosetti D.C. et al., 2000). Come l'*enhancer* di FGF4, anche l'*enhancer* di Sox2 possiede un motivo ottamerico che può essere regolato dall'interazione sinergica tra Oct4 e Sox2 (Tomioka M. et al., 2002). Il gene *Utf1* specifico per le cellule staminali, è regolato dall'azione sinergica di Oct4/Sox2 (Nishimoto M. et al., 1999).

Tali osservazioni rendono evidente la versatilità di Oct4, che può agire sia come soppressore di geni responsabili della differenziazione delle cellule staminali embrionali (ES) sia come transattivatore di geni noti nel mantenimento della pluripotenzialità delle cellule ES. Oct4 può quindi essere considerato uno dei fattori principali nel determinare il destino delle cellule staminali embrionali rivolte all'auto-rinnovo o alla differenziazione.

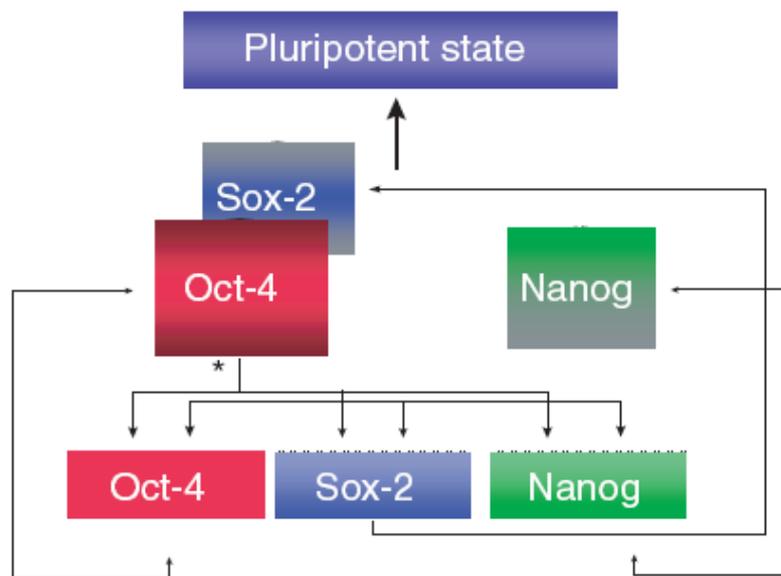


Fig. 1.15- Modello per la costituzione e la regolazione dello stato pluripotente; Oct4, Nanog e Sox2 sono regolatori chiave dello stato pluripotente.

Studi condotti su cellule di topo e cellule umane indicano Oct4 come un componente di un *network* di fattori di trascrizione, insieme alla proteina *omeobox* Nanog e al fattore di trascrizione Sox2, appartenente alla famiglia delle proteine HMG (*high mobility group*), (Fig.1.15); questi cooperativamente mantengono la pluripotenzialità nelle cellule staminali embrionali (Babaie Y. et al., 2007). Oltre a regolare la trascrizione di altri geni, Oct4, Nanog e Sox2 attivano uno la trascrizione dell'altro. Oct4 e Sox2 riconoscono insieme e legano una sequenza consenso altamente conservata, che è essenziale per l'espressione di Nanog sia nelle cellule staminali embrionali di topo che in quelle umane (Kuroda T. et al., 2005). Nanog è noto per l'interazione con una varietà di complessi modificanti la cromatina; l'analisi del suo *network* di interazioni ha rivelato il legame con repressori proteici. Un recente studio di interazione proteica ha identificato l'interazione di Oct4 e Nanog con un nuovo repressore che consiste in un complesso contenente alcune proteine associate alla cromatina (Liang J. et al., 2008).

In un recentissimo lavoro, il gruppo di ricerca di Yu et al. (2007) ha dimostrato che i quattro fattori Oct4, Sox2, Nanog e Lin28 sono sufficienti a riprogrammare le cellule somatiche umane in cellule staminali pluripotenti dotate delle caratteristiche essenziali delle cellule staminali embrionali. Queste cellule staminali umane la cui pluripotenza è stata indotta, presentano un cariotipo normale, mostrano attività telomerasica, esprimono marcatori di superficie cellulare e geni che caratterizzano le cellule ES umane, e mantengono inoltre il potenziale di sviluppo per la differenziazione in cellule di tutti e tre i foglietti germinali primari (Yu J. et al., 2007).

Oct4/A sembra essere coinvolto anche nel processo della tumorigenesi. Fenotipicamente, le cellule embrionali umane allo stadio del preimpianto sono simili alle cellule del cancro per molti aspetti, specialmente per la loro capacità di proliferare indefinitamente. Entrambi questi tipi cellulari vanno incontro a deprogrammazione per poi entrare in uno stato proliferativo, divengono immortali e, auto-rinnovandosi, acquisiscono capacità invasive. Per esempio, hanno messo in luce il ruolo chiave di Oct4 nei tumori delle cellule germinali testicolari umane (Gidekel S. et al., 2003); hanno inoltre rilevato l'espressione di Oct4 in linee cellulari di carcinoma mammario (Monk M. et al., 2001). Allo stesso modo delle cellule embrionali staminali, le cellule tumorali esibiscono un unico *pattern* di espressione genica e dovrebbero dunque sottostare al controllo di uno o più dei regolatori

per eccellenza, come Oct4.

1.6.4- Espressione e funzione di Oct-4B

E' stato più volte dimostrato che, durante lo sviluppo embrionale, entrambe le isoforme del fattore di trascrizione Oct-4 vengono espresse e, differenze di espressione di Oct4/A e Oct-4B, sono state evidenziate sia a livello spaziale sia a livello temporale (Cauffman G. et al., 2006).

L'espressione di Oct-4A, in embrioni allo stadio di sei cellule, viene rilevata, tramite analisi di immunofluorescenza, come una macchia molto debole a livello citoplasmatico, invece, un'espressione significativa a livello nucleare, si manifesta solo in embrioni a stadi più tardivi di sviluppo e in blastocisti.

L'espressione di Oct-4B è stata debolmente rintracciata nella regione periferica dello zigote, anche se, tramite saggi di immunofluorescenza, non si è rilevata la presenza della proteina allo stadio successivo, ovvero in embrioni allo stadio di due cellule; probabilmente il *pattern* di espressione ottenuto rappresenta il *pool* materno di proteine Oct-4B in via di degradazione: questo spiegherebbe il motivo per cui nello zigote l'isoforma B sembra essere presente mentre nello stadio successivo non si ottiene nessun segnale relativo alla proteina stessa. Nell'embrione allo stadio di quattro cellule, si ottiene un chiaro segnale citoplasmatico, riconducibile ad Oct-4B, che si protrae per tutti i successivi stadi di preimpianto dell'embrione.

La differenza più significativa, a livello di espressione proteica, tra l'isoforma A e l'isoforma B di Oct-4, risulta quella legata alla localizzazione cellulare: Oct-4A è una proteina maggiormente rappresentata a livello nucleare mentre Oct-4B è una proteina prettamente citoplasmatica. La localizzazione dell'isoforma B, che pur essendo un fattore di trascrizione risiede nel comparto citoplasmatico delle cellule, risulta abbastanza insolita: infatti, come già detto, entrambe le isoforme di Oct-4 presentano, internamente alla loro sequenza aminoacidica, una sequenza di localizzazione nucleare (*NLS*), che indurrebbe quindi ad aspettarsi un segnale, relativo ad Oct-4B, anche a livello del compartimento nucleare. Attualmente risulta ignoto il motivo per cui l'isoforma B di Oct-4 non traslochi a livello nucleare anche se si ritiene che questo potrebbe implicare ulteriori funzioni di Oct-4B, oltre

all'attivazione trascrizionale, al momento ancora non conosciute.

Quello che risulta assodato è che la proteina Oct-4B viene codificata dal DNA genomico embrionale e in uno stadio precedente rispetto ad Oct-4A. Tuttavia l'importanza della sua funzione nel mantenimento della staminalità e del *self-renewal* cellulare, risulta attualmente ancora incerta anche perché non si è riscontrata la presenza di alcuna proteina equivalente ad Oct-4B in topo: infatti il gene POU5F1 murino, che risulta anch'esso cruciale per lo sviluppo embrionale e che condivide l'87% di omologia di sequenza con il gene per Oct-4 umano, non codifica alcuna proteina simile a Oct-4B (Nichols J. et al., 1998).

Per quanto riguarda la funzione dell'isoforma B di Oct-4, Lee J. e i suoi collaboratori hanno dimostrato che probabilmente questa proteina è coinvolta nei processi di trasduzione dei segnali cellulari e che inoltre la sua sola espressione non è sufficiente per mantenere uno stato di indifferenziazione e di *self-renewal* cellulare. La mancanza di questa attività in Oct-4B, potrebbe essere associata sia alla perdita dell'abilità transattivante sia all'incapacità di legare specifiche sequenze a livello del DNA a causa della sua localizzazione citoplasmatica (Lee J. et al., 2006).

2- SCOPO

Lo scopo principale del lavoro affrontato e descritto in questa tesi è stato lo studio del fattore di trascrizione Oct-4A, in particolare la ricerca di nuovi interattori molecolari potenzialmente coinvolti nella sua regolazione.

Per quanto concerne l'isoforma A di Oct-4, grazie ai risultati ottenuti dagli studi di ricerca degli ultimi anni, si è riusciti a comprendere, con maggiore chiarezza, il ruolo chiave che esso svolge nella regolazione della pluripotenzialità e nella capacità di autorinnovo delle cellule staminali embrionali ed adulte umane. Infatti, esso risulta altamente espresso sia a livello delle cellule staminali embrionali murine sia a livello delle cellule staminali embrionali umane e la sua espressione si modifica quando le cellule differenziano e perdono quindi la propria pluripotenzialità. Inoltre, l'espressione di Oct-4A, è stata rilevata in numerose popolazioni di cellule staminali umane isolate da vari tessuti adulti e, poiché è possibile associare Oct-4A allo stadio di cellula pluripotente e indifferenziata, esso viene considerato uno specifico marcatore di staminalità in vari tessuti dell'adulto (Yeom Y.I et al., 1996; Niwa H., 2001).

Per quanto riguarda l'attività di attivazione trascrizionale di Oct-4A, sono stati identificati numerosi geni target sulla cui espressione essa può esercitare un'azione attivatoria o repressoria. Oct-4A infatti funziona da attivatore trascrizionale di quei geni coinvolti nel mantenimento di uno stato pluripotente indifferenziato mentre risulta reprimere l'espressione dei geni che vengono attivati durante il differenziamento cellulare.

L'interazione proteina-proteina risulta, chiaramente, alla base del modello di interazione cooperativa e sinergica finalizzata a determinare l'attività trascrizionale dei geni regolati e il comportamento delle proteine interagenti all'interno di una specifica *pathway* segnalatoria.

A questo scopo, l'attività sperimentale eseguita e descritta in questo lavoro di tesi si è focalizzata sugli effetti funzionali dell'attivazione della via RAS-MAPK, in particolare si è voluto chiarire le correlazioni esistenti con la localizzazione sub-cellulare e la degradazione via proteosoma. Nella seconda parte del lavoro svolto l'attenzione è stata rivolta allo studio specifico degli effetti dell'interazione tra Erk e Oct4-A.

L'obiettivo di questa Tesi è, in ultima analisi, quello di fare luce su tre aspetti, la fosforilazione, la localizzazione e la degradazione, che possono rappresentare importanti

meccanismi di regolazione dell'attività biologica di Oct-4A. Lo studio di tali meccanismi è fondamentale allo scopo di comprendere il funzionamento di Oct-4A e il suo ruolo nel mantenimento della staminalità e nel differenziamento.

3- MATERIALI E METODI

3.1- Modelli cellulari e mantenimento in coltura

Le linee cellulari utilizzate per gli esperimenti di questa tesi sono:

Ntera2: sono una linea di cellule staminali di carcinoma embrionale umano che condividono molte proprietà caratteristiche delle cellule staminali embrionali (Thomson J.A. Et al. 1998; Reubinoff B.E. et al., 2000). Tali peculiarità ne hanno consentito un largo utilizzo, ad esempio, nello studio dello sviluppo neurale umano oppure nella valutazione di cambiamenti precoci dello sviluppo (Pleasure S.J. et al., 1993; Miyazono M. et al., 1995). Ai fini del presente studio sono state scelte anche per la loro alta espressione del fattore di trascrizione Oct-4A.

Le cellule Ntera2 sono state coltivate utilizzando il terreno commerciale *DMEM* (Dulbecco's Modified Eagle's Medium, Lonza) ad alta concentrazione di glucosio (4.5g/l) e addizionato del 10 % di FBS (Fetal Bovine Serum) e mantenute in incubatore a 37°C, in un'atmosfera umida contenente 5% di CO₂.

Quando le cellule Ntera2 raggiungono un grado di confluenza pari a circa l'80%, vengono staccate meccanicamente, utilizzando uno *scraper*, dalla piastra su cui sono adese e passate 1:4. Viene utilizzato un metodo meccanico per staccare le cellule dalla piastra e non il metodo enzimatico con la tripsina, perché questo enzima sembrerebbe indurre, nelle cellule Ntera2, differenziamento cellulare.

HeLa derivate da carcinoma umano. Le cellule sono state mantenute in terreno di coltura DMEM (Dulbecco's Modified Eagle's Medium) addizionato del 10% di FBS (Fetal Bovin Serum), 2 mM L-glutamina, 50 µg/ml gentamicina in incubatore a 37°C, in un'atmosfera umida contenente 5% di CO₂. A differenza delle Ntera2 non esprimono a livello proteico Oct-4A e sono state quindi utilizzate per gli esperimenti di co-trasfezione transiente.

3.2- Trasfezioni

Le cellule impiegate negli esperimenti di trasfezione transiente (NTera2 e HeLa) vengono piastrate il giorno prima nei rispettivi terreni di coltura già descritti precedentemente in piastre del diametro di 100 mm² (Falcon) ad una densità dell'80% e 50% rispettivamente.

Il giorno successivo viene cambiato il terreno di coltura e dopo quattro ore viene effettuata la trasfezione.

Le trasfezioni sono state eseguite con il metodo della lipofectamina utilizzando il reagente LipofectamineTM Plus Reagent (Invitrogen) seguendo le indicazioni fornite dal produttore.

Le cellule sono state analizzate tramite immunoblotting dopo 24-48 ore dalla trasfezione.

Di seguito vengono riportati i plasmidi utilizzati in questi esperimenti di trasfezione transiente:

-p-MEK1^{CA} contenente la sequenza codificante per una forma costitutivamente attiva di MEK1 ottenuta mutando i residui di serina alla posizione 218 e 222 in acido aspartico, clonata nel vettore p-USE. Il plasmide presenta anche la sequenza codificante per il "tag" HA (emoagglutinina).

-p-FLAG-Oct4A contenente la sequenza codificante per Oct-4A assieme alla sequenza codificante per il peptide DYKDDDDK (FLAG) che ne consente l'isolamento

-p-USE utilizzato come controllo negli esperimenti di co-trasfezione al posto del plasmide p-MEK1^{CA}.

3.3- Clonaggio in vettore di espressione della proteina Oct-4A

Dopo aver eseguito l'estrazione dell'RNA, il cDNA corrispondente alla sequenza completa del fattore di trascrizione Oct-4A viene amplificato in PCR usando l'intero cDNA e oligonucleotidi opportunamente disegnati per inserire all'estremità 5' e all'estremità 3' i siti di riconoscimento per le endonucleasi di restrizione *BamHI* ed *EcoRI*. Successivamente il frammento ottenuto dall'amplificazione viene clonato nel vettore di espressione *pGEX4T1*.

3.3.1- Estrazione di RNA

Per l'isolamento dell'RNA totale dalle cellule della polpa dentale è stata utilizzato il kit "TRIzol[®] Plus RNA Purification Kit" (Invitrogen), nel rispetto delle condizioni sperimentali imposte dal produttore. Il "TRIzol[®] Reagent", che consiste in una soluzione monofasica di fenolo e isotiocianato di guanidina, mantiene l'integrità dell'RNA e porta invece alla distruzione delle cellule e alla dissoluzione dei componenti cellulari. Tutti i passaggi che prevedono l'utilizzo di RNA devono essere eseguiti in contenitori privi di RNAsi e, per evitare la degradazione dell'acido ribonucleico ad opera delle ribonucleasi, l'operatore deve sempre indossare guanti puliti e utilizzare soluzioni preparate con acqua demineralizzata sterile.

3.3.2- Reazione di retrotrascrizione

L'RNA estratto deve essere convertito in cDNA, DNA complementare a singolo filamento, che fungere da stampo nelle successive reazioni di amplificazione qualitativa e quantitativa. La formazione del cDNA avviene grazie all'appaiamento di oligonucleotidi commerciali (*oligo(dT)₁₂₋₁₈ primers*) all'mRNA precedentemente isolato, fornendo così l'aggancio, rappresentato da un gruppo ossidrilico iniziale, per l'attività dell'enzima trascrittasi inversa. I primers *oligo(dT)₁₂₋₁₈* che si legano alla coda di poli(A) dell'mRNA consentono di trascrivere l'intero messaggero di Oct-4A.

Innanzitutto, viene preparata una pre-mix contenente RNA (1 ng - 5 µg), 1 µl di primers oligo(dT) (500 µg/ml; Invitrogen), 1 µl di dNTPs (10 Mm, Invotrogen), H₂O mQ per un volume finale di 12 µl. La pre-mix viene successivamente denaturata a 65°C per 5 minuti nel termociclatore. Per ogni campione vengono di seguito aggiunti: 4 µl di First-Strand Buffer 5x (Invitrogen), 2 µl di DTT (0.1 M), 1µl di RNasi-OUT (40U/µl; Invitrogen) ed 1 µl della retrotrascrittasi "SuperScript[™] II" (200U/µl; Invitrogen). La reazione di retrotrascrizione viene eseguita alle seguenti condizioni:

| | |
|------|--------------------------------------------------------------------|
| 42°C | 2 minuti (prima di aggiungere la "SuperScript [™] II RT") |
| 42°C | 60 minuti |

70°C

15 minuti

3.3.3- PCR (*Polymerase Chain Reaction*)

La PCR (*Polymerase Chain Reaction*) è una tecnica efficace che permette l'amplificazione esponenziale *in vitro* di una determinata sequenza di cDNA, precedentemente preparato con reazione di retrotrascrizione, sfruttando l'attività di sintesi della *Taq* polimerasi, estratta dal batterio *Thermus aquaticus*.

È necessario conoscere le estremità della sequenza da amplificare con sufficiente precisione per poter sintetizzare degli oligonucleotidi (*primers senso* e *antisense*) che saranno ibridizzati ad esse. Il prodotto della reazione di amplificazione sarà una molecola discreta di dsDNA con estremità corrispondenti a quelle 5' degli oligonucleotidi utilizzati.

In questo caso al posto della *Taq*, è stato utilizzato l'enzima *Pfu* (*Pfu DNA Polymerase*) derivante dall'Archea *Pirococcus furiosus*, che grazie alla sua attività esonucleasica 3'→5', garantisce una maggiore fedeltà di trascrizione, necessaria ai fini del clonaggio.

L'amplificazione si realizza in un processo ciclico in tre tempi e i requisiti essenziali della PCR sono: (1) due inneschi oligonucleotidici di sintesi (~ 20 nucleotidi ciascuno) complementari a regioni dei filamenti opposti che fiancheggiano la sequenza di DNA bersaglio e che, dopo l'ibridazione con il DNA di partenza, abbiano la proprie estremità ossidriliche 3' orientato l'una verso l'altra; (2) una sequenza bersaglio nel campione di DNA che si trovi tra l'uno e l'altro innesco e sia lunga da 100 a ~ 35000 bp; (3) una DNA-polimerasi termostabile capace di resistere a 95°C e più; (4) i quattro deossiribonucleotidi.

Il tipico processo della PCR comprende numerosi cicli per amplificare una sequenza di DNA e ciascun ciclo prevede tre fasi successive:

- ***Denaturazione***. La prima fase del sistema di amplificazione della PCR è costituito dalla denaturazione termica del campione di DNA mediante riscaldamento in provetta a 94-95°C, al fine di separare le due catene di DNA rendendo possibile l'accesso dei *primers* al filamento complementare. La provetta conterrà, oltre al DNA di partenza, anche un ampio eccesso dei due *primers*, una DNA-polimerasi

termostabile e i quattro desossiribinucleotidi.

- Rinaturazione. In questa seconda fase si abbassa lentamente la temperatura della miscela per permettere ai *primers* di appaiarsi alla sequenza complementare di DNA. La temperatura utilizzata in questo caso viene scelta in base alla temperatura di *melting*, cioè la temperatura massima oltre la quale il *primer* si denatura e perde la capacità di appaiarsi in modo specifico alla sequenza di cDNA.
- Sintesi. Nella terza fase si innalza la temperatura fino a raggiungere quella ottimale (*optimum* di temperatura) per il funzionamento catalitico della DNA-polimerasi scelta nella propria reazione di amplificazione. La sintesi ha inizio all'estremità ossidrilica 3' di ciascun innesco. In questo caso, la *Pfu* polimerasi utilizzata per l'amplificazione ha una *optimum* di temperatura di circa 72°C (Fig. 3.1).

In genere i risultati migliori si ottengono con frammenti non superiori a 1.5 kb. Il numero di cicli non deve superare la cinquantina, poiché al di sopra l'amplificazione non è più esponenziale ma raggiunge un *plateau* dovuto a carenza di oligonucleotidi, carenza di dNTPs, aumento di gruppi pirofosfato.

Di seguito vengono riportate le sequenze dei *primers* utilizzati, con i siti di riconoscimento specifici per gli enzimi di restrizione BamHI ed EcoRI. Infatti, per la fase successiva di clonaggio nel sito *polylinker* del vettore, l'inserto deve avere le terminazioni compatibili con le estremità del vettore linearizzato in cui deve essere inserito:

primer senso (Oct-4A F):

F 5' - GATC**GGATCC**ATGGCGGGACACCTGGCT - 3'
} BamHI

primer antisenso (Oct-4A R):

R 5' - GATC**GAATTC**TCAGTTTGAATGCATGGG - 3'
} EcoRI

Per ogni campione da amplificare è stata preparata una miscela di reazione di 50 µl

contenente:

- 5 µl di Pfu buffer 10x con MgSO₄
- 1 µl di dNTPs 10 mM
- 0.3 µl di primer F 100pmoli/µl
- 0.3 µl di primer R 100pmoli/µl
- 40.4 µl di H₂O mQ
- cDNA stampo (50 pg – 1µg/50µl)

All'ultimo momento è stato aggiunto 1µl di *Pfu* (Fermentas 2.5 U/µl) e il campione è stato denaturato a 95°C per 2 minuti. La reazione è quindi stata impostata per 40 cicli alle seguenti condizioni:

- denaturazione 35 secondi a 94°C
- rinaturazione 40 secondi a 62.5°C
- sintesi 2 minuti e 30 secondi a 72°C

L'ultimo ciclo prevede uno stadio finale di 10 minuti a 72°C.

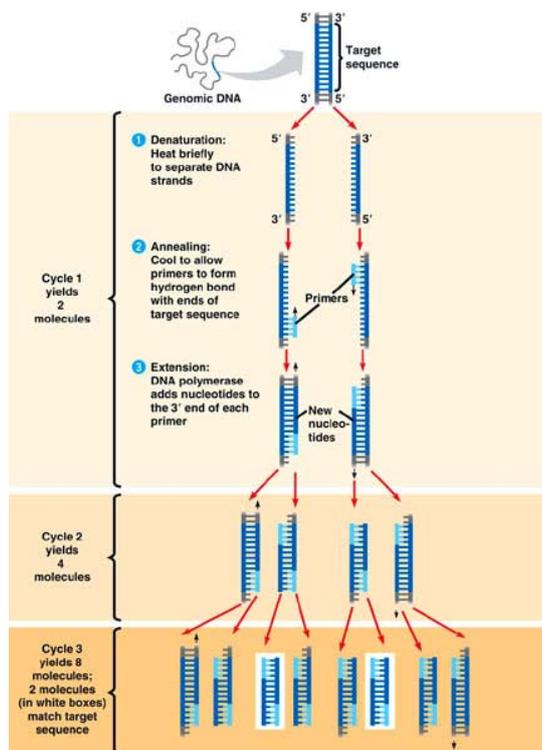


Fig. 3.1- Illustrazione schematica della tecnica della PCR

3.3.4- Elettroforesi su gel di agarosio per acidi nucleici

L'elettroforesi su gel di agarosio è un metodo semplice e veloce che permette di separare, e quindi identificare, frammenti di DNA in base al loro peso molecolare.

L'agarosio, che viene estratto dalle alghe, è un polisaccaride lineare e neutro formato da unità di D-galattosio e di 3,6-anidro-L-galattosio legate alternativamente da legami glicosidici.

L'agarosio, se fuso e gelificato, forma una matrice la cui porosità dipende dalla sua concentrazione iniziale e che consente la separazione dei frammenti di DNA contenuti nei campioni caricati.

I frammenti migrano, nel campo elettrico che attraversa il gel, dal polo negativo a quello positivo, in funzione delle cariche elettriche conferitegli dai gruppi fosfato (PO_4^{3-}). La velocità di migrazione dipende:

- dalle dimensioni dei frammenti;
- dalla percentuale di agarosio nel gel;
- dal voltaggio applicato.

Frammenti lineari più piccoli migrano più velocemente rispetto a quelli più grandi, mentre, a parità di peso molecolare, il DNA circolare migra più velocemente di un DNA lineare, in quanto assume una conformazione detta superavvolta (supercoiled DNA).

Per eseguire la separazione elettroforetica di acidi nucleici è necessario preparare un gel sospendendo l'agarosio nel tampone *TBE* alla concentrazione 0.5x (ovvero diluito 0,5 volte). Il *TBE* è costituito da Tris, EDTA e Acido Borico: il Tris viene utilizzato nel tampone per mantenere il valore di pH costante, l' EDTA è un chelante di cationi bivalenti e risulta importante perché la maggior parte delle nucleasi richiede cationi bivalenti per il corretto funzionamento; l'Acido Borico fornisce invece un'appropriata forza ionica al tampone. Successivamente la soluzione viene addizionata di Etidio Bromuro (5 μ l/100ml), ~~che~~ agente intercalante molto utile per evidenziare il DNA durante la corsa elettroforetica. L'Etidio Bromuro infatti, se stimolato da raggi ultravioletti (260-360 nm), emette fluorescenza arancione a 560nm: è quindi possibile, osservando il gel ad un transilluminatore, individuare i diversi campioni caricati.

Il gel d'agarosio agisce da separatore molecolare: le maglie reticolate del gel permettono il passaggio di molecole di ugual dimensione o più piccole, ma ostacolano quello di molecole di dimensione maggiori, rallentandone la corsa. La misura delle maglie è funzione della concentrazione iniziale d'agarosio: con concentrazioni più basse di agarosio, si otterranno maglie larghe che sono maggiormente adatte a separare frammenti di grandi dimensioni, mentre concentrazioni più elevate di agarosio porteranno alla formazione di maglie strette, utilizzate per separare frammenti di piccole dimensioni

In questo lavoro di tesi si è comunemente lavorato con gel d'agarosio al 1.5%, adatti a separare frammenti da 100 a 1000 paia di basi. Prima di essere caricati, ai campioni viene aggiunto un tampone di caricamento, contenete blu di Bromofenolo, che serve ad appesantire i campioni, permettendone la permanenza nei pozzetti, e funge da *marker* del fronte di migrazione. Inoltre, per poter stabilire la lunghezza dei frammenti di DNA in studio, si confrontano i campioni con uno specifico standard di riferimento costituito da una serie di frammenti di DNA di lunghezza nota (INVITROGEN); normalmente la separazione elettroforetica viene condotta per circa 30 minuti a 150 V.

3.3.5- Estrazione da gel

Per eseguire l'estrazione dal gel di agarosio della banda corrispondente al campione d'interesse (vettore o inserto), è stato utilizzato il kit "JETquick Gel Extraction" (GENOMED).

Tale metodica necessita dell'utilizzo di due soluzioni:

| | |
|--------------|-----------------------------------------------------------------------------------------------------|
| Soluzione L1 | NaClO ₄ , Sodio acetato, TBE (solubilizzatore) (da conservare a temperatura ambiente) |
| Soluzione L2 | Etanolo, NaCl, EDTA e Tris/HCl (da conservare a temperatura ambiente) |

Tabella 3.1: Soluzioni utilizzate nella tecnica dell'estrazione di bande da gel

Dopo la corsa elettroforetica e l'individuazione della banda relativa al campione oggetto di studio, si procede tagliando, mediante l'utilizzo di un bisturi sterile, la banda dal gel di agarosio che viene poi trasferita in una eppendorf. Come consigliato da protocollo, per ogni 100 mg di peso della banda, vanno aggiunti 300 μ l di soluzione L1. Quindi si incuba il tutto a 50°C per 15 minuti sul *thermomixer* 5436 (Eppendorf) agitando, mediante l'uso di un vortex, ogni 3 minuti per assicurare la completa solubilizzazione dell'agarosio. Il campione viene prelevato, trasferito in una colonnina (JETQUICK spin column), che viene successivamente inserita in un tubo di raccolta, e sottoposto ad una centrifugazione nella microcentrifuga a 12.200 x g per 1 minuto a 20°C. Scartato l'eluato, si aggiungono alla colonnina 500 μ l di soluzione L2. Si centrifuga nuovamente a 12.200 x g per 1 minuto, si elimina l'eluato e si ripete la centrifugazione alle stesse condizioni. Alla colonnina, che viene posta in una nuova eppendorf, vengono successivamente addizionati 50 μ l di H₂O sterile scaldata a 65°C. Infine si centrifuga a 12200 x g per 2 minuti e si ottiene così il campione purificato che viene conservato a - 80°C (Fig. 3.2).

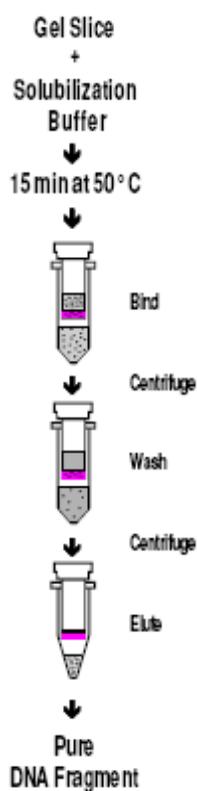


Fig. 3.2- Illustrazione schematica della tecnica di estrazione da gel.

3.3.6- Quantificazione su gel d'agarosio

Il campione (vettore o inserto) viene caricato su un gel di agarosio e al termine della corsa elettroforetica viene quantificato. La concentrazione del campione viene stimata paragonando l'intensità della banda ad esso corrispondente con l'intensità delle bande a concentrazione nota del *marker*. Il marker utilizzato è: *Mass RulerTM DNA Ladder Low Range* (Fermentas).

3.3.7- Il vettore di espressione procariotica pGEX4T1

Il vettore di espressione *pGEX4T1* (Fig.3.3), disponibile in commercio (Amersham Biosciences), è un plasmide di circa 4.9 kb. Tale vettore contiene:

- un promotore *tac* per l'espressione ad alti livelli di una proteina di fusione (GST-proteina). *Tac*, che è un promotore forte, è un costrutto realizzato *in vitro* che comprende la regione -10 (ovvero 10 coppie di nucleotidi a monte del sito di inizio della trascrizione) del promotore *lac* e la regione -35 del promotore *trp*. L'induzione o l'accensione del promotore *tac* si può ottenere chimicamente con l'IPTG (isopropil β-D tiogalattoside);
- un'origine di replicazione batterica *oriC*;
- il gene *amp^r*, che conferisce resistenza all'antibiotico ampicillina necessario per la selezione positiva dei batteri trasformati;
- un sito di taglio per la trombina;
- l'operone batterico *lacI^q*, per l'induzione controllata del promotore. Infatti il prodotto genico di *lacI^q* è una proteina-repressore che si lega all'operatore del promotore *tac* che impedisce l'espressione dell'inserto fino al momento dell'induzione con l'IPTG ed esercita quindi uno stretto controllo sull'espressione dell'inserto;
- un sito di clonaggio multiplo (*MCS, Multiple Cloning Site*) (Fig. 3.3).

La peculiarità di questo plasmide sta nel possedere, a monte del sito di clonaggio, il gene che codifica per l'enzima glutatione S-transferasi (*GST*) che andrà ad inserirsi *in frame* all'estremità N-terminale della proteina di interesse.

Il frammento di DNA, costituito dalla sequenza codificante per la proteina di interesse (cDNA), viene inserito nel vettore per la successiva trasformazione delle cellule di *E. coli*. È da notare che, sebbene l'espressione in *E. coli* sia efficiente, nei procarioti non è presente il macchinario responsabile delle modificazioni post-traduzionali.

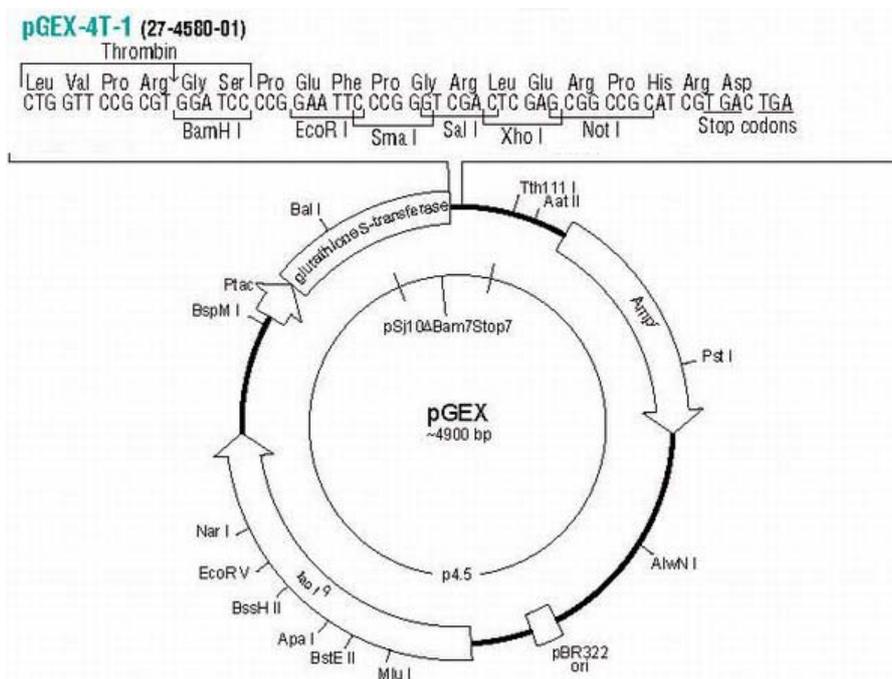


Fig. 3.3- Schema del vettore dove è inserito il cDNA completo della proteina Oct-4A. www.pasteur.ac.ir

3.3.8- Digestione e defosforilazione del vettore

Il vettore di espressione pGEX4T1 è stato digerito e defosforilato in un'unica reazione. Per eseguire una doppia digestione del vettore con gli enzimi di restrizione specifici, *BamHI* ed *EcoRI* (Fermentas), è stata preparata una soluzione di reazione composta da:

- DNA plasmidico 20 ng totali;
- 4 µl di buffer Tango 10x, uso finale 2x (Fermentas), adatto per entrambi gli enzimi di restrizione;
- 2 µl di BamHI (10 U/µl);
- 1 µl di EcoRI (10 U/µl);
- 12 µl di H₂O per un volume finale di 20 µl.

Il tutto è stato incubato a 37°C per 1 ora in stufa. Successivamente è stata aggiunta 1 unità di fosfatasi *Shrimp Alkaline Phosphatase (SAP)* (Fermentas) (1U/μl) e si è messo nuovamente ad incubare a 37°C per un'altra mezz'ora.

Al termine del processo è stata seguita l'inattivazione al calore degli enzimi coinvolti nella reazione, 80°C per 20 minuti. Infatti, mentre *EcoRI* e la *SAP* si inattivano a 65°C, *BamHI* viene inattivato solo parzialmente a 80°C.

L'eliminazione dei gruppi fosfato presenti alle estremità generate in seguito al taglio proteolitico degli enzimi di restrizione, consente di ridurre notevolmente la probabilità che il vettore si richiuda su se stesso e ricircularizzi.

Il vettore così digerito e purificato dagli enzimi, è stato sottoposto ad una corsa elettroforetica su gel di agarosio al fine di separare la frazione di vettore in forma superavvolta da quella in forma lineare. Eseguendo quindi una purificazione della banda dal gel, è stato possibile estrarre solamente il vettore in forma linearizzata che è stato poi utilizzato per la successiva reazione di ligasi.

3.3.9- Ligasi vettore-inserito

La reazione di ligasi è stata allestita con quantità di vettore (pGEX4T1) e inserto (cDNA di Oct-4A) in rapporto 1:3 in un volume finale di 20 μl contenente ATP 1 mM, 1 μl dell'enzima ligasi del fago T4, "T4 DNA Ligase" (Fermentas) (5 U/μl) ed il tampone 10x fornito dalla ditta. La miscela è stata incubata a 16°C *over night*. La ligasi deve poi essere inattivata a 65°C per 10 minuti.

Le estremità coesive del vettore e dell'inserto, essendo tra loro complementari, si appaiano e vengono legate covalentemente nella reazione catalizzata dalla ligasi. La miscela di ligasi è stata utilizzata per trasformare le cellule *E. coli* BL21 (Promega).

In parallelo è stata eseguita una *ligation* con il solo vettore senza l'inserto, in modo da ottenere un controllo negativo che consentisse di verificare quanto del vettore si era ricircularizzato.

3.3.10- Preparazione di cellule competenti

Una colonia di *E.coli* BL21 viene piastrata su una piastra di agar priva di ampicillina e successivamente incubata a 37°C per circa 16-20 ore. A questo punto si preleva una colonia per farla crescere in terreno LB incubandola a 37°C *over night*. Il giorno successivo si prelevano 500 µl di LB, nel quale sono cresciuti i batteri, per essere inoculati in 50 ml di Medium A (terreno LB addizionato di MgSO₄, H₂O e glucosio allo 0.2%) in una beuta da 500 ml.

La crescita batterica viene effettuata a 37°C fino ad ottenere una lettura alla lunghezza d'onda di 600 nm, di circa 0.3-0.4 OD. Si raccolgono i batteri cresciuti in falcon, le quali vengono successivamente centrifugate a 1500 x g per 10 minuti a 4°C. Da questo passaggio in poi, tutte le operazioni vengono effettuate in ghiaccio. Il pellet viene risospeso in 500 µl di Medium A a cui vengono successivamente aggiunti 2.5 ml di "Storage solution" (terreno LB addizionato di MgSO₄, H₂O 12mM, glicerina 36%, PEG 12%). Dopo aver mescolato bene si fanno delle aliquote che andranno conservate a – 80°C.

3.3.11- Trasformazione dei batteri *E. coli* BL21

A 50 µl di batteri BL21 resi competenti, sono stati aggiunti 10 µl della reazione di ligasi effettuata tra vettore-inserito. Quindi, le cellule batteriche sono state sottoposte ad una serie di fasi, di seguito riportate:

- 30 minuti in ghiaccio (shock termico)
- 90 secondi a 42°C
- 5 minuti in ghiaccio

Successivamente sono stati addizionati 200 µl di Medium SOC (Bacto[®]-tryptone 2.0 g, Bacto[®]-yeast extract 0.5 g, NaCl 1M 1 ml, KCl 1M 0.25 ml, Mg²⁺ soluzione stock 2M 1 ml, glucosio 2M filtrato- sterilizzato 1 ml), portato in precedenza a temperatura ambiente, e i batteri sono stati incubati a 37°C per un'ora nell'agitatore. In seguito i batteri sono stati piastrati su piastre di LB agar e ampicillina, che sono state poste in stufa a 37°C *over night*.

3.4- Miniprep

Questa metodica consente di estrarre quantità di DNA plasmidico fino a 20 µg da una coltura batterica. La tecnica richiede l'utilizzo di quattro diverse soluzioni:

| | |
|--------------|------------------------------------------------------------------------------------------|
| Soluzione P1 | Glucosio 50 mM, Tris-HCl 25 mM pH=8, EDTA 10 mM, RNasi-A 100 µg/ml (da conservare a 4°C) |
| Soluzione P2 | NaOH 0.2 M, SDS 0.1% (da preparare al momento) |
| Soluzione P3 | Potassio acetato 3 M, acido formico 4.9% (da conservare a temperatura ambiente) |
| Soluzione P4 | Miscela fenolo/cloroformio/alcol isoamilico (25:24:1) (da conservare a 4°C) |

Tabella 3.2- Soluzioni utilizzate nella tecnica delle miniprep

Ogni colonia batterica, derivante dalla trasformazione con il plasmide d'interesse, viene inoculata in 4 ml di terreno LB (Bacto[®]-tryptone 10g/L, Bacto[®]-yeast extract 5g/L, NaCl 5g/L), addizionato dell'appropriato antibiotico selettivo (ampicillina 50 µg/ml), e lasciata over night a 37°C in agitazione a 220 rpm. Il giorno seguente vengono trasferiti 1.5 ml della coltura totale batterica in una eppendorf e si procede, con questi, all'estrazione del DNA plasmidico. Ciò che resta della coltura viene in parte utilizzata per preparare una serie di stock in glicerolo (250 µl di coltura batterica+250 µl di glicerolo 40%, conservati a -80°C), e in parte conservata in frigo a 4°C, per brevi periodi.

La sospensione batterica è stata quindi centrifugata a massima velocità per circa 1-2 minuti per pellettare i batteri che, dopo aver scartato il surnatante, sono stati risospesi in 100 µl di soluzione P1; si agita il tutto, mediante l'uso di un vortex, per 10 secondi e si aggiungono 200 µl di soluzione P2, che serve da lisante. Dopo aver invertito per qualche minuto, alla miscela di batteri vengono addizionati 150 µl di soluzione P3 e, dopo aver nuovamente invertito per un paio di volte la eppendorf, si incuba 5 minuti in ghiaccio. Successivamente si centrifuga la soluzione per 5 minuti a massima velocità e al surnatante, che è stato trasferito in una nuova eppendorf, vengono aggiunti 450 µl di soluzione P4. Si centrifuga nuovamente per 3 minuti sempre alla massima velocità. A questo punto la soluzione apparirà suddivisa in due fasi distinte: una fase acquosa e una fase organica. Prelevata la fase acquosa, nella quale è

contenuto il DNA plasmidico, vengono addizionati 270 μ l di isopropanolo, necessario per far precipitare il DNA. La soluzione viene agitata per 10 secondi, sempre mediante l'uso di un vortex, e successivamente incubata a temperatura ambiente per 5 minuti. Il campione viene centrifugato per 10 minuti alla massima velocità e, una volta scartato il surnatante, vengono aggiunti 50 μ l di etanolo 70% per lavare il DNA plasmidico. La soluzione viene centrifugata per 3 minuti alla massima velocità e si cerca poi di rimuovere l'etanolo nel modo più accurato possibile. La eppendorf viene comunque lasciata per qualche minuto sotto cappa a flusso laminare, per favorire la completa evaporazione dell'etanolo dal campione. Infine il pellet viene risospeso in 15 μ l di acqua mQ.

3.5- Miniprep eseguite mediante l'utilizzo del kit Eurogold (Euroclone)

Dopo che le colone batteriche, trasformate con il plasmide d'interesse, sono state espanse in 5 ml di terreno LB addizionato di ampicillina (100x) *over night* a 37°C in agitazione, vengono centrifugate a 3.600 x g per 20 minuti. Successivamente viene eliminato accuratamente il surnatante e, a ciascun pellet, vengono addizionati 250 μ l di “*Solution I/RNase A*”. Al fine di risospesare adeguatamente il pellet di cellule batteriche, queste vengono agitate per alcuni secondi mediante l'utilizzo di un vortex. In seguito vengono aggiunti, alle eppendorf contenenti le cellule batteriche risospese in “*Solution I/RNase A*”, 250 μ l di “*Solution II*” e i campioni vengono messi in agitazione su un ruota meccanica finchè i lisati non acquisiscono una colorazione chiara. E' molto importante non agitare troppo vigorosamente i campioni per evitare il danneggiamento del DNA plasmidico che porterebbe ad una riduzione della sua purezza. A questo punto i campioni vengono incubati per 2 minuti a temperatura ambiente.

Alle eppendorf contenenti i lisati batterici vengono aggiunti 350 μ l di “*Solution III*” e, successivamente, vengono poste in agitazione su una ruota meccanica fin tanto che non compare, nel campione, un precipitato bianco. In seguito si centrifuga il tutto a 10.000 x g per 10 minuti a temperatura ambiente.

Il surnatante di ciascuna eppendorf, viene trasferito nella corrispondente colonnina “*HiBind® miniprep column*”, inserita a sua volta in un tubo di raccolta da 2 ml (*collection tube*); le colonnine e i rispettivi tubi di raccolta vengono centrifugati a 10.000 x g per 1

minuto a temperatura ambiente, in modo tale che il lisato passi attraverso la membrana filtrante e che il DNA plasmidico si leghi ad essa.

Viene eseguito poi un primo lavaggio aggiungendo in ciascuna colonnina 500 μ l di “*HB buffer*”; le colonnine e i rispettivi tubi di raccolta vengono centrifugati a 10.000 x g per 1 minuto a temperatura ambiente e, in seguito, l'eluato viene scartato. Successivamente vengono eseguiti altri due lavaggi, alle medesime condizioni, utilizzando 750 μ l di *Wash Buffer* (HB buffer) addizionato con etanolo.

Segue un'ulteriore centrifugazione dei campioni sempre a 10.000 x g per un minuto a temperatura ambiente. Questo passaggio è molto importante perché permette di far asciugare la matrice delle colonnine, a cui è legato il DNA plasmidico, e di rimuovere quindi eventuali residui di etanolo presenti.

Infine segue la fase di eluizione del DNA plasmidico, che si esegue ponendo ciascuna colonnina in una nuova eppendorf e aggiungendo 50 μ l di acqua deionizzata. Il tutto viene centrifugato a 5.000 x g per 1 minuto ottenendo così il DNA plasmidico in soluzione acquosa.

Normalmente la prima eluizione contiene il 75-80% del DNA legato alla matrice; si può eventualmente procedere con una seconda eluizione per raccogliere eventuali residui di DNA plasmidico che risulta ancora legato alla matrice.

3.5.1- Analisi di restrizione

L'analisi di restrizione viene eseguita per verificare quale dei cloni, espansi e sottoposti a estrazione del DNA plasmidico, sono stati correttamente trasformati e risultano quindi in possesso dell'inserito d'interesse. Si esegue una doppia digestione del DNA plasmidico con gli enzimi BamHI e EcoRI: in un volume finale di 20 μ l, vengono aggiunti:

- 13 μ l di DNA plasmidico
- 4 μ l di Buffer Tango 10X (l'utilizzo di questo buffer permette di eseguire simultaneamente la doppia digestione con BamHI e EcoRI)
- 2 μ l di enzima BamHI (10 U/ μ l)
- 1 μ l di enzima EcoRI (10 U/ μ l; gli enzimi BamHI e EcoRI devono essere addizionati in rapporto 2:1)

L'eppendorf contenente la soluzione viene poi lasciata in stufa a 37°C per circa un'ora. Successivamente il campione viene posto a 80°C nel *thermomixer* (5436 Eppendorf) per permettere l'inattivazione degli enzimi di restrizione, anche se l'enzima BamHI viene inattivato solo parzialmente. Prima di caricare su gel di agarosio 1% (0.45g di agarosio disciolti in 45 ml di TBE 0.5X +2.5 µl di Etidio-Bromuro), al campione deve essere addizionato il *DNA Loading Dye 6X* (Tris-HCl 10 mM pH=7.6, Blu-bromofenolo 0.03%, Glicerolo 60%, EDTA 60 mM).

Una volta terminata la corsa, il gel viene osservato al transilluminatore e si verifica, quindi, se nei campioni caricati vi sono bande ad un'altezza paragonabile a quella dell'inserto d'interesse.

3.6- Midiprep

Questa tecnica permette di ottenere quantità di DNA plasmidico finqua 100 L'estrazione del DNA plasmidico è stata eseguita utilizzando il “QIAGEN Plasmid Midi Kit” (QIAGEN).

La metodica richiede l'utilizzo di diverse soluzioni:

| | |
|--------------|------------------------------------------------------------------------------------------------------------------------|
| Soluzione P1 | Tris-HCl 50 mM pH=8, EDTA 10 mM, RNasi-A 100 µg/ml |
| Soluzione P2 | NaOH 200 mM, SDS 0.1% (w/v) |
| Soluzione P3 | Potassio acetato 3 M pH=5.5 |
| Tampone QBT | NaCl 750 mM, MOPS [Acido 3-(N morfolino) propansulfonico] 50 mM pH=7, Isopropanolo 15% (v/v), Triton X-100 0.15% (v/v) |
| Tampone QC | NaCl 1 M, MOPS 50 mM pH=7, Isopropanolo 15% (v/v) |
| Tampone QF | NaCl 1.25 mM, Tris-HCl 50 mM pH=8.5, Isopropanolo 15% (v/v) |

Tabella 3.3- Soluzioni utilizzate nella tecnica delle midiprep

Si recuperano gli stock in glicerolo dei cloni che, in seguito ad analisi di restrizione enzimatica, sono risultati contenere un inserto della lunghezza attesa; quest'ultimi vengono successivamente espansi in 50 ml di terreno LB, addizionato dell'appropriato antibiotico selettivo (ampicillina 50 µg/ml). Il giorno seguente, i batteri vengono centrifugati per 20 minuti a 3.600 rpm alla temperatura di 4°C e risospesi in 4 ml di soluzione P1 fredda. Si trasferisce il tutto in un tubo autoclavato da ultracentrifuga. La reazione di lisi viene indotta aggiungendo 4 ml di soluzione P2: il campione tende ora ad assumere un aspetto viscoso, parzialmente colorato di blu e deve essere invertito lentamente 4-6 volte finché non acquisisce una colorazione omogenea; a questo punto si incuba per 5 minuti a temperatura ambiente. Per neutralizzare la reazione di lisi vengono aggiunti 4 ml di soluzione P3 fredda (si inverte la falcon lentamente per un paio di volte finché la soluzione non ritorna di colore bianco) e successivamente si incuba per 30 minuti in ghiaccio, per facilitare la precipitazione del DNA genomico, proteine, SDS e detriti cellulari. Dopo aver centrifugato per 30 minuti a 13.000 rpm a 4°C, il surnatante viene caricato in una colonnina (QUIAGEN-tip 100), precedentemente equilibrata con 4 ml di tampone QBT. Vengono poi effettuati due lavaggi della colonnina con 10 ml di Buffer QC. Il DNA viene eluito in 5 ml di tampone QF, successivamente addizionato a 3.5 ml di isopropanolo e centrifugato per 60 minuti a 3.600 rpm alla temperatura di 4°C. Il DNA precipitato, corrispondente al pellet, viene lavato con 2 ml di etanolo freddo al 70%, dopo aver rimosso con attenzione il surnatante. Il campione viene nuovamente centrifugato a 3.600 rpm per un'ora a temperatura ambiente. Dopo aver accuratamente eliminato l'etanolo presente nella soluzione, sia tramite aspirazione sia tramite evaporazione a temperatura ambiente, il precipitato viene risospeso in 50 o 100 µl di acqua mQ sterile e quantificato allo spettrofotometro (Fig. 3.4).

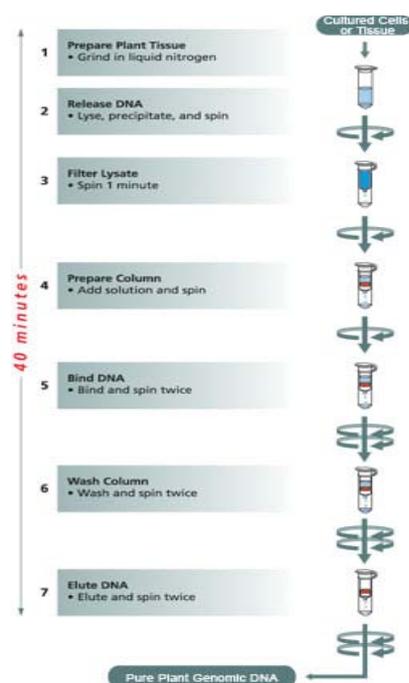


Fig. 3.4- Illustrazione schematica del metodo di purificazione del DNA plasmidico midiprep. www.fermentas.com

3.7- Maxiprep

Questa tecnica consente di estrarre quantità di DNA plasmidico fino a 500 µg. L'estrazione del DNA plasmidico è stata eseguita utilizzando il “QIAGEN Plasmid Maxi Kit” (QIAGEN). La tecnica richiede l'utilizzo di una serie di soluzioni:

| | |
|--------------|------------------------------------------------------------------------------------------------------------------------|
| Soluzione P1 | Tris-HCl 50 mM pH=8, EDTA 10 mM, RNasi-A 100 µg/ml |
| Soluzione P2 | NaOH 200 mM, SDS 0.1% (w/v) |
| Soluzione P3 | Potassio acetato 3 M pH=5.5 |
| Tampone QBT | NaCl 750 mM, MOPS [Acido 3-(N morfolino) propansulfonico] 50 mM pH=7, Isopropanolo 15% (v/v), Triton X-100 0.15% (v/v) |
| Tampone QC | NaCl 1 M, MOPS 50 mM pH=7, Isopropanolo 15% (v/v) |
| Tampone QF | NaCl 1.25 mM, Tris-HCl 50 mM pH=8.5, Isopropanolo 15% (v/v) |

Tabella 3.4- Soluzioni utilizzate nella tecnica delle maxiprep

La procedura utilizzata dal kit è identica a quella seguita per le midiprep, l'unica differenza consiste nella maggiore quantità di coltura batterica di partenza (100 ml) e di conseguenza nell'uso di colonnine più capienti (QUIAGEN-tip 500) e di volumi maggiori delle soluzioni elencate in tabella.

3.8- Sequenziamento

Il DNA plasmidico, incluso dell'inserto codificante la proteina Oct-4A e ottenuto al termine dell'esecuzione delle Midiprep, è stato sequenziato in Germania presso l'azienda MWG.

Per il sequenziamento, il campione è stato preparato nel modo seguente:

- 150 ng/μl di plasmide (con l'inserto) per un totale di 30 μl (l'analisi richiede un volume minimo di 15 μl);
- *primers* alla concentrazione di 2 pmoli/μl.

I *primers* utilizzati per il sequenziamento sono interni alla sequenza nucleotidica di Oct-4A e vengono di seguito riportati:

primer senso:

F 5' - GACAACAATGAAAATCCTTCAGGAGA - 3'

primer antisenso:

R 5' - TTCTGGCGCCGGTTACAGAACCA - 3'

Il *primer senso* cade a cavallo degli esoni 2 e 3, mentre l'*antisenso* si localizza a livello dell'esone 4.

3.9- Trasformazione del ceppo batterico KRX

Le cellule batteriche *KRX* sono state trasformate con i seguenti plasmidi: *pGEX4T1* e *pGEX-Oct-4A*.

Per ogni plasmide con cui si intende trasformare le cellule batteriche, è stato scongelato in ghiaccio uno stock di *KRX* (Promega). Per la trasformazione dei batteri sono stati aggiunti 30 ng di plasmide (il range è 1-50 ng) in un volume non superiore a 5 μ l. Il campione è stato agitato delicatamente ed incubato in ghiaccio per 10 minuti (l'intervallo è 5-30 minuti). Si è quindi eseguito lo shock termico a 42°C per 15 minuti. Dopo una seconda incubazione in ghiaccio di 2 minuti, sono stati aggiunti 450 μ l di *SOC*. A questo punto si è incubato a 37°C per un'ora in agitazione a 220 rpm. 100 μ l prelevati dal campione, sono stati piastrati su LB agar e ampicillina e incubati a 37°C *over night* in stufa.

3.10- Induzione dell'espressione proteica

Ogni singola colonia, derivante dalla coltura *over night* di batteri *KRX* trasformati, è stata inoculata in 5 ml di *Terrific Broth* (Bacto[®]-tryptone 12.0 g, Bacto[®]-yeast extract 24.0 g, glicerolo 4 ml, potassio fosfato 0.89M 100 ml) contenente ampicillina ed è stata poi portata nell'agitatore a 37°C *over night* per favorirne l'espansione. Il giorno seguente, 2 ml della coltura *over night* sono stati prelevati, inoculati in 100 ml di *Terrific Broth* e fatti crescere in incubatore a 37°C per ~2 ore e mezza in agitazione. Una volta che la sospensione batterica ha raggiunto una OD_{600nm} di 1, misurata allo spettrofotometro, è stata indotta l'espressione proteica tramite l'aggiunta di IPTG (isopropil β -D tiogalattoside) alla concentrazione finale di 1 mM. L'IPTG stimola il promotore *tac* situato a monte della sequenza codificante la proteina Oct-4A, permettendone la trascrizione. I batteri sono stati quindi incubati per ulteriori 3 ore circa a 37° C, in agitatore, per permettere l'espressione della proteina.

Successivamente i batteri sono stati lisati, sonicati e l'espressione della proteina è stata verificata mediante elettroforesi su gel di poliacrilamide (SDS-PAGE) seguita da una colorazione del gel con blu Comassie [Comassie Brilliant Blue R-250, 0.25% (p/v), acido acetico glaciale 10% (v/v) e metanolo 50% (v/v)].

3.11- Saggio di GST Pull-Down

Il saggio del GST Pull-Down è una tecnica *in vitro* che permette l'analisi di interattori molecolari mediante l'utilizzo della proteina d'interesse legata al *GST (bait)* per identificare possibili *partner* di legame (*prey*). La proteina *bait* (esca), purificata da un appropriato sistema di espressione (es. *E. coli* o cellule d'insetto infettate con baculovirus), viene immobilizzata sulla resina di affinità costituita dal glutatione, un ligando specifico per il *tag* della proteina. L'esca funge da secondo supporto d'affinità, per confermare l'interazione, precedentemente sospettata, con una certa proteina o per identificare nuovi interattori proteici della proteina esca. La proteina che fa da preda (*prey*) può essere ottenuta in diversi modi: proteine precedentemente purificate, lisati cellulari o reazioni trascrizione/traduzione *in vitro*.

In questa tesi è stato applicato il saggio del Pull-Down come saggio iniziale di screening per identificare nuovi potenziali interattori della proteina Oct-4A ancora sconosciuti.

Per l'esecuzione del saggio è stato utilizzato il kit *ProFound Pull-Down GST Protein:Protein Interaction Kit* della PIERCE.

3.11.1- Preparazione del *Bait*

Dopo aver indotto l'espressione della proteina Oct-4A con l'*IPTG* nelle cellule batteriche *KRX*, vengono trasferiti 5 ml di coltura batterica in un tubo sterile da centrifuga.

Si centrifuga a 3600 rpm per 20 minuti a 4°C e si scarta il surnatante. Successivamente il pellet di cellule viene risospeso in 1 ml di *TBS* freddo (Tris Buffered Saline: NaCl 136.9 mM, Tris base 20 mM) per un volume di coltura iniziale di 5 ml. Il pellet va risospeso bene pipettando più volte, ma non troppo energicamente per evitare la lisi delle cellule. La sospensione cellulare così ottenuta viene trasferita in una eppendorf e centrifugata a 5000 x g per 5 minuti a 4°C. Si elimina poi con cura il surnatante e si risospende, pipettando più volte, il pellet in 200 µl di *TBS* freddo per un volume di coltura iniziale di 5 ml. Come consigliato da protocollo, è stato aggiunto un cocktail di inibitori delle proteasi, *Protease inhibitor cocktail* (SIGMA). Si aggiungono 200 µl di *Lysis Buffer ProFound* fornito dal kit (per un volume di coltura iniziale di 5 ml) e si inverte ripetutamente per permettere alla soluzione di miscelarsi

correttamente.

A questo punto si incuba la soluzione in ghiaccio per 30 minuti, invertendo di tanto in tanto, e si centrifuga a 12000 x g per 5 minuti a 4°C. Si recupera il surnatante e lo si trasferisce in una nuova eppendorf che viene conservata in frigorifero a – 80°C. Il campione così preparato, è denominato “*bait* pGEX-Oct-4A”.

3.11.2- Equilibratura della resina

Per prima cosa, viene definito il numero di colonnine necessarie per il saggio (una per ciascun campione, inclusi i controlli). Viene quindi preparata la *Wash Solution* (circa 8 ml per colonnina) in cui *TBS* e *ProFound Lysis Buffer*, fornito dal kit, sono in rapporto 1:1.

Dopo aver accuratamente risospeso, mediante l'uso del vortex, la resina di glutatione, ad ogni colonnina vengono aggiunti 50 µl di resina e 400 µl di *Wash Solution*. Le colonnine vengono chiuse ad entrambe le estremità con tappini appositi e invertite alcune volte in modo da equilibrare la resina. Si rimuovono i due tappini e si pone la colonnina in un tubo di raccolta, *Collection Tube*; poi si centrifuga in microcentrifuga a 1250 x g per 1 minuto a 4°C. Dopo aver scartato la *Wash Solution* dal tubo di raccolta, viene reinserita la colonnina, chiusa all'estremità inferiore, e aggiunti altri 400 di *Wash Solution*; una volta chiusa anche l'estremità superiore, si inverte la colonnina più volte. Rimossi entrambi i tappini, si centrifuga nuovamente a 1250 x g per 1 minuto a 4°C e si ripete lo step del lavaggio per un totale di 5 lavaggi.

3.11.3- Immobilizzazione del *Bait*

Dopo aver applicato il tappo inferiore, ad ogni colonnina viene aggiunto il *bait*, cioè la proteina di fusione con il *GST*, che in questo esperimento è *pGEX-Oct-4A - GST*. Nel caso in cui il *bait* sia rappresentato da lisati cellulari, vanno aggiunti almeno 300 µl di lisato contenente la proteina legata a *GST* (ciascuna colonnina ha una capacità massima di 800µl); nel caso si disponga, invece, della proteina di fusione con il *GST* purificata in precedenza, va usato un volume sufficiente per assicurare l'aggiunta di circa 100-150 µg di proteina *bait* (da

indicazione, 25 μ l di resina-glutatione immobilizzata possono teoricamente legare 200 μ g di proteina legata a *GST*, anche se, in realtà, tale proteina può rivelare una capacità di legame più bassa dovuta a differenze conformazionali).

Dopo aver chiuso ogni colonnina con il proprio tappo, si incubano le colonnine a 4°C lasciandole in leggero movimento su una piattaforma rotante per circa 1 ora di tempo (un legame massimale può richiedere un periodo d'incubazione più lungo, che può essere determinato per ogni nuova proteina di fusione con il *GST*). Al termine dell'incubazione, si tolgono i tappini alle due estremità e ciascuna colonnina viene inserita in un nuovo tubo di raccolta. Si centrifuga in microcentrifuga a 1250 x g per 1 minuto a 4°C e l'eluato così raccolto, denominato *bait flow-through*, viene conservato a - 80°C. Si richiude la colonnina sul lato inferiore e si aggiungono 400 μ l di *Wash Solution* per lavare la resina con la proteina legata. Tappata anche l'estremità superiore, le colonnine vengono invertite alcune volte per miscelare bene e successivamente, rimossi entrambi e tappini, vengono inserite nei tubi di raccolta. Si centrifuga a 1250 x g per 1 minuto a 4°C e poi si scarta la *Wash Solution* dal tubo di raccolta; l'operazione va ripetuta per un totale di 5 lavaggi (lavaggi multipli con inversione della resina sono richiesti per eliminare effettivamente le interazioni aspecifiche).

3.11.4- Preparazione del *Prey*

Il *prey* si ottiene preparando un lisato cellulare a partire da cellule di mammifero. Le cellule utilizzate sono cellule staminali umane della polpa dentale, al terzo passaggio di coltura.

Le cellule, cresciute in adesione sulla superficie della piastra, vengono staccate utilizzando il *CTC* e, dopo essere state raccolte in un tubo sterile, vengono centrifugate a 1000 rpm per 5 minuti. Dopo avere scartato il surnatante di coltura, il pellet di cellule viene risospeso e lavato in *TBS* freddo. Si centrifuga nuovamente a 1000 rpm per 5 minuti e si elimina il surnatante. Il pellet viene poi risospeso in un volume di *TBS* freddo e *Lysis Buffer* proporzionale al peso del pellet stesso: vanno infatti aggiunti 2.5 ml di *TBS* freddo e 2.5 ml di *Lysis Buffer* per grammo di peso cellulare. Dopo aver risospeso il pellet accuratamente, si trasferisce tutto in eppendorf e si aggiunge un volume di *Lysis Buffer* uguale a quello del *TBS* usato per la risospensione [vengono addizionati anche gli inibitori delle proteasi, *Protease*

Inhibitor Cocktail, (SIGMA)].

La soluzione viene quindi incubata in ghiaccio per 30 minuti, invertendo i tubi di tanto in tanto, e successivamente centrifugata nella microcentrifuga a 12000 x g per 5 minuti a 4°C per chiarificare l'estratto cellulare.

Il surnatante con il lisato cellulare, denominato “*prey-lisato BIC3*”, viene raccolto in una nuova eppendorf e conservato a – 80°C.

3.11.5- Cattura del *Prey*

Il *prey* così preparato viene aggiunto alla colonnina contenente la proteina *bait* immobilizzata sulla resina (ogni colonnina ha una capienza massima di 800 µl di *prey*).

La colonnina viene incubata in stanza fredda a 4°C per almeno 1 ora (3-4 ore), posizionandola in leggero movimento sulla ruota (un legame massimale può richiedere un tempo di incubazione prolungato e può essere determinato per ogni nuovo campione *prey*). Conclusa l'incubazione del *prey* con la proteina *bait* legata alla resina, si rimuovono i tappini dalle colonnine, che vengono poste in tubi di raccolta, *Collection Tube*.

Si centrifuga quindi a 1250 x g per 1 minuto a 4°C e si conserva in ghiaccio l'eluato definito *prey flow-through*. Dopo aver chiuso inferiormente la colonnina, si aggiungono 400 µl di *Wash Solution*, si applica il tappo sul lato superiore e si inverte ripetutamente. Tolti entrambi i tappi, si ripone la colonnina nel tubo di raccolta, si centrifuga a 1250 x g per 1 minuto a 4°C e successivamente si scarta la *Wash Solution* che dalla colonnina è stata trasferita al tubo.

Si ripetono questi passaggi, sempre usando la *Wash solution*, per un totale di 5 lavaggi, che permettono di eliminare eventuali interattori aspecifici della proteina *bait*.

3.11.6- Eluizione *Bait-Prey*

Le proteine legate alla resina di glutatione vengono eluite in 250 µl di *Glutatione Elution Buffer*. Per preparare 1 ml di *Glutatione Elution Buffer* 100 mM, si sciolgono 31 mg di glutatione precedentemente pesati sulla bilancia analitica, in 1 ml di *Wash Solution*

utilizzando il vortex. Si chiude la base della colonnina con il tappino e si aggiungono 250 μ l di buffer di eluizione. Si applica il tappo superiore e si incuba la colonnina per 5-10 minuti lasciandola in agitazione a temperatura ambiente. Dopo aver rimosso entrambi i tappini, ogni colonnina viene posta in un nuovo tubo di raccolta e si centrifuga in microcentrifuga a 1250 x g per 1 minuto a 4°C. L'eluato così ottenuto corrisponde al complesso proteico *bait* - *prey* e viene conservato a -80°C.

Prima di essere caricati su gel di poliacrilamide, i campioni vengono dializzati utilizzando la membrana da dialisi *Slide-A-Lyzer*[®] *Mini Dialysis Unit* (Pierce) per circa 50 minuti nel buffer carbonato-bicarbonato a pH 9.4 (200 ml di Na₂CO₃ 0.1 M + 800 ml di NaHCO₃ 0.1 M). La dialisi consente di eliminare il glutatione che, in alte concentrazioni, oltre a mascherare le proteine a basso peso molecolare, determina anche la formazione di uno *smear* durante la corsa elettroforetica (Fig. 3.5).

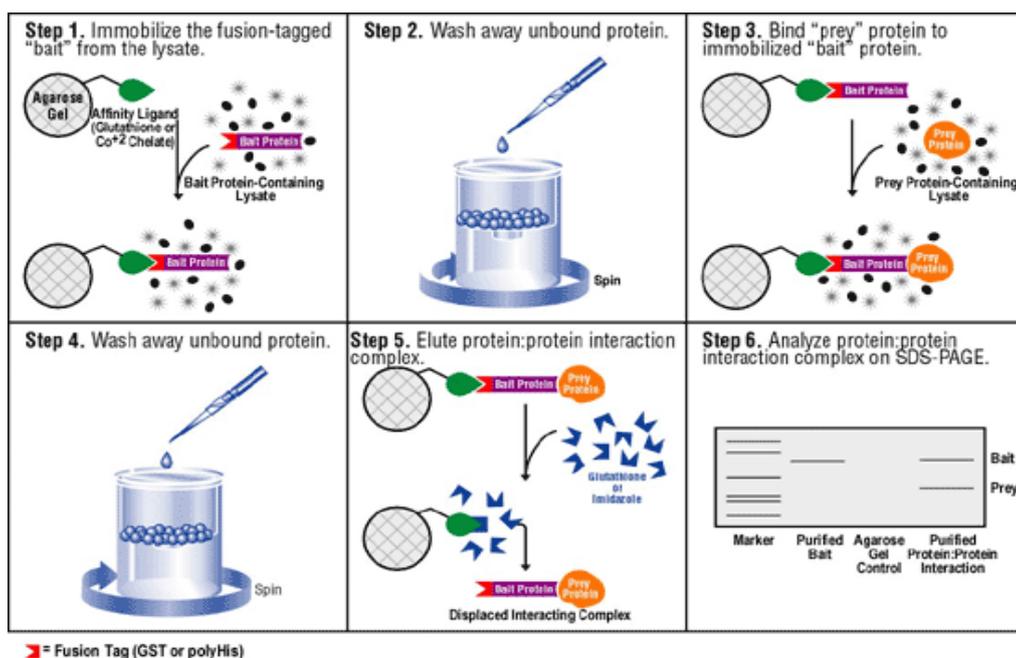


Fig. 3.5- Illustrazione schematica delle fasi di cui si compone il saggio GST Pull Down. www.piercenet.com

3.12- Frazionamento nucleo-citoplasmatico

Quando si vogliono ottenere estratti proteici nucleari o citoplasmatici è importante ricordarsi di proteggere le proteine dalla degradazione enzimatica. Infatti, in seguito a disgregazione, le cellule liberano proteasi che possono digerire velocemente le proteine degli estratti. Per minimizzare tale rischio è importante aggiungere alle soluzioni utilizzate nell'estrazione degli inibitori di proteasi ed effettuare le preparazioni a freddo. Centrifughe, eppendorf, siringhe, soluzioni e tutti gli altri materiali necessari per l'esecuzione dell'estrazione vanno raffreddati prima dell'uso.

Il principio su cui si basa questo tipo di estrazione è l'utilizzo di due tamponi a diversa forza ionica: con il primo tampone, iposmotico, le cellule si gonfiano fino a scoppiare e liberano così il loro contenuto citoplasmatico in soluzione (le membrane nucleari non vengono danneggiate perché sono più resistenti); si possono recuperare i nuclei cellulari dai quali le proteine vengono liberate mediante incubazione con una soluzione ipertonica.

La metodica richiede l'utilizzo di due soluzioni:

| | |
|-------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Soluzione T1 | HEPES 10 mM (pH 7.9), KCl 10 mM, MgCl ₂ 0.1 mM, EDTA 0.1 mM (pH 8) |
| Soluzione T2 | HEPES 20 mM (pH 7.9), NaCl 420 mM, MgCl ₂ 1.5 mM, EDTA 0.1 mM (pH 8), glicerolo 5% |
| Al momento dell'utilizzo le soluzioni vanno completate con: | cocktail di inibitori delle proteasi, " <i>Protease inhibitor cocktail</i> " 100x (SIGMA) Na-ortovanadato 2 mM } Per inibire le fosfatasi NaF 1 mM } |

Tabella 3.5- Soluzioni utilizzate nella preparazione degli estratti nucleari e citoplasmatici.

I pellet vengono quindi risospesi in 200 µl di *Soluzione T1* fredda, pipettando più volte, ma non troppo energicamente per evitare la lisi cellulare, al fine di ottenere una soluzione quanto più omogenea possibile. I pellet risospesi vengono trasferiti in eppendorf e centrifugati a 800 x g per 10 minuti a 4°C per sedimentare i nuclei. Si raccoglie il surnatante, contenente le proteine citosoliche, e lo si trasferisce in una nuova eppendorf: questo rappresenta l'estratto

proteico citoplasmatico.

I pellet, che corrispondono ai nuclei sedimentati, vengono risospesi in 200 μ l di *Soluzione T1* fredda e centrifugati a 800 x g per 10 minuti a 4°C, per sedimentarli ulteriormente. Il surnatante viene accuratamente scartato e i pellet, che rappresentano sempre i nuclei sedimentati, vengono sottoposti ad un'ulteriore centrifugazione per permettere l'eliminazione degli ultimi microlitri di *Soluzione T1*.

Successivamente si risospendono i pellet in 200 μ l di *Soluzione T2*, pipettando più volte ma non troppo energicamente per evitare la lisi nucleare finché non si ottiene una soluzione omogenea.

I campioni vengono incubati per 30 minuti in ghiaccio e vengono agitati, per qualche secondo, ogni 5 minuti mediante l'utilizzo di un vortex. Successivamente vengono centrifugati a 15.000 x g per 20 minuti a 4°C per sedimentare le membrane. Si raccoglie il surnatante, contenente le proteine nucleari, e lo si trasferisce in nuove eppendorf: questo rappresenta l'estratto proteico nucleare.

Viene successivamente eseguito il dosaggio proteico degli estratti nucleari e citoplasmatici precedentemente ottenuti, mediante il metodo di dosaggio con *BCA*.

Per l'estrazione delle proteine nucleari e citoplasmatiche è stato inoltre utilizzato anche il kit ProteoJect (Fermentas) seguendo le istruzioni fornite dalla casa produttrice che rispetto alla metodica precedentemente descritta ha consentito di ottenere dei frazionamenti più puri.

3.13- Co-immunoprecipitazione

La tecnica della Co-immunoprecipitazione permette di isolare, mediante precipitazione, complessi di proteine sfruttando l'utilizzo di un anticorpo specifico che riconosce la proteina d'interesse, ritenuta appartenente ad un ampio complesso proteico. Utilizzando un anticorpo che riconosce la proteina oggetto di studio, è possibile isolare dalla miscela proteica, che può essere rappresentata da un lisato totale, un lisato citoplasmatico o un lisato nucleare, l'intero complesso di proteine e avere, quindi, la possibilità di identificare nuovi interattori della proteina d'interesse.

Per gli esperimenti eseguiti in questo lavoro di tesi, si è fatto uso sia di lisati cellulari

totali, sia di lisati nucleari e citoplasmatici ottenuti tramite frazionamento. Le cellule Ntera2 sono state lisate utilizzando un *Lysis Buffer* contenente: Hepes pH 7.6 10mM, NaCl 250 mM, EDTA 5mM, NP40 0.1% a cui va aggiunto prima dell'uso NaF 500 mM, l'inibitore delle fosfatasi.

Prima di eseguire la lisi cellulare per ottenere gli estratti proteici totali, le cellule Ntera2 sono state incubate per 16 ore con terreno di coltura privo di siero e poi stimolate con 50 ng e 100 ng di EGF, fattore di crescita epidermico, rispettivamente per 30 minuti e 2 ore.

Per eseguire il saggio della co-immunoprecipitazione sono stati utilizzati:

- il kit “Seize[®] X protein G Immunoprecipitation Kit” (PIERCE)
- la resina coniugata all'anticorpo anti-FLAG (SIGMA)

Entrambe le metodiche si compongono di diversi:

- *Fase di Pre-clearing*: questa fase consiste nell'incubare la resina, coniugata alle Immunoglobuline appartenenti alla specie in cui è stato prodotto l'anticorpo utilizzato per la successiva immunoprecipitazione, con i lisati da testare al fine di eliminare le proteine che interagiscono in modo aspecifico. La resina è costituita da proteina A o G che legano con alta affinità la regione *Fc* della maggior parte delle classi e sottoclassi delle immunoglobuline di topo, coniglio, capra e di uomo. Gli eluati ottenuti, che rappresentano i lisati privati delle componenti che si legano in modo aspecifico al complesso resina/Ig, sono dunque pronti per subire i successivi passaggi.
- *Fase di Immunoprecipitazione*: ai lisati vengono aggiunti 5 µg di anticorpo specifico e successivamente vengono incubati *over night* in agitazione a 4°C. Successivamente viene eseguita l'incubazione con la resina in agitazione a 4°C per 4 ore.
- *Fase di Eluizione*: terminata l'incubazione si eseguono 3 lavaggi con 500-1000 µl di *Wash Buffer* e si procede poi con l'eluizione degli immuno-complessi. Nel caso della resina coniugata all'anticorpo anti-FLAG, l'eluizione è stata eseguita per competizione utilizzando una soluzione 25 mM di TRIS-HCL contenente 0,2 mg/ml di peptide FLAG.

3.14- Immunoblotting

La tecnica dell'*Immunoblotting*, chiamata anche tecnica del *Western blotting*, si utilizza generalmente dopo una elettroforesi per identificare la presenza e la posizione elettroforetica di una data proteina, contenuta in una miscela proteica, di cui si possiede l'anticorpo specifico.

3.14.1- Elettroforesi su gel di poliacrilamide con Sodio Dodecil Solfato (SDS-PAGE)

L' elettroforesi su gel di poliacrilamide con Sodio Dodecil Solfato (SDS-PAGE), permette la separazione delle proteine in base al loro peso molecolare, sfruttando la presenza di un campo elettrico e la carica negativa conferita dall'SDS. Infatti, il Sodio Dodecil Solfato si lega alle proteine in modo proporzionale al loro peso molecolare e indipendentemente dalla loro sequenza.

Il gel che si utilizza è composto da due parti:

stacking gel (gel di compattazione);

running gel (gel di corsa).

Il gel di compattazione [acrilamide:N,N'-metilen-bis-acrilamide = 30:0.8 (p/p) al 3.5% (p/v) in acrilamide, Tris-HCl 125 mM pH 6.8, SDS 0.1% (p/v)] serve a compattare i diversi campioni proteici, in modo tale che le proteine arrivino uniformi al *running gel* [poliacrilamide a varie percentuali, Tris-HCl 375 mM pH 8.8, SDS 0.1% (p/v)] attraverso cui avviene la loro separazione per peso molecolare.

Ai campioni proteici viene aggiunta la soluzione di caricamento *Laemmli 4X* [Tris-HCl 250 mM, pH 6.8, SDS 8% (p/v), glicerolo 40% (v/v), blu di bromofenolo 0.08% (v/v), DTT (DiTioTritolo) 400 mM]; successivamente i campioni vengono denaturati per 5 minuti a 95°C. La corsa è condotta in *Running Buffer* per SDS-PAGE [Tris 25 mM, glicina 192 mM, SDS 0.1% (p/v)].

Le proteine separate nel gel possono essere evidenziate usando diverse tecniche di colorazione, tra le quali la soluzione *Comassie* [*Comassie Brilliant Blue R-250*, 0.25% (p/v), acido acetico glaciale 10% (v/v) e metanolo 50% (v/v)] e la colorazione argentea (descritta in

seguito) che sono state ampiamente utilizzate in questa tesi. L'identificazione di proteine, tramite l'utilizzo di anticorpi specifici, richiede invece il trasferimento su una membrana di nitrocellulosa mediante la tecnica di “blotting” di seguito descritta.

3.14.2- Blotting

Una volta terminata la corsa delle proteine su gel di poliacrilamide, si allestisce un tipico “sandwich” costituito da:

- tre o quattro fogli di carta 3mm, precedentemente imbibiti nel tampone di trasferimento, *Towbin Buffer* (Tris base 25 mM, Glicina 192 mM, SDS 0.1%, Metanolo 20%);
- la membrana di nitrocellulosa, equilibrata prima in acqua mQ e poi trasferita per qualche minuto nel *Towbin Buffer*;
- il gel di poliacrilamide;
- altre tre o quattro fogli di carta 3mm, sempre precedentemente equilibrati nel tampone di trasferimento (Fig.3.6).

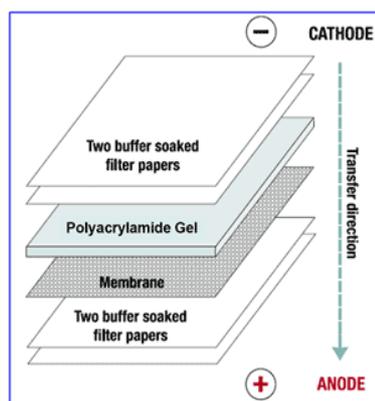


Fig. 3.6- Caratteristico allestimento a "sandwich" utilizzato nel sistema di blotting in semi-dry. www.fermentas.com

Il trasferimento su membrana di nitrocellulosa è stato eseguito in condizioni “semi-dry” mediante “*ECL Semi-dry transfert unit*” (Amersham Biosciences), che sfrutta la presenza di un campo elettrico come mostrato in figura. Le condizioni utilizzate per il trasferimento sono state le seguenti:

- 25 V;

- ~ 300 mA;
- tempo di trasferimento: 1-1.30 h.

In seguito la membrana è stata bloccata con *BSA 5%* (*Bovin Serum Albumin - SIGMA*) in TBS (Tris Buffered Saline: NaCl 136.9 mM, Tris base 20 mM) per un'ora a temperatura ambiente.

Dopo due rapidi lavaggi con *TBS-Tween 20 0.1%*, la membrana è stata incubata over night, in agitazione, a 4°C con l'anticorpo primario diluito in *TBS-Tween 20 0.1%+BSA 5%*. La membrana è stata dunque sottoposta ad una serie di lavaggi con 25 ml di *TBS-Tween 20 0.1%* per togliere l'eccesso di anticorpo non legato, successivamente è stata incubata per un'ora a temperatura ambiente e in agitazione con l'anticorpo secondario diluito sempre in *TBS-Tween 20 0.1%+BSA 5%*. L'anticorpo secondario è coniugato con l'enzima HRP che catalizza l'ossidazione del substrato Lumigen, contenuto nel reagente *ECLTM* (Amersham). La luce risultante viene rilevata mediante una lastra autoradiografica.

3.15- Immunofluorescenza

La tecnica dell'immunofluorescenza permette di visualizzare, con il supporto di un microscopio a fluorescenza, la presenza e la localizzazione di una data proteina mediante l'utilizzo di un anticorpo primario specifico e di un anticorpo secondario coniugato ad un fluorocromo. I fluorocromi sono molecole che assorbono radiazioni non visibili (ultravioletto) e che emettono nel visibile, cioè ad una lunghezza d'onda superiore rispetto a quella incidente e che, per questa loro caratteristica, vengono sfruttati nella visualizzazione delle proteine cellulari.

A tale scopo, le cellule sono state piastrate su vetrini portaoggetti e lasciate proliferare per 2-3 giorni. Dopo aver eseguito un lavaggio con TBS+Tween 1%, le cellule sono state fissate in paraformaldeide 4% per 20 minuti in agitazione. La paraformaldeide è un agente crosslinkante che, essendo liposolubile, attraversa la membrana creando una rete di ponti

intermolecolari, in genere tra gruppi amminici liberi. E' stato quindi effettuato un lavaggio in TBS+Tween 1% per 10-15 minuti, un'incubazione in TBS+glicina 1M (per tamponare la paraformaldeide) per 20 minuti ed un ulteriore lavaggio in TBS+Tween 1% per 10-15 minuti. A questo punto, per permeabilizzare la membrana e consentire così l'accesso dell'anticorpo primario nella cellula, è stata eseguita un'incubazione in TBS+Triton X-100 0.5% per 10 minuti, seguita da un lavaggio in TBS+Tween 1% per 10-15 minuti.

La fase seguente consiste nel blocco antigenico con TBS+Tween 1%+BSA 5% e nell'incubazione con anticorpo primario *over night* a 4°C mouse monoclonal anti-Oct-4A (Santa Cruz Biotechnology) diluito in TBS+Tween 1%+BSA 1.5% e utilizzato ad una concentrazione finale di 0,8 µg/ml. Seguono 3 lavaggi con TBS+Tween 1% e l'incubazione, un'ora a temperatura ambiente, con l'anticorpo secondario coniugato ad uno specifico fluorocromo [per Oct-4A anti-mouse FITC coniugated (SIGMA), diluito in TBS+Tween 1%+BSA 1.5% (1:400). Infine vengono eseguiti 3 lavaggi con TBS+Tween 1% e 2 lavaggi in acqua mQ, necessari ad eliminare i sali contenuti nel TBS. Il vetrino coprioggetti viene quindi adagiato su di un vetrino portaoggetti, su cui è stata posta una goccia di Vectashield (Vector)+DAPI [4'-6-Diamino-2-fenilindolo, un intercalante utilizzato per identificare i nuclei; (SIGMA)].

L'analisi microscopica delle cellule è stata eseguita mediante microscopio a epifluorescenza (Leica DMI6000B).

3.16- Analisi di Spettrometria di Massa

La spettrometria di massa è stata utilizzata per l'identificazione dei fosfo-peptidi presenti all'interno della proteina di fusione Oct-4A-FLAG in seguito a co-trasfezione con il plasmide MEK1^{CA}. A tal fine le cellule HeLa sono state a co-trasfettate con i due plasmidi sopra indicati e sottoposte a co-immunoprecipitazione utilizzando la resina coniugata all'anticorpo anti-FLAG (SIGMA). L'eluizione è stata eseguita come precedentemente descritto ed il campione ottenuto è stato utilizzato per l'analisi spettrometrica.

La spettrometria di massa è una metodologia che in genere viene utilizzata per l'identificazione delle proteine e prevede in un primo momento la digestione enzimatica delle proteine stesse. Le analisi con varie tecniche di spettrometria di massa consentono di ottenere i valori dei pesi molecolari dei peptidi che si formano dopo la digestione; l'identificazione delle proteine avviene andando a confrontare i valori ottenuti con quelli che derivano dalla digestione teorica di ogni proteina e che sono elencati in apposite banche dati. Per ogni analisi e sottomissione in banca dati si otterranno vari risultati, ognuno caratterizzato da un punteggio (score) che indica quanto è affidabile l'identificazione. Alcuni spettrometri sono in grado di analizzare velocemente e con sufficiente precisione i peptidi derivanti dalla digestione enzimatica della proteina (MALDI-TOF), ma non consentono di discriminare tra due peptidi aventi lo stesso peso molecolare e diversa sequenza aminoacidica. In questi casi sono necessari degli spettrometri di massa che sono in grado di lavorare in modalità MS/MS e che prevedono, in un primo momento, l'isolamento del peptide e la determinazione del suo peso molecolare, e successivamente la sua frammentazione; dalla massa molecolare degli ioni-frammento e dalla loro distribuzione è poi possibile risalire alla sequenza del peptide.

3.16.1- Dai risultati dell'analisi spettrometrica all'identificazione del polipeptide

Negli ultimi vent'anni sono stati messi a punto numerosi metodi di elaborazione spettrometrica dei dati al fine della identificazione delle proteine; lo sviluppo delle prime strategie per il sequenziamento dei polipeptidi tramite tecniche spettrometri che MS/MS, sono

state infatti sviluppate agli inizi degli anni '80. Dopo solo dieci anni, con la crescita esponenziale dei database di sequenze genomiche e proteiche, si è assistito ad un radicale cambiamento della strategia. Il motivo del cambiamento è stato essenzialmente determinato dalla possibilità di sfruttare i database di sequenze genomiche e proteiche per effettuare il lavoro di identificazione delle proteine.

Esiste un'ampia varietà di sistemi dedicati alla valutazione del grado di affidabilità con cui una proteina viene identificata. In ognuno di questi sistemi vengono definiti parametri sui quali si basa l'assegnazione di un punteggio, *score*, ai risultati di un'identificazione.

Due esempi di tali sistemi sono rappresentati dal software *PeptideSearch* (Mann M. et al., 1994) e dal software *PepFrag* (Fenyo D. et al., 1998), che sfruttano un semplice sistema di scoring basato sul numero di masse coincidenti quando lo spettro sperimentale viene messo a confronto con lo spettro teorico.

Uno dei principali programmi che sfruttava gli spettri di massa per effettuare il *Peptide Mass Fingerprint (PMF)* è stato il software *MOWSE (MOlecular Weight Search)* (Pappin D.J. Et al., 1993), da cui in seguito è stato sviluppato *Mascot* che rappresenta tutt'ora uno dei software più utilizzati per il riconoscimento di sequenze proteiche in banche dati.

In questo lavoro di tesi, il software utilizzato per l'identificazione è stato proprio *Mascot*.

Mascot è uno dei software maggiormente utilizzati per l'ottenimento del *PMF* di proteine ed è disponibile come *web server* liberamente utilizzabile, anche se con alcune limitazioni per la quantità di dati analizzabili simultaneamente.

Nella figura di seguito riportata (Fig. 3.7), è rappresentata la finestra di immissione dati del *web server Mascot*, dalla quale si può impostare il tipo di database su cui effettuare la ricerca, il tipo di enzima proteolitico utilizzato per la digestione, la tipologia di modificazioni post-traduzionali (ad esempio l'ossidazione delle metionine, la fosforilazione delle funzioni alcoliche o fenoliche di serina, treonina e tirosina) che il programma deve tenere in considerazione per effettuare la ricerca.

Fig. 3.7- Finestra di avvio del web server Mascot per il PMF di proteine.

Il primo passo della ricerca consiste nel confrontare il set di masse derivanti dal database delle sequenze presenti in banca dati con il set di masse sperimentali. Ogni valore calcolato, che cade in un range di massa vicino ad un valore sperimentale, viene considerato come un *match*. Il range di massa viene pre-definito dallo sperimentatore come tolleranza consentita.

Nel processamento dei dati sperimentali tramite Mascot, *MOWSE* assegna ad ogni *match* un peso statistico, usando parametri specifici prefissati. I fattori che determinano il peso statistico sono calcolati in una matrice che viene costruita durante il processamento del database di sequenze proteiche. La matrice dei fattori di frequenza, denominata F , viene costruita mettendo in ogni riga un intervallo di 100 Da nella massa del peptide e in ogni colonna un intervallo di 10 kDa nella massa intatta della proteina. Quando le sequenze presenti nel database vengono processate, i generici elementi f_{ij} della matrice sono incrementati, in modo tale da accumulare dati sulla forma della distribuzione delle masse dei peptidi come funzione della massa della proteina. Gli elementi di F sono in seguito normalizzati dividendo i singoli elementi della colonna da 10 kDa per il valore più grande in quella stessa colonna per dare il fattore della matrice MOWSE denominato m :

$$m_{ij} = \frac{f_{ij}}{(f_{ij})_{\max \text{ in column } j}}$$

Dopo il confronto tra masse sperimentali e masse teoriche il punteggio per ogni proteina è calcolato secondo la formula:

$$Score = \frac{50,000}{M_{Prot} \times \prod_n m_{i,j}}$$

in cui il denominatore è costituito dal prodotto tra M_{prot} (il peso molecolare della sequenza) e il prodotto tra i fattori MOWSE (per ogni match tra i dati sperimentali e le masse dei peptidi calcolati in base alla sequenza della proteina presente nel database).

A tutti i match positivi ottenuti utilizzando i dati spettrali, siano essi riferiti alle masse dei peptidi o a quelle dei frammenti, sono assegnati punteggi calcolati come $-10 \cdot \log(P)$ dove P è la probabilità assoluta che l'identificazione sia corretta.

Per fare un esempio, ad una probabilità di 10-20 corrisponde uno score di 200. Un tipico valore di soglia normalmente utilizzato è quello per cui un evento risulta significativo se la frequenza di eventi casuali è minore del 5%.

In figura 3.8 è riportato un tipico esempio di grafico proveniente dalla pagina dei risultati di *PMF* tramite Mascot: è un istogramma relativo alla distribuzione dei punteggi di identificazione.

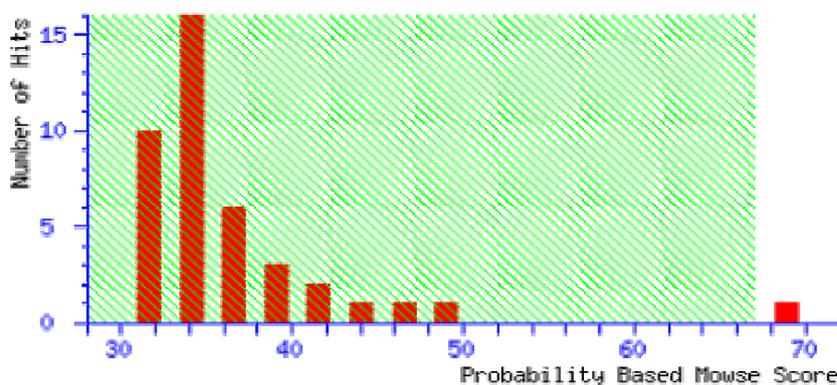


Fig. 3.8- Grafico di distribuzione dei punteggi di identificazione utilizzando Mascot

La zona evidenziata con il tratteggio verde corrisponde alla zona di non significatività e

contiene tutti i risultati considerati non significativi nel modello probabilistico utilizzato. In riferimento sempre alla figura 3.8, la proteina con score 108 risulta invece chiaramente distinguibile dalla zona di non significatività (il cui limite corrisponde ad uno score di 67) e rappresenta quindi un match significativo o comunque quello con il più alto grado di significatività o omologia.

Occorre aggiungere che è importante distinguere tra un match significativo e un match corretto. Infatti, se si ripete la ricerca precedente aumentando il valore di tolleranza di massa da $\pm 0,1$ Da a ± 1.0 Da si ottiene quanto rappresentato nella figura 3.9. In questo caso, il potere discriminante della ricerca viene largamente ridotto e il miglior match, pur essendo ancora corretto, risulta ora poco significativo.

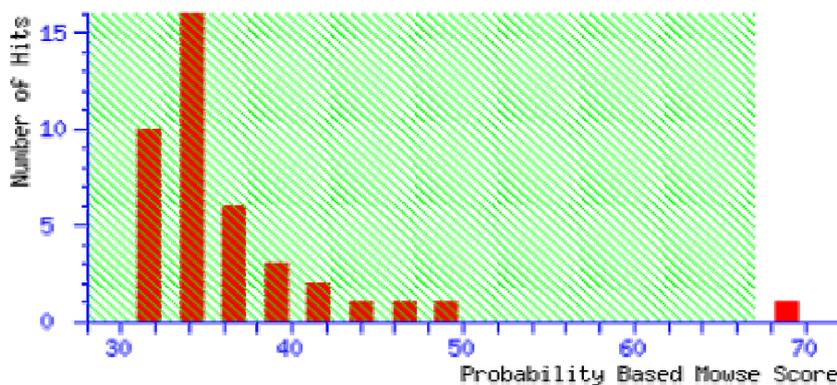
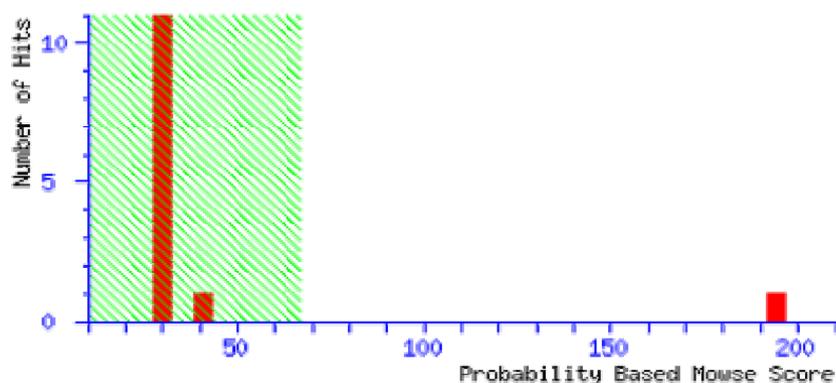


Fig. 3.9- Grafico di distribuzione dei punteggi di identificazione utilizzando Mascot, aumentando la tolleranza di massa.

Un esempio dei risultati ottenibili con Mascot è riportato in figura 3.10.



1. [PML HUMAN](#) Mass: 97455 Score: **194** Expect: 1e-14 Queries matched: 15
 Probable transcription factor PML (Tripartite motif-containing protein 19) (RING finger protein 71)
- [MURC IDILO](#) Mass: 52994 Score: 51 Expect: 2 Queries matched: 5
 UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-acetylmuramoyl-L-alanine synthetase) - I
- [DP01 RICHE](#) Mass: 104386 Score: 50 Expect: 2.8 Queries matched: 6
 DNA polymerase I (EC 2.7.7.7) (POL I) - Rickettsia helvetica
- [THIO PONPY](#) Mass: 11877 Score: 41 Expect: 20 Queries matched: 3
 Thioredoxin (Trx) - Pongo pygmaeus (Orangutan)
- [RBL2 RHOS4](#) Mass: 50569 Score: 40 Expect: 28 Queries matched: 4
 Ribulose biphosphate carboxylase (EC 4.1.1.39) (RuBisCO) - Rhodobacter sphaeroides (strain ATCC 17
- [RBL2 RHOSH](#) Mass: 50487 Score: 40 Expect: 28 Queries matched: 4
 Ribulose biphosphate carboxylase (EC 4.1.1.39) (RuBisCO) - Rhodobacter sphaeroides (Rhodopseudomon
- [GPAL YEAST](#) Mass: 54042 Score: 40 Expect: 29 Queries matched: 4
 Guanine nucleotide-binding protein alpha-1 subunit (GPI-alpha) - Saccharomyces cerevisiae (Baker's

Fig. 3.10- Esempio dei risultati di una ricerca di PMF con Mascot. Il grafico della distribuzione dei punteggi sperimentali indica come match valido solamente una proteina, con il punteggio (194) evidenziato in rosso nella lista

3.16.2- Digestione del campione

Il campione ottenuto in seguito all'immunoprecipitazione con resina anti-FLAG (SIGMA) delle cellule HeLa co-trasfettate con i plasmidi FLAG-Oct-4A e MEK1^{CA}, è stato concentrato mediante speed vac, ridotto con DTT e alchilato con iodoacetamide. Una piccola quantità è stata digerita con tripsina per l'identificazione, il restante è stato digerito con l'enzima Glu-C per eseguire la fosfo-analisi. Di seguito viene riportata la sequenza dei peptidi teorici derivanti dalla digestione con Glu-C e i siti ipotetici di fosforilazione da parte di Erk.

| Num | From-To | MH+ | Sequence |
|-----|---------|---------|---------------------------------------------------------------------|
| 19 | 239-246 | 957.52 | NRVRGNLE |
| 11 | 136-145 | 1159.63 | SQDIKALQKE |
| 2 | 57- 68 | 1411.67 | VWGIPPCPPPYE |
| 18 | 225-238 | 1686.01 | TLVQARKRKRTSIE |
| 25 | 344-360 | 1741.86 | AFPPVSVTTLGSPMHSN |
| 7 | 105-125 | 2115.99 | SNSDGA <u>SP</u> EPCTVTPGAVKLE |
| 3 | 69- 91 | 2353.08 | FCGGMAYCGPQVGVLVPOGGLE |
| 14 | 189-210 | 2731.48 | ALQLSFKNMCKLRPLLQKWVEE |
| 20 | 247-270 | 2776.48 | NLFLQCPKPTLQQISHIAQQLGLE |
| 21 | 271-296 | 3255.67 | KDVVRVWFNRRQKGRSSDYAORE |
| 23 | 300-341 | 4189.01 | AAG <u>SP</u> FSGGPVSFPLAPGPHFGTPGYG <u>SP</u> HFTALYSSVFPPE |
| 13 | 148-188 | 4675.58 | QFAKLLKQKRITLGYTQADVGLTLGVLFQKVFSSQTTICRFE |
| 1 | 1- 56 | 5321.48 | MAGHLASDFAF <u>SP</u> PPGGGGDGGPGGPEPGWVDPRTWLSFQGGPPGGPGIGPGVGGPSE |

3.16.3- Identificazione con “MS peptide mapping” e analisi della sequenza

Il campione digerito con tripsina è stato purificato utilizzando una colonna C18 (MILLIPORE) e l'eluato è stato analizzato con Bruker Autoflex III MALDI TOF/TOF.

Dopo aver eseguito l'analisi MALDI MS/MS i peptidi sono stati sottoposti al sequenziamento parziale.

Gli spettri ottenuti dall'analisi MS e MS/MS sono stati combinati e utilizzati per la ricerca nel database NCBI utilizzando il software Mascot precedentemente descritto.

3.16.4- Arricchimento dei fosfo-peptidi tramite purificazione con biossido di titanio (TiO₂)

Il campione digerito con Glu-C è stato concentrato tramite Speed Vac e quindi purificato su di una micro-colonna di TiO₂. Successivamente l'eluato è stato purificato con una micro-colonna POROS R3 per l'analisi MS.

La miscela di peptidi è stata quindi analizzata per determinarne la massa ed i peptidi risultanti appartenenti a Oct-4A sono stati parzialmente sequenziali mediante analisi MS/MS.

Gli spettri ottenuti dall'analisi MS e MS/MS sono stati combinati per la ricerca nel database NCBI utilizzando il software Mascot.

4- RISULTATI

4.1- Verifica dell'espressione del fattore di trascrizione Oct-4A nelle cellule NTera2

4.1.1- Immunofluorescenza

Il saggio dell'immunofluorescenza è stato eseguito per verificare l'espressione dell'isoforma A del fattore di trascrizione Oct-4 nelle cellule NTera2. Infatti l'anticorpo utilizzato riconosce gli amminoacidi da 1 a 134, che corrispondono all'estremità N-terminale della proteina, ovvero il dominio proteico che distingue l'isoforma A di Oct-4 da quella B.

Le cellule sono state fissate in paraformaldeide e permeabilizzate utilizzando il Triton X-100 come descritto nel capitolo “Materiali e metodi”.

Come si può notare dall'immagine di seguito riportata (Fig. 4.1), le cellule NTera2 presentano una positività per Oct-4A soprattutto a livello nucleare.

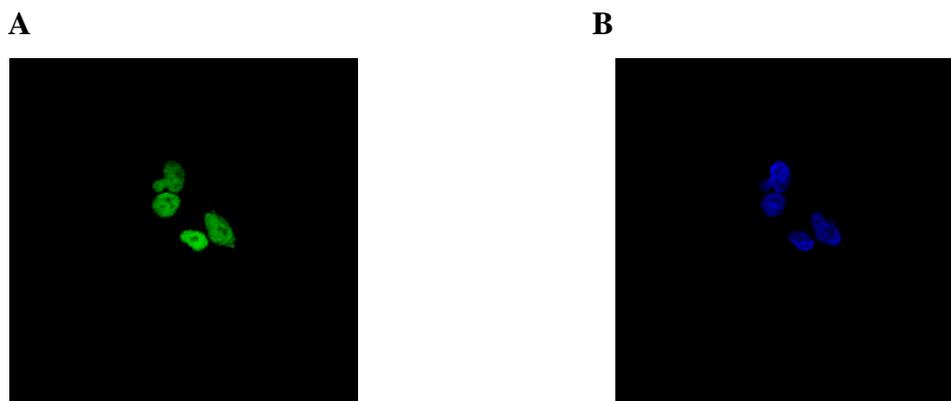


Fig. 4.1- A) Immunofluorescenza per Oct-4A eseguita sulle cellule NTera2 B) DAPI (Ingrandimento 40X).

4.1.2- Immunoblotting

L'espressione dell'isoforma A di Oct-4 è stata indagata anche tramite Immunoblotting, che viene eseguito al termine di una corsa elettroforetica e consente di individuare le proteine d'interesse valutando anche il peso molecolare (vedi capitolo "Materiali e metodi").

L'anticorpo specifico utilizzato per Oct-4A è lo stesso utilizzato nell'analisi di immunofluorescenza. Nell'immagine di seguito mostrata (Fig. 4.2) si può notare la presenza di una banda specifica all'altezza attesa e corrispondente ad un peso molecolare approssimativo di 50 kDa, come riportato in letteratura (Liedtke et al., 2008).

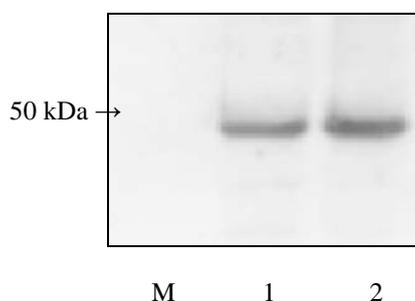


Fig. 4.2- Immunoblotting eseguito su 2 lisati totali delle cellule NTera2 (lane 1 e 2) utilizzando l'anticorpo specifico per l'isoforma A di Oct-4. M: marker di peso molecolare

4.2- Analisi della sequenza aminoacidica di Oct-4A

Nota la presenza di numerosi siti di fosforilazione nella sequenza aminoacidica di Oct-4A (Brehm A. et al., 1996), si è pensato che un buon spunto per meglio comprenderne i meccanismi di regolazione potesse essere indagare le possibili sequenze di fosforilazione.

Per identificare i siti di fosforilazione presenti nella sequenza aminoacidica di Oct-4A, è stata effettuata una ricerca sul database *Human Protein References Database*, in particolare in *PhosphoMotif Finder*, un portale esclusivamente dedicato alla ricerca di sequenze riconosciute da protein-chinasi e fosfatasi. Inserendo nel database la sequenza aminoacidica

corrispondente alla proteina d'interesse, si è ottenuto come risultato che Oct-4A possiede numerosi potenziali siti di fosforilazione da parte della protein-kinasi *Erk*. *Erk* è una *MAP-chinasi* e risulta coinvolta nella via di segnalazione mediata da *Ras*, una proteina monomericamente *GTPasica* che agisce da trasduttore del segnale in una molteplicità di vie ma in particolare in una via principale che stimola le cellule alla proliferazione e al differenziamento.

Di seguito sono evidenziati, nella sequenza aminoacidica, e riportati nella tabella sottostante i siti di fosforilazione di Oct-4A per *Erk*, identificati in seguito alla ricerca nel database *PhosphoMotif Finder*:

```

1  MAGHLASDFA FSEPPGGGGD  GPGGPEPGWV  DPRTWLSFOG  PPGGPGIGPG  VGGPSEVWGI
61  PPCPPPYEFC  GGMAYCGPQV  GVGLVPQGGL  ETSQPEGEAG  VGVESNSVGA  SPEPCTVTPG
121  AVKLEKEKLE  QNPEESQDIK  ALQKELEQFA  KLLKQKRITL  GYTQADVGLT  LGVLFQKQVFS
181  QTTICRFEAL  QLSFKNMCKL  RPLLQKWVEE  ADNNENLQEI  CKAETLVQAR  KRKRTSIENR
241  VRGNLENLFL  QCPKPTLQOI  SHIAQQQLGLE  KDVVVRVWFCN  RRQKGRSSS  DYAQREDFEA
301  AGSPFSGGPV  SFPLAPGPHF  GTPGYGSPHF  TALYSSVPFP  EGEAFPPVSV  TTLGSPMHSN

```

Fig. 4.3- Sequenza aminoacidica della proteina Oct-4A. In rosso sono evidenziati i siti di fosforilazione per *Erk* corrispondenti al motivo consenso XX(S/T)P, identificati tramite la ricerca nel database *PhosphoMotif Finder*.

| Posizione nella proteina Oct-4A | Sequenza riconosciuta della proteina Oct-4A | Motivo generale riconosciuto da ERK | Caratteristiche del motivo |
|---------------------------------|---------------------------------------------|-------------------------------------|--------------------------------|
| 10-13 | AFSP | XXpSP | Motivo substrato di ERK1, ERK2 |
| 11-13 | FSP | X[pS/pT]P | Motivo substrato di ERK1, ERK2 |
| 12-13 | SP | pSP | Substrato di fosforilazione |
| 109-112 | GASP | XXpSP | Motivo substrato di ERK1, ERK2 |
| 110-112 | ASP | X[pS/pT]P | Motivo substrato di ERK1, ERK2 |
| 111-112 | SP | pSP | Substrato di fosforilazione |
| 301-304 | AGSP | XXpSP | Motivo substrato di ERK1, ERK2 |
| 302-304 | GSP | X[pS/pT]P | Motivo substrato di ERK1, ERK2 |
| 303-304 | SP | pSP | Substrato di fosforilazione |
| 325-328 | YGSP | XXpSP | Motivo substrato di ERK1, ERK2 |
| 326-328 | GSP | X[pS/pT]P | Motivo substrato di ERK1, ERK2 |
| 327-328 | SP | pSP | Substrato di fosforilazione |
| 353-356 | LGSP | XXpSP | Motivo substrato di ERK1, ERK2 |
| 354-356 | GSP | X[pS/pT]P | Motivo substrato di ERK1, ERK2 |
| 355-356 | SP | pSP | Substrato di fosforilazione |

Tabella 4.1- In tabella sono riportate le posizioni e le sequenze aminoacidiche di Oct-4A riconosciute dalla protein-chinasi Erk ed i relativi substrati di fosforilazione.

4.3- Erk interagisce *in vivo* e *in vitro* con Oct-4A

Vista la presenza di numerosi potenziali siti di fosforilazione da parte di questa chinasi, abbiamo voluto valutare la possibilità di un'interazione fisica tra queste due proteine, sia *in vitro* con il saggio del GST Pull-Down, che *in vivo* mediante la Co-Immunoprecipitazione.

4.3.1- Saggio del GST Pull-Down: amplificazione del cDNA che codifica per la proteina Oct-4A

Verificata l'espressione del fattore di trascrizione Oct-4A nelle cellule Ntera-2, si è potuto procedere con l'estrazione dell' RNA, come descritto nel capitolo "Materiali e metodi", e successivamente con la retrotrascrizione dell'RNA messaggero codificante per la proteina Oct-4A in cDNA. Il cDNA di Oct-4A è stato amplificato utilizzando dei *primers* specifici provvisti dei siti di riconoscimento per le endonucleasi di restrizione *EcoRI* e *BamHI*. Il *primer forward* contiene la tripletta *ATG*, che rappresenta il sito d'inizio della traduzione, e in combinazione al *primer reverse* permette l'amplificazione completa dell'intera porzione codificante di Oct-4A.

Primer Forward (Oct-4A F)



Primer Reverse (Oct-4A R)



```

1   CCTTCGCAAG   CCCTCATTTC   ACCAGGCCCC   CGGCTTGGGG   CGCCTTCCTT   CCCCATGGCG
61  GGACACCTGG   CTTCGGATTT   CGCCTTCTCG   CCCCCCTCCAG   GTGGTGGAGG   TGATGGGCCA
121 GGGGGGCCGG   AGCCGGGCTG   GGTTGATCCT   CGGACCTGGC   TAAGCTTCCA   AGGCCCTCCT
181 GGAGGGCCAG   GAATCGGGCC   GGGGGTTGGG   CCAGGCTCTG   AGGTGTGGGG   GGTGGAGTG
241 TGCCCCCGC   CGTATGAGTT   CTGTGGGGGG   ATGGCCTACT   GTGGGCCCCA   GGTGGAGTG
301 GGCTAGTGC   CCAAGGCGG   CTGAGAGACC   TCTCAGCCTG   AGGGCGAAGC   AGGAGTCGGG
361 GTGGAGAGCA   ACTCCGATGG   GGCCTCCCCG   GAGCCCTGCA   CCGTCACCCC   TGGTGCCTGG
421 AAGCTGGAGA   AGGAGAAGCT   GGAGCAAAAC   CCGAGGAGT   CCCAGGACAT   CAAAGCTCTG
481 CAGAAAGAAC   CAGAAAGAAC   TCGAGCAATT   TGCCAAGCTC   CTGAAGCAGA   CCTGGGATAT
541 ACACAGGCCG   ATGTGGGGCT   CACCCTGGGG   GTTCTATTTG   GGAAGGTATT   CAGCCAAACG
601 ACCATCTGCC   GCTTTGAGGC   TCTGCAGCTT   AGCTTCAAGA   ACATGTGTAA   GCTGCGGCC
661 TTGCTGCAGA   AGTGGGTGGA   GGAAGCTGAC   AACAAATGAAA   ATCTTCAGGA   GATATGCAAA
721 GCAGAAACCC   TCGTGCAGGC   CCGAAAGAGA   AAGCGAACCA   GTATCGAGAA   CCGAGTGAGA
781 GGCAACCTGG   AGAATTTGTT   CCTGCAGTGC   CCGAAACCCA   CACTGCAGCA   GATCAGCCAC
841 ATCGCCAGC   AGCTTGGGCT   CGAGAAGGAT   GTGGTCCGAG   TGTGGTTCTG   TAACCGGCGC
901 CAGAAGCGCA   AGCGATCAAG   CAGCGACTAT   GCACAACGAG   AGGATTTTGA   GGCTGCTGGG
961 TCTCCTTTCT   CAGGGGGACC   AGTGTCTTTT   CCTCTGGCCC   CAGGGCCCA   TTTTGGTACC
1021 CCAGGCTATG   GGAGCCCTCA   CTFCACTGCA   CTGTACTCCT   CGGTCCCTTT   CCCTGAGGGG
1081 GAAGCCTTTC   CCCCTGTCTC   CGTCACCACT   CTGGGCTCTC   CCATGCATTC   AAACTGAGGT
1141 TTGCTGCCCTT   CTAGGAATGG   GGGACAGGGG   GAGGGGAGGA   GCTAGGAAA   GAAAACCTGG
1201 AGTTTGTGCC   AGGGTTTTTG   GGATTAAGTT   CTTCAATCAC   TAAGGAAGGA   ATTGGGAACA
1261 CAAAGGTGG   GGCAGGGGA   GTTTGGGCA   ACTGGTTGGA   GGAAGGTGA   AGTTCAATGA
1321 TGCTCTTGAT   TTTAATCCCA   CATCATGTAT   CACTTTTTTC   TTAATAAAG   AAGCCTGGGA
1381 CACAGTAGAT   AGACACACTT   AAAAAAAAAA   A

```

Fig. 4.4- Sequenza nucleotidica di Oct-4A. Sono evidenziate in rosso le sequenze a cui si appaiano i primers utilizzati per l'amplificazione del cDNA successivamente clonato in pGEX4T1.

Per la reazione di amplificazione del cDNA di Oct-4A, è stata utilizzata la *Pfu* polimerasi che sintetizza il filamento complementare al DNA stampo a partire dai due primers *Oct-4A F* e *Oct-4A R*. La *Pfu* polimerasi possiede un'ottima termostabilità e grazie alla sua attività esonucleasica 3' → 5' garantisce una maggiore fedeltà di trascrizione, necessaria ai fini del clonaggio.

L'immagine di seguito riportata (Fig. 4.5) rappresenta l'analisi del prodotto ottenuto dall'amplificazione mediante corsa elettroforetica su gel di agarosio, come descritto nel capitolo "Materiali e metodi"; i campioni sono stati sottoposti a due PCR, allestite mantenendo le medesime condizioni ma differenti nel numero di cicli (30 e 40 cicli). In entrambi i campioni analizzati, si può notare una banda intensa collocata al di sopra dell'ultimo *marker* (1030 bp), di altezza circa pari a 1100 bp. Dato che il numero di bp dell'amplificato che si dovrebbe ottenere utilizzando la coppia di primers *Oct-4A F* e *Oct-4A R* è di ~ 1093 bp, l'esito ottenuto conferma il successo dell'amplificazione.

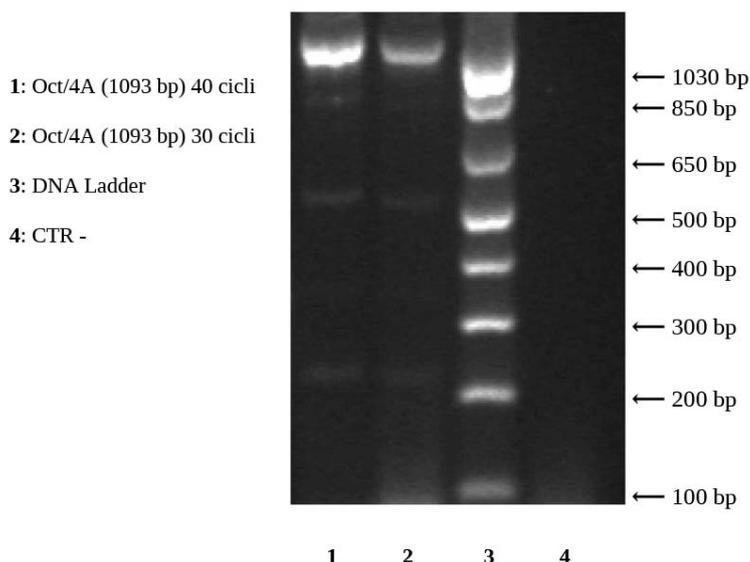


Fig. 4.5- PCR del cDNA corrispondente alla regione codificante di Oct-4A . La lane 1 corrisponde a 40 cicli di amplificazione , la lane 2 a 30 cicli. Nel controllo negativo (lane 4) il campione è stato sostituito con H₂O.

4.3.2- Saggio del GST Pull-Down: analisi di restrizione dei cloni trasformati con il vettore “pGEX4T1-Oct-4A”

La sequenza di cDNA corrispondente alla porzione codificante per la proteina Oct-4A è stata clonata all'interno del sito *polylinker* presente nel vettore di espressione pGEX4T1; il vettore è stato precedentemente digerito con le endonucleasi EcoRI e BamHI, per cui possiede i siti di riconoscimento all'interno del sito multiplo di clonaggio, e defosforilato. Al fine di ottenere estremità coesive, anche l'inserito è stato digerito con gli stessi enzimi di restrizione. Successivamente è stata eseguita la reazione di ligazione vettore-inserito e la trasformazione delle cellule batteriche BL21.

In seguito, dai cloni così ottenuti, è stato possibile eseguire l'estrazione del DNA plasmidico, tramite la metodica delle Miniprep (vedi “Materiali e metodi”), a cui è seguita l'analisi di restrizione su gel di agarosio.

Nella figura 4.6 viene riportato il plasmide digerito, contenente un inserto di altezza pari a quella attesa: si può notare una banda all'altezza di circa 1100 bp, che corrisponde all'inserto codificante per Oct-4A (lunghezza teorica di 1093 bp) ed una banda più alta, circa di 5000 bp, corrispondente al plasmide pGEX4T1 linearizzato (lunghezza teorica di 4969 bp).

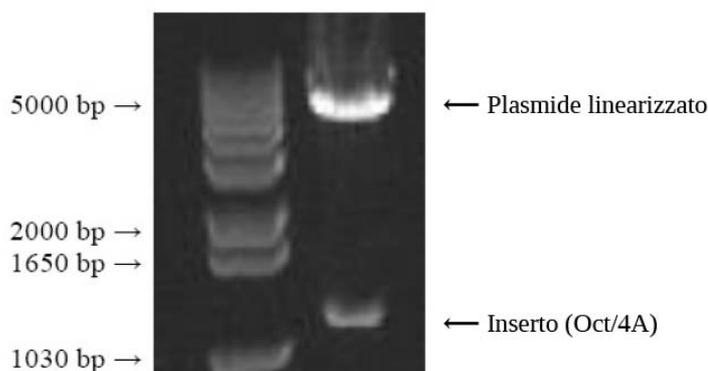


Fig. 4.6- Analisi di restrizione su gel di agarosio: è possibile notare una banda corrispondente al plasmide linearizzato e una banda corrispondente all'inserito Oct-4A

4.3.3- Saggio del GST Pull-Down: sequenziamento del plasmide “pGEX4T1-Oct4A”

Il sequenziamento è stato realizzato sul DNA plasmidico estratto tramite la metodica delle Midiprep ed è stato eseguito utilizzando i primers riportati nel capitolo “Materiali e metodi”. La tecnica delle Midiprep, oltre a permettere di ottenere una quantità maggiore di DNA plasmidico rispetto alla metodica delle miniprep, garantisce un maggior grado di purezza del campione, necessaria per il sequenziamento. L'analisi di sequenziamento è stata realizzata presso l'azienda tedesca *MWG*.

Per alcuni dei plasmidi ricombinanti ottenuti, l'inserito corrispondente alla sequenza di Oct-4A clonata è stato completamente sequenziato per verificarne la bontà in termini di sequenza nucleotidica e per accertarne il corretto orientamento all'interno della struttura del vettore pGEX4T1.

Viene riportato di seguito (Fig. 4.9), a titolo esemplificativo, uno dei cromatogrammi relativo al sequenziamento dell'inserto codificante per Oct-4A. Questo cromatogramma risulta dal sequenziamento effettuato mediante l'utilizzo del *primer forward*:

5'- GACAACAATGAAAATCCTTCAGGAGA - 3'

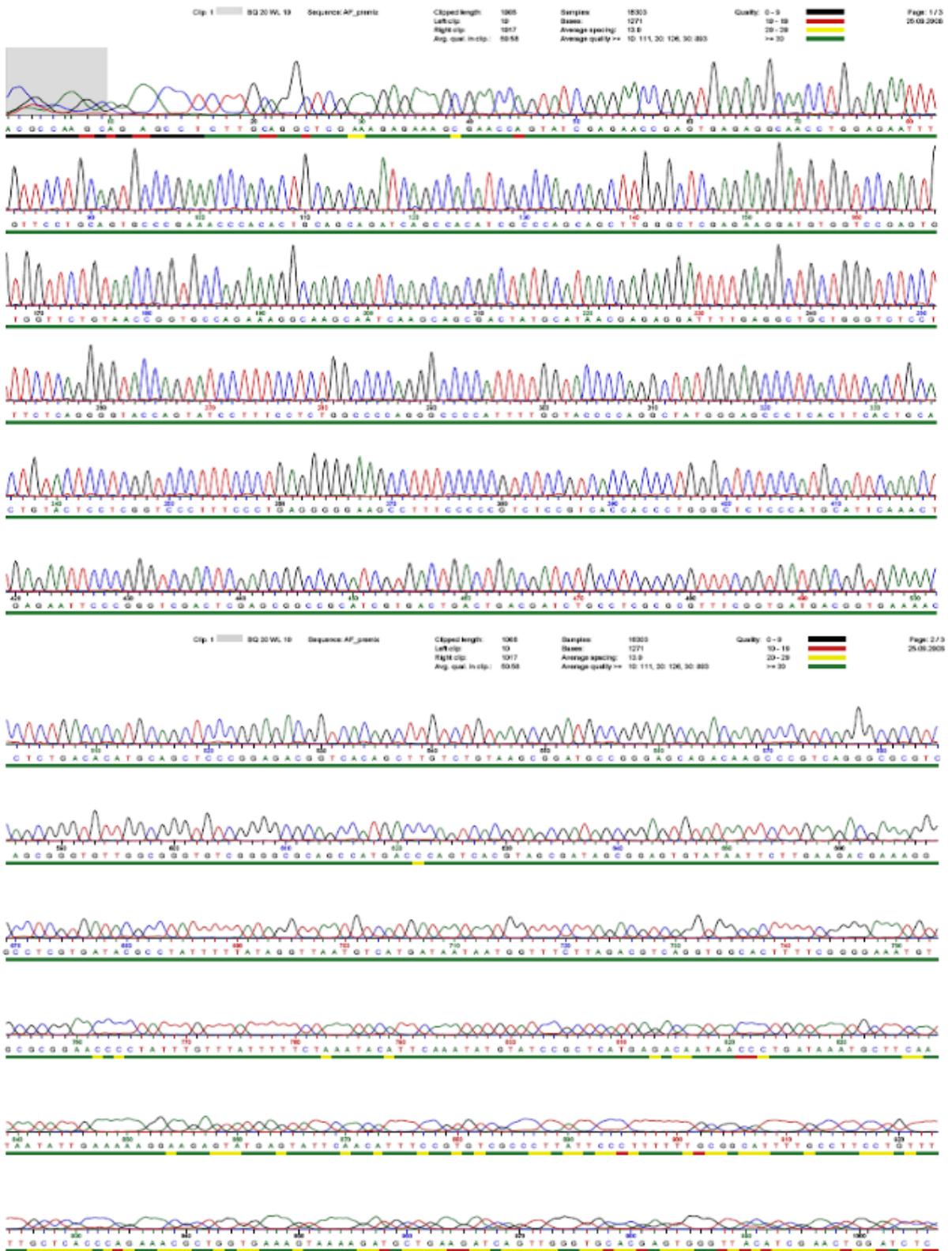


Fig. 4.7- Cromatogramma relativo ad un sequenziamento eseguito come verifica della sequenza e del corretto orientamento dell'inserito all'interno del plasmide pGEX4T1.

4.3.4- Saggio del GST Pull-Down: induzione dell'espressione di Oct-4A

Il plasmide pGEX4T1 contenente l'inserito di cDNA che codifica per Oct-4A, è stato utilizzato per trasformare le cellule batteriche KRX. Successivamente si è potuto procedere con le prove di induzione dell'espressione proteica.

Il vettore pGEX4T1 si caratterizza per la presenza di un *open reading frame* codificante per la glutatione-S-transferasi (*GST*) posizionata a monte del sito multiplo di clonaggio; questo rappresenta dunque un sistema che permette l'espressione inducibile della proteina d'interesse in fusione con GST.

E' stata quindi indotta l'espressione della proteina di fusione GST-Oct-4A, aggiungendo alla coltura batterica di KRX trasformato IPTG alla concentrazione finale di 1mM.

Successivamente i batteri sono stati lisati, sonicati e si è potuto analizzare l'espressione della proteina di fusione GST-Oct-4A mediante elettroforesi su gel di poliacrilamide (*SDS-PAGE*). La figura di seguito riportata (Fig. 4.8) è un esempio di analisi SDS-PAGE: si può notare una banda netta, corrispondente alla proteina GST-Oct-4A nel campione indotto, che risulta molto più intensa rispetto al campione non indotto; questo dimostra che le condizioni di induzione applicate hanno permesso di ottenere un buon livello di espressione della proteina.

Oct-4A, che presenta un peso molecolare apparente di ~ 50 kDa, in fusione con GST (PM ~ 25 kDa), raggiunge un valore di ~ 75 kDa.

Per controllare la correttezza del sistema allestito, è stato utilizzato come controllo il vettore pGEX4T1 vuoto, privo cioè dell'inserito.

La proteina di fusione GST-Oct-4A è stata purificata a partire da 5 ml di coltura batterica indotta, secondo il protocollo descritto nel capitolo "Materiali e metodi".

Per prevenire la degradazione delle proteine del lisato batterico, è stato addizionato un cocktail di inibitori delle proteasi e tutte le fasi del processo di purificazione sono state eseguite con i campioni sempre tenuti in ghiaccio o a 4°C.

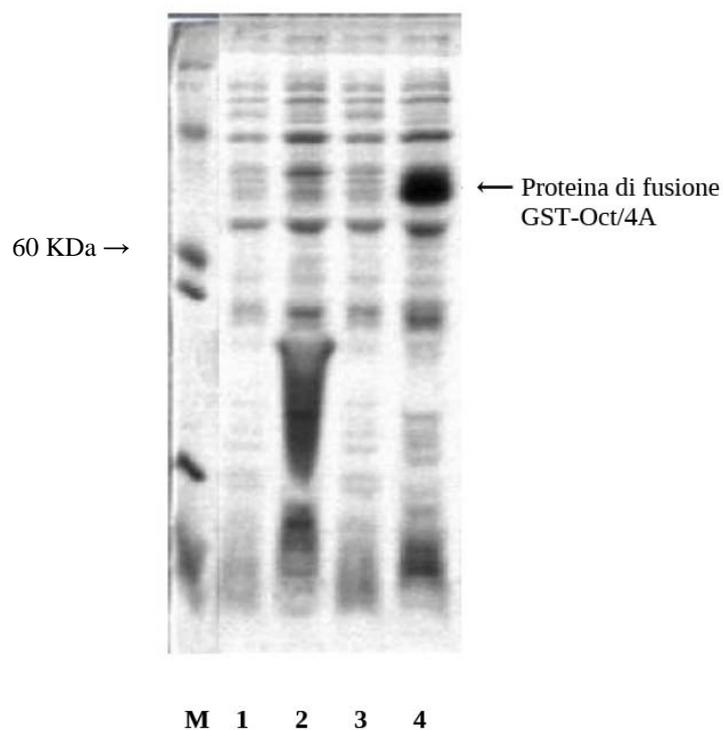


Fig. 4.8- Analisi dell'induzione proteica in lisati batterici tramite gel di poliacrilamide 12% (SDS-PAGE) colorato con Coomassie. I batteri sono stati precedentemente trasformati con i plasmidi pGEX-4T1 (lane 2) e pGEX-4T1/Oct-4A (lane 4). Sono stati inseriti due controlli (lane 1 e 3) rappresentati da batteri non sottoposti all'induzione. M: marker di peso molecolare.

4.3.5- GST Pull-Down

Nonostante il saggio di GST Pull-Down sia stato messo appunto, come strumento di analisi di interattori molecolari, ormai da diversi anni (Current Protocols, Wiley 1996), è un sistema che vanta un'eccellente affidabilità, proprio in virtù del suo enorme utilizzo e del suo progressivo perfezionamento.

Come per ogni altra tecnica, anche nel saggio di GST Pull-Down non mancano peculiari limiti; un'importante limitazione è rappresentata dalla proteina di fusione *GST-bait* che viene prodotta e purificata da cellule batteriche. La produzione della proteina di fusione *GST-bait* in un sistema procariotico, fa sì che essa non sia sottoposta alle dovute modificazioni post-traduzionali, soprattutto quelle a livello della struttura terziaria e quaternaria; da ciò ne consegue la possibile perdita di alcuni interattori dell'esca.

Di seguito vengono riportati l'illustrazione schematica della proteina di fusione ed i risultati ottenuti (Fig. 4.9a, 4.9b)

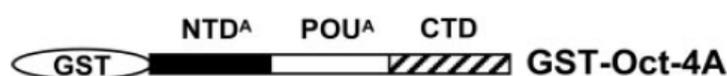


Fig. 4.9a- Illustrazione schematica della proteina di fusione GST-Oct-4A. Lee J. et al., 2006.

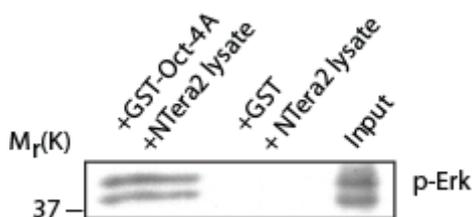


Fig. 4.9b - Saggio del GST pull-down condotto utilizzando la proteina di fusione GST-Oct-4A o la proteina di controllo GST, immobilizzate su beads di agarosio e incubate con i lisati di cellule Ntera2. L'interazione in vitro tra Erk e Oct-4A è stata testata con l'immunoblotting utilizzando un

4.3.6- Co-immunoprecipitazione

Per verificare se l'interazione tra Oct-4A ed Erk, dimostrata *in vitro* con il saggio del GST pull-down, avviene anche *in vivo* è stata eseguita la co-immunoprecipitazione utilizzando i lisati delle cellule NTera2. La co-immunoprecipitazione è una tecnica largamente utilizzata per dimostrare l'interazione tra due o più proteine e sfrutta l'affinità esistente tra una data proteina e il suo anticorpo specifico che viene coniugato ad una fase solida, tipicamente proteina A o proteina G. Come si può vedere nella figura 4.10 l'anticorpo anti-Oct-4A è in grado di immunoprecipitare Erk, così come l'anticorpo anti-Erk immunoprecipita Oct-4A confermando l'ipotesi dell'interazione endogena.

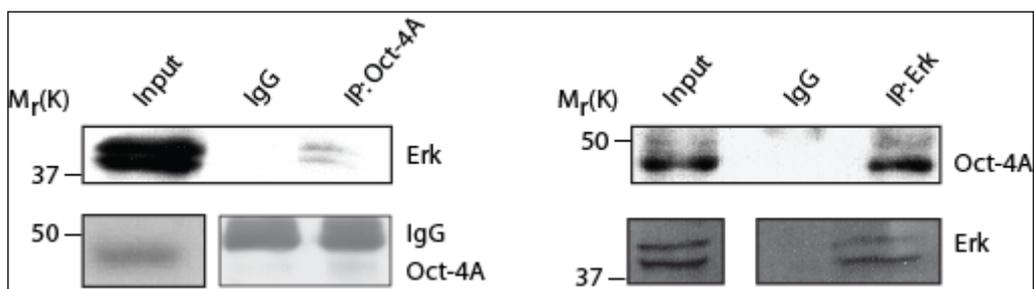


Fig. 4.10 – I lisati delle cellule NTera2 sono stati sottoposti a immunoprecipitazione utilizzando un anticorpo anti-Oct-4A o anti-Erk. L'interazione endogena tra Oct-4A ed Erk è stata valutata mediante immunoblotting con anticorpi anti-Erk e anti-Oct-4A rispettivamente. La specificità dell'interazione è stata testata utilizzando IgG di controllo.

4.4- Erk media la down-regolazione di Oct-4A

Per verificare se la pathway delle MAP kinasi interviene nella regolazione di Oct-4A, abbiamo innanzitutto testato gli effetti dell'attivazione e dell'inibizione di questa via di segnalazione sull'espressione di Oct-4A.

A tal scopo le cellule NTera-2 sono state trattate con concentrazioni crescenti

dell'inibitore PD98059 per 4 ore, osservando un aumento dose-dipendente nei livelli di espressione di Oct-4A concomitante con l'inibizione dell'attività di Erk (Fig. 4.11A, 4.12). Il trattamento con FGF2, attivatore della via di Erk, determina invece una diminuzione dei livelli di espressione di Oct-4A concomitante con l'attivazione di Erk (Fig. 11B, 4.12).

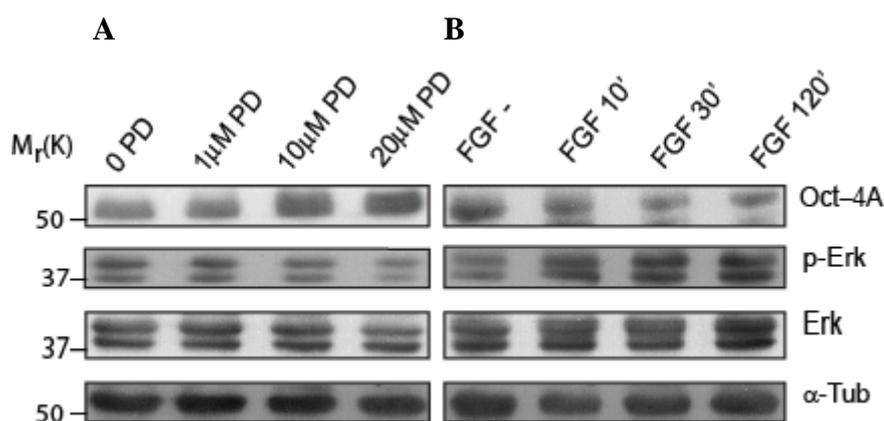


Figura 4.11- Le cellule NTERA2 sono state trattate con concentrazioni crescenti di PD98059 (A) o con FGF2 (50 NG ml⁻¹) per i tempi indicati. I lisati totali sono stati quindi sottoposti a immunoblotting utilizzando gli anticorpi indicati.

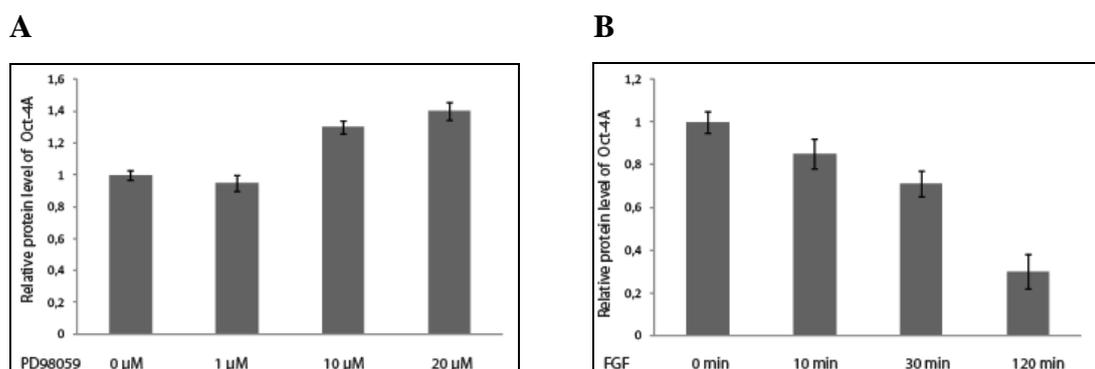


Figura 4.12- Quantificazione dell'espressione di Oct-4A sulle cellule NTERA2 trattate con concentrazioni crescenti di PD98059 (A) o con 100 ng/ml di FGF2 per diversi tempi (B). I grafici riportano i risultati di 3 esperimenti indipendenti (n=3) e la relativa deviazione standard.

Inoltre, per confermare ulteriormente la regolazione dell'espressione di Oct-4A da parte di Erk, abbiamo eseguito delle trasfezioni transienti con una forma costitutivamente attiva di

MEK1 (MEK1^{CA}) che determina un'attivazione di Erk. Le cellule trasfettate con MEK1^{CA} presentano una quantità minore di Oct-4A rispetto alle cellule non trasfettate (Fig. 4.13, 4.14).

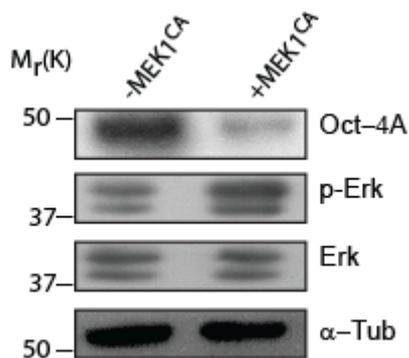


Figura 4.13- I lisati delle cellule Ntera2 trasfettate con il plasmide di controllo o il plasmide MEK1^{CA} sono stati analizzati mediante immunoblotting utilizzando gli anticorpi indicati.

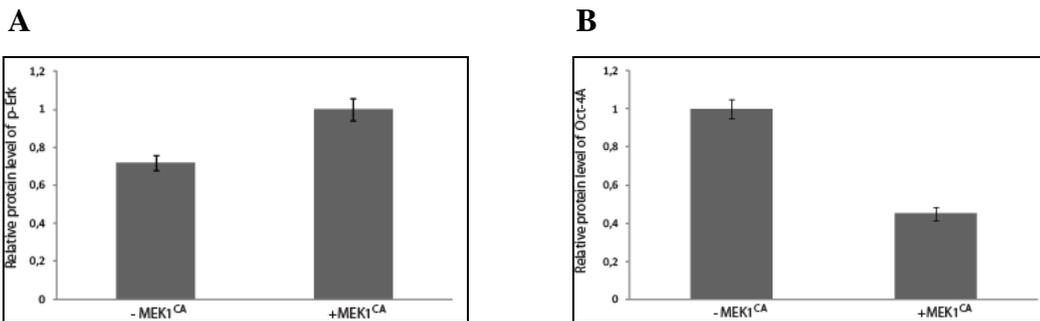


Figura 4.14- Quantificazione dell'espressione di p-Erk e Oct-4A sulle cellule trasfettate con il plasmide di controllo o MEK1^{CA}. I grafici riportano i risultati di 3 esperimenti indipendenti (n=3) e la relativa deviazione standard.

4.5- L'attivazione di Erk induce l'esclusione nucleare di Oct-4A

Visto che la traslocazione nucleo-citoplasmatica di molti fattori di trascrizione è regolata dalla fosforilazione diretta da parte di specifiche chinasi, si è volute valutare l'ipotesi di una fosforilazione di Oct-4A da parte di Erk e le eventuali implicazione sulla sua localizzazione sub-cellulare. A tale fine è stato eseguito il frazionamento nucleo-citoplasmatico delle cellule Ntera2, inducendo l'attivazione di Erk tramite il trattamento con FGF2 per 30 minuti, in presenza o in assenza dell'inibitore PD98059. Come si può vedere nelle figure 4.15 e 4.16, i livelli di espressione di Oct-4A si riducono nel nucleo ma aumentano nel citoplasma in seguito al trattamento con FGF2, andamento che viene inibito dall'aggiunta dell'inibitore PD98059. Questi risultati suggeriscono che l'attivazione di Erk è implicata nella traslocazione nucleo-citoplasmatica di Oct-4A.

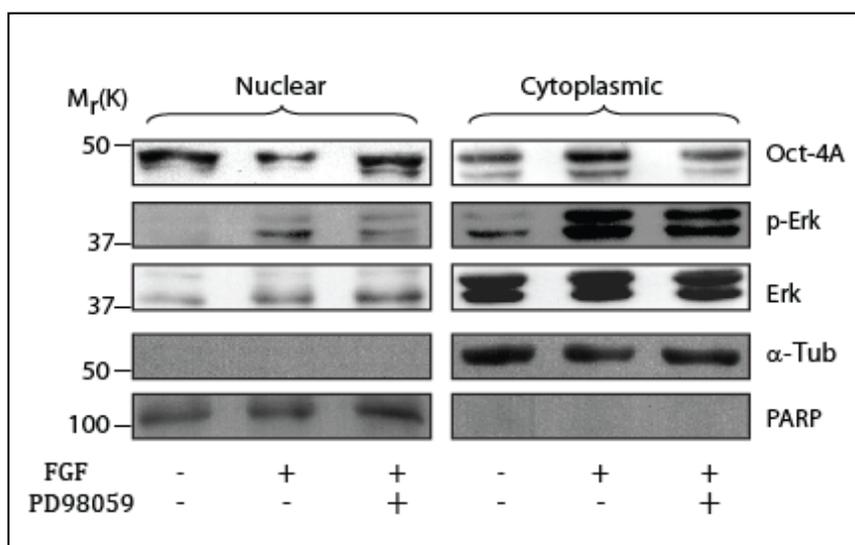


Figura 4.15- Le cellule Ntera2 sono state mantenute in assenza di siero per 18 ore, quindi sono state trattate con FGF2 (50 ng ml⁻¹) per 30 minuti in presenza (+) o assenza (-) dell'inibitore PD98059. Le frazioni nucleari (nuclear) e citoplasmatiche (cytoplasmic) sono state analizzate per immunoblotting utilizzando gli anticorpi indicati.

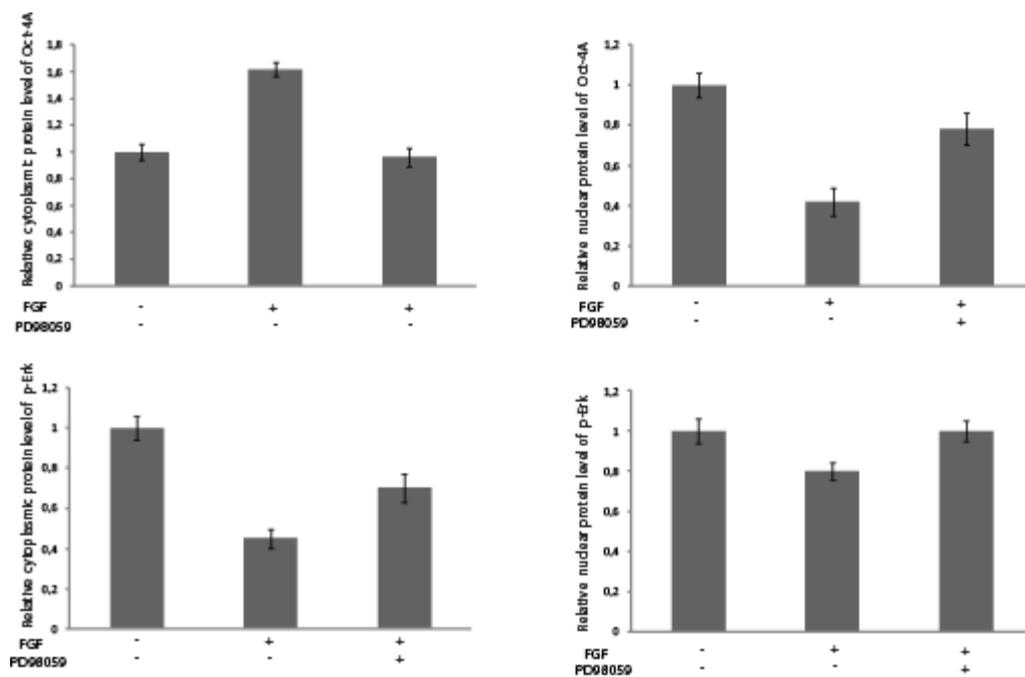


Figura 4.16- Gli estratti nucleari e citoplasmatici delle cellule NTERA2 trattate con FGF2 e PD98059 come indicato, sono stati analizzati mediante immunoblotting. La quantità relativa di Oct-4A (normalizzata rispetto alla tubulina o PARP) e p-Erk (normalizzata rispetto a Erk) sono state misurate. I grafici riportano i risultati derivati da tre esperimenti indipendenti ($n=3$) e la relativa deviazione standard.

4.6- L'attivazione di Erk incrementa i livelli di ubiquitinazione di Oct-4A

Di recente è stato dimostrato che Oct-4A viene degradata mediante ubiquitinazione da parte di WWP2 (Xu et al., 2009). Per verificare se l'attivazione di Erk influisce sui livelli di ubiquitinazione di Oct-4A, abbiamo analizzato l'ubiquitinazione endogena nelle cellule HeLa trasfettate con Flag-Oct-4A o Flag-Oct-4A e MEK1^{CA} in presenza o assenza dell'inibitore del proteosoma Lactacystina (LC). I risultati ottenuti dall'immunoprecipitazione per FLAG mostrano che i livelli di ubiquitinazione di Oct-4A aumentano quando viene co-espresso MEK1^{CA}. Tale incremento diventa più evidente dopo il trattamento con Lactacystina. Sono stati calcolati quindi i livelli di "ubiquitinazione relativa" normalizzando le bande della forma ubiquitinata con le bande relative alla forma non ubiquitinata di Oct-4A. L' "ubiquitinazione relativa" indica che è presente un incremento di circa due volte dei livelli di ubiquitinazione di Oct-4A nelle cellule co-trasfettate con MEK1^{CA} (Fig.17). Questi risultati suggeriscono che l'attivazione di Erk induce un incremento dell'ubiquitinazione di Oct-4A che potrebbe spiegare la down-regolazione osservata in seguito al trattamento con FGF2 e la co-trasfezione con MEK1^{CA}.

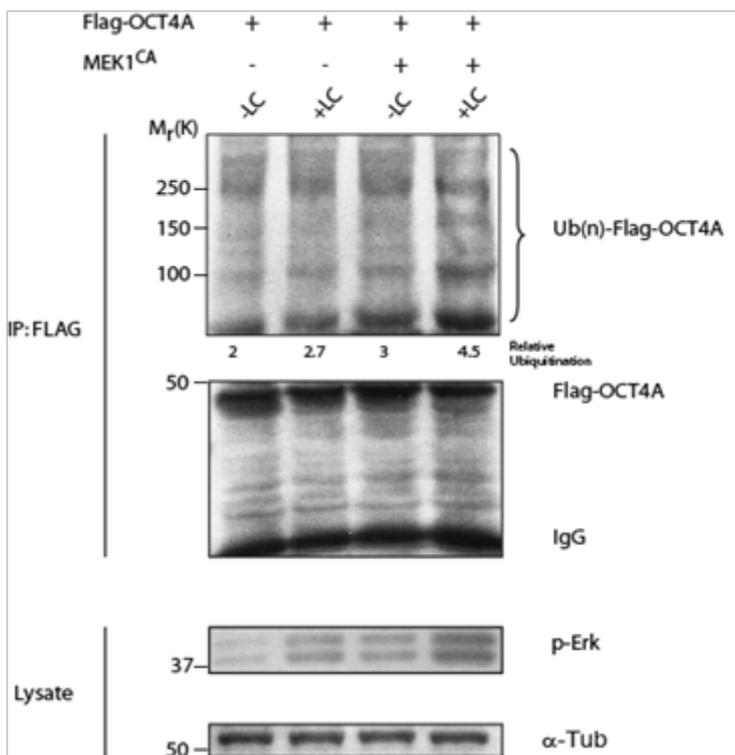


Figura 4.17- Immunoprecipitazione di Flag-Oct-4A eseguita sui lisati delle cellule HeLa in presenza o assenza dell'inibitore del proteosoma Lactacystina (LC) e co-trasfettate con MEK1^{CA} dove indicato. Sugli immunoprecipitati è stato eseguito l'Immunoblot per l'ubiquitina e Flag, sui lisati utilizzati per l'immunoprecipitazione è stato invece eseguito l'Immunoblot per p-Erk e la Tubulina. L' "ubiquitinazione relativa" corrisponde al rapporto tra la forma ubiquitinata e quella non ubiquitinata.

4.7- Oct-4A è fosforilata sulla Serina 111

Per meglio comprendere il meccanismo dell'esclusione nucleare di Oct-4A mediata da ERK, si è deciso di valutare la possibilità di una fosforilazione diretta. A sostegno di questa ipotesi l'immunoblot eseguito con l'anticorpo anti-Oct-4A sui lisati totali delle cellule Ntera2, rivela uno "shift" evidente della banda dopo il trattamento con CIP (Calf Intestinal Phosphatase) rispetto alle cellule in proliferazione che si annulla in seguito al trattamento con FGF2 (Fig. 4.18).

Per indagare ulteriormente l'ipotesi precedentemente formulata, è stato eseguito il frazionamento delle cellule HeLa trasfettate con FLAG-Oct-4A o FLAG-Oct-4A e MEK1^{CA} in presenza o assenza dell'inibitore PD98059. Gli estratti nucleari e citoplasmatici sono stati immunoprecipitati con la resina anti-FLAG e quindi analizzati mediante immunoblotting utilizzando sia un anticorpo anti-FLAG che un anticorpo anti-p-Serina (Fig. 4.19). I risultati evidenziano la presenza di una forma di Oct-4A fosforilata su uno o più residui di serina che si localizza principalmente nel citoplasma in seguito all'attivazione di Erk diminuendo nel nucleo. Inoltre l'aumento a livello citoplasmatico concomitante alla diminuzione a livello nucleare viene inibito dall'aggiunta dell'inibitore PD98059. Messi assieme questi risultati suggeriscono non solo che Erk potrebbe fosforilare Oct-4A, ma anche che questa fosforilazione potrebbe essere implicata nella sua traslocazione nucleo-citoplasmatica.

Si è quindi deciso di eseguire un'analisi spettrometrica per verificare se la fosforilazione avviene effettivamente su uno o più degli ipotetici siti di fosforilazione da parte di Erk. A tal fine, gli estratti totali delle cellule HeLa co-trasfettate con Oct-4A e MEK1^{CA}, sottoposti a immunoprecipitazione, sono stati analizzati mediante spettrometria di massa come descritto nei Materiali e Metodi (Fig. 4.20, 4.21). I risultati ottenuti hanno

che la Serina 111, uno dei possibili siti di fosforilazione da parte di Erk, viene effettivamente fosforilata nelle nostre condizioni sperimentali.

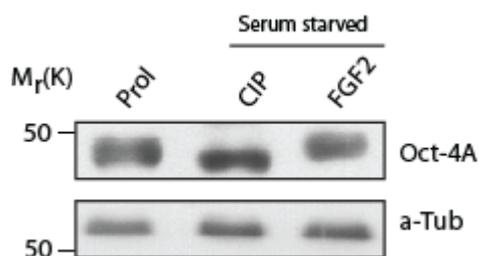


Figura 4.18- Immunoblot per Oct-4A eseguito sui lisati delle cellule Ntera2 in proliferazione, trattate con Calf Intestinal Phosphatase o FGF2.

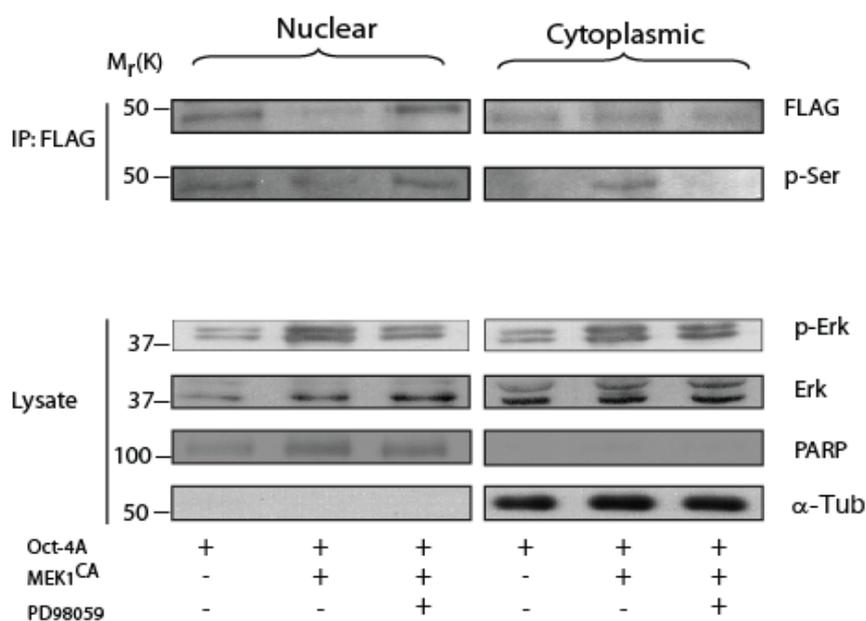


Figura 4.19- Gli estratti nucleari e citoplasmatici delle cellule Hela co-trasfettate con il plasmide FLAG-Oct-4A e il plasmide di controllo o MEK1^{CA}, sono stati immunoprecipitati utilizzando beads di agarosio coniugate all'anticorpo anti-FLAG. Dove indicato le cellule trasfettate sono state trattate con PD98059 per 4 ore prima della lisi. Gli immunoprecipitati sono stati analizzati mediante immunoblotting con gli anticorpi anti-phospho-serina e anti-FLAG. I lisati non immunoprecipitati sono stati sottoposti a immunoblotting con gli anticorpi indicati.

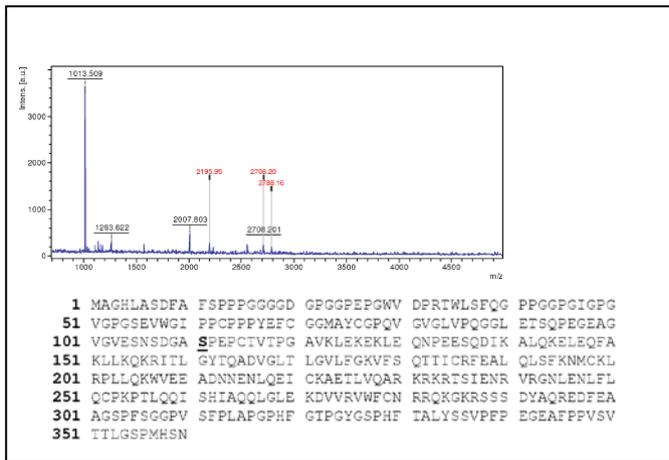


Figura 4.20- I lisati delle cellule HeLa cotrasfettate con Oct4A e MEK1^{CA} sono stati sottoposti a immunoprecipitazione con la resina anti-FLAG. L'eluato è stato utilizzato per l'analisi spettrometrica per l'identificazione dei fosfo-peptidi. In figura vengono riportati lo spettro del campione purificato su micro-colonna di TiO₂ e la posizione della fosfo-serina identificata.

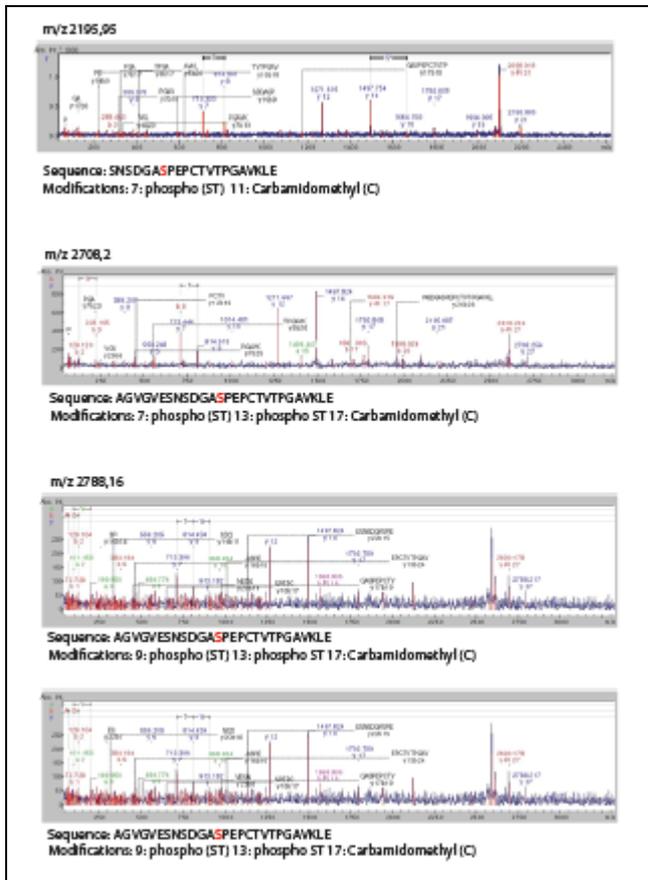


Figura 4.21- Spettri MS/MS annotati.

5- DISCUSSIONE

Molti studi presenti in letteratura hanno più volte rimarcato il ruolo chiave di Oct-4A nell'auto-mantenimento cellulare, nella pluripotenzialità e nella successiva determinazione della linea germinale.

L'espressione di Oct-4A sembra essere limitata alle cellule indifferenziate e in via di differenziamento dell'embrione: il fatto che l'espressione di Oct-4A sia confinata alle cellule della massa interna (ICM) fino al momento in cui il trofoectoderma si differenzia, suggerisce che la proteina è richiesta per il mantenimento dello stato pluripotente e che la sua presenza è proibitiva per lo stadio di differenziamento. Uno studio recentemente pubblicato (2009) ha dimostrato però che la *down*-regolazione dell'espressione di Oct-4A è responsabile della perdita del mantenimento della pluripotenza e la sua *over*-espressione induce differenziamento nelle cellule staminali embrionali (ES). Questi dati suggeriscono che sono necessari particolari livelli di espressione di Oct-4A affinché la proteina sia in grado di sostenere la pluripotenzialità cellulare (Kashyap V. et al., 2009).

I dati forniti dalla letteratura inducono a considerare Oct-4A come un importante fattore di regolazione embrionale, i cui geni target identificati comprendono una serie di geni noti per le loro criticità durante le fasi di sviluppo precoci (FGF4 e hCG): Infatti una riduzione nell'espressione di Oct-4A, induce una diminuzione dell'attivazione di alcuni geni target, come FGF4, o una de-regolazione di altri, come α e β hCG, dovuta alla limitante quantità di Oct-4A (Pesce M. et al., 1999; Pesce M. et al., 2001).

E' stato dimostrato che la proteina Oct-4A viene espressa anche in popolazioni di cellule staminali umane derivate da vari tessuti dell'adulto e che essa esercita un ruolo fondamentale nel supportare il loro potenziale differenziativo multilinea sia *in vitro* sia *in vivo*. Tali studi hanno dunque portato a considerare Oct-4A come un *marker* di staminalità in diversi tessuti adulti (Zangrossi S. et al., 2007).

L'attività di regolazione trascrizionale mediata da Oct-4A è associata all'azione di una serie, non ancora del tutto definita, di fattori in un unico e complesso circuito regolatorio. Oct-4A funziona, infatti, in cooperazione con altri fattori regolatori per attivare o inibire bersagli genici specifici in particolari tipi cellulari. Oct-4A viene considerato uno dei componenti principali di questo *network* di regolazione trascrizionale che esercita un controllo biologico

preciso tramite interazioni proteina-proteina. Studi approfonditi sulla funzionalità di Oct-4A possono certamente aiutare a comprendere con maggiore chiarezza i complessi meccanismi molecolari e cellulari responsabili della pluripotenzialità delle cellule staminali.

Infatti, sebbene i networks trascrizionali controllati da Oct-4A siano stati largamente delineati (Boyer LA et al., 2005; Loh YH et al., 2006), tuttavia i meccanismi regolatori di questo fattore di trascrizione rimangono ancora poco chiari.

Lo scopo principale di questo lavoro di tesi è stato quindi quello di ricercare nuovi interattori molecolari potenzialmente coinvolti nella regolazione di Oct4-A. I saggi di interazione funzionale, sia *in vitro* (GST pull-down) che *in vivo* (co-immunoprecipitazione), da noi eseguiti, hanno portato all'identificazione di Erk come nuovo interattore.

L'identificazione di potenziali interattori dell'isoforma A di Oct-4 tramite il saggio di GST pull-down ha recentemente portato alla scoperta di una nuova proteina regolatrice di Oct-4A, PKM2, un isozima della piruvato chinasi con funzionalità glicolitica (Lee J. et al., 2008).

Un'altro ipotetico interattore di Oct-4A che è stato identificato da Li X. et al. in un lavoro pubblicato nel 2008, è la proteina karioperina-alpha (*KPNA-2*). *KPNA-2* è un membro della famiglia delle importine alpha coinvolto nel trasporto attivo, dal citoplasma al nucleo, di proteine contenenti la sequenza segnale di localizzazione nucleare (NLS).

Oct-4A, come l'isoforma B di Oct-4, presenta nella sua sequenza aminoacidica una sequenza di localizzazione nucleare funzionale (NLS), che viene riconosciuta da *KPNA-2* e che è quindi responsabile della sua traslocazione e della sua localizzazione nucleare (Li X. et al., 2008).

L'interazione tra Erk e Oct-4A non è mai stata riportata in letteratura ed è anche per questo motivo che abbiamo voluto indagare ulteriormente gli effetti della sua attivazione su Oct-4A e quindi, di conseguenza, sui meccanismi che regolano il mantenimento della staminalità e il differenziamento.

Erk fa parte della via segnalatoria Ras-MAPK che risulta implicata in molti processi cellulari, quali proliferazione, differenziamento, sopravvivenza e metabolismo. La deregolazione di questa *pathway* è associata a diverse patologie e risulta correlata sia alla

comparsa di disordini proliferativi e tumori sia ad uno sviluppo embrionale anormale (Malumbres M. et al., 2003; Bentires-Alj M. et al., 2006; Lloyd A., 2006).

Erk è una *MAP-chinasi*, proteina chinasi attivata da mitogeni, coinvolta nella via di trasduzione del segnale attivata da *Ras*, una proteina monomerica con attività GTPasica. *Ras* è a sua volta attivata da recettori tirosina chinasi che in risposta a ligandi specifici, quali *EGF* (*Epidermal Growth Factor*), *FGF* (*Fibroblast Growth Factor*), *PDGF* (*Platelets Derived Growth Factor*), dimerizzano e si autofosforilano attivando *Ras* e la via di trasduzione mediata da questa proteina (Fig. 5.1).

Una caratteristica tipica delle MAP-chinasi è che la sua attivazione completa richiede la fosforilazione sia di una treonina che di una tirosina, che sono separate nella proteina da un singolo aminoacido. La proteina chinasi che catalizza entrambe queste fosforilazioni si chiama *MAP-chinasi-chinasi*, che nella via di segnalazione di *Ras* nei mammiferi si chiama *Mek*. La necessità di una fosforilazione sia in tirosina che in treonina assicura che la MAP-chinasi sia mantenuta inattiva a meno di un'attivazione specifica da parte di una MAP-chinasi-chinasi. La stessa MAP-chinasi-chinasi è attivata da una fosforilazione catalizzata dalla prima chinasi del modulo a tre componenti, la *MAP-chinasi-chinasi-chinasi*, che nella via di segnalazione di *Ras* è chiamata *Raf* e viene attivata da *Ras* attiva. Una volta attivata, la MAP-chinasi trasmette il segnale a valle fosforilando varie proteine bersaglio, comprese proteine che regolano geni e altre proteina chinasi. La MAP-chinasi entra nel nucleo e fosforila uno o più componenti di un complesso che regola geni. Ciò attiva la trascrizione di una serie di geni precoci immediati, così chiamati perché si accendono nel giro di minuti dopo la stimolazione della cellula da parte di un segnale extracellulare. Alcuni di questi geni codificano altre proteine che regolano a loro volta geni diversi: in questo modo la via di segnalazione Ras-MAP-chinasi trasmette segnali dalla superficie cellulare al nucleo e altera lo schema di espressione genica in modo significativo. Fra i geni attivati da questa via ci sono quelli necessari per la proliferazione cellulare, quali quelli che codificano le *cicline G₁*.

Le MAP-chinasi sono in genere attivate soltanto temporaneamente in risposta a segnali extracellulari e il periodo di tempo in cui rimangono attive può influenzare profondamente la natura della risposta. Queste proteine sono inattivate per defosforilazione e la rimozione specifica dalla tirosina o dalla treonina è sufficiente ad inattivare l'enzima.

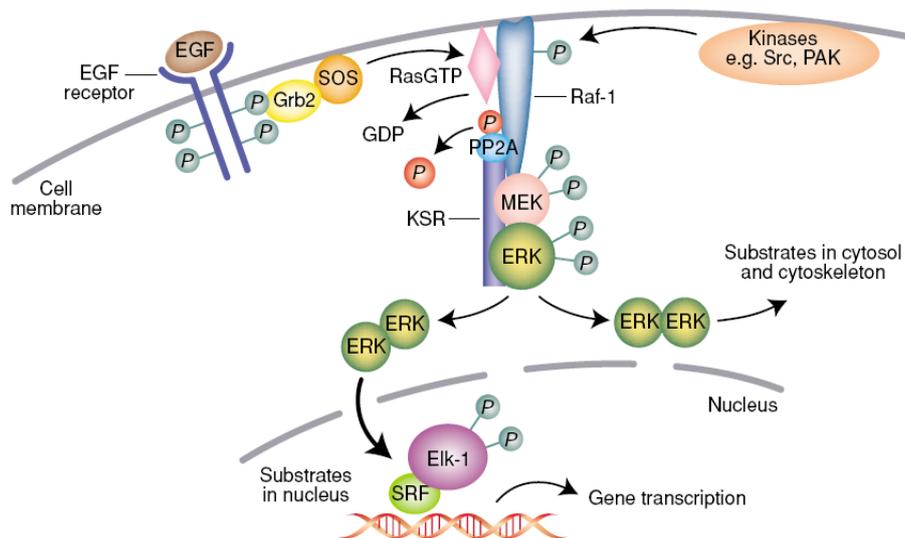


Fig 5.4- Rappresentazione schematica dell'organizzazione e della funzione della via segnalatoria Ras-Raf-Mek-Erk. Il legame con EGF induce la dimerizzazione e l'autofosforilazione di residui di tirosin presenti nella porzione citosolica del recettore. Queste fosfo-tirosine si comportano da siti d'ancoraggio per molecole segnalatorie le quali includono il complesso Grb2-Sos che attiva Ras stimolando lo scambio del GDP con GTP. Questo scambio determina un cambiamento conformazionale di Ras che ora può legarsi a Raf-1, reclutandolo dal citosol alla membrana cellulare, dove ha luogo l'attivazione Raf-1. Raf attivata fosforila e attiva Mek, una proteina chinasi-chinasi, che, a sua volta, fosforila e attiva Erk, una MAP-chinasi. Erk attivata possiede molti substrati bersaglio nel citosol, quali proteine citoscheletriche, fosfolipasi A2, recettori per estrogeni, Sos, Stats e altri. Erk è in grado di entrare nel nucleo e controllare l'espressione genica, fosforilando fattori di trascrizione come Elk-1 e altre proteine della famiglia Ets. www.expertreviews.org

Sebbene i requisiti necessari al mantenimento delle cellule staminali embrionali in uno stato di autorinnovo e di pluripotenzialità, stiano diventando via via più chiari, ciò che rimane ancora oscuro è il processo tramite cui le cellule embrionali entrano in differenziamento. Quello che si è dimostrato con certezza è che FGFs (Fibroblast Growth Factors) e la cascata segnalatoria Ras-Erk rappresentano stimoli critici per la proliferazione e il differenziamento in molti tipi cellulari (Roux P. et al., 2004).

I risultati ottenuti negli esperimenti eseguiti in questo lavoro di tesi, hanno dimostrato che l'attivazione della via segnalatoria RAS-MAPK induce la traslocazione nucleocitoplasmatica di Oct-4A. L'esportazione dei fattori di trascrizione dal nucleo è stata implicata nel controllo di numerosi processi nucleari, primo fra tutti la trascrizione genica. Chiaramente,

il sequestro di un fattore di trascrizione nel citoplasma o addirittura la sua successiva degradazione, rappresentano potenti meccanismi che regolano negativamente la sua attività.

I nostri dati suggeriscono inoltre un più ampio coinvolgimento della via RAS-MAPK nella regolazione di Oct-4A evidenziando la sua implicazione nel processo di ubiquitinazione e degradazione via proteosoma.

L'analisi spettrometrica ha evidenziato che la Serina 111 risulta essere fosforilata nelle nostre condizioni sperimentali. Essendo questo un potenziale sito di fosforilazione da parte di Erk, stiamo al momento valutando, mediante analisi di mutagenesi, la sua possibile implicazione nei meccanismi di traslocazione nucleo-citoplasmatica, ubiquitinazione e degradazione via proteosoma osservati nei nostri esperimenti.

Considerando l'importanza della *pathway* Ras-MAPK in moltissimi eventi cellulari, quale il differenziamento cellulare, e il ruolo fondamentale che Oct-4A ricopre nel processo di auto-mantenimento e nella pluripotenzialità delle cellule staminali (sia embrionali sia che adulte), l'interazione Oct-4A-Erk potrebbe rivelarsi estremamente utile per meglio comprendere i meccanismi che regolano il mantenimento della staminalità.

6- BIBLIOGRAFIA:

Alison M.R., Poulsom R., Forbes S. and Wright N.A., *An introduction to stem cells*. J Pathol. 2002; 197: 419:423.

Alonso L. and Fuchs E., *Stem cells of the skin epithelium*. Proc Natl Acad Sci USA 2003; 100(1): 11830-11835.

Ambrosetti D.C., Basilico C. and Dailey L., *Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites*. Mol Cell Biol. 1997; 17: 6321-6329.

Atlasi Y., Mowla S.J., Ziaee S.A.M., Gokhale P.J., Andrews P.W., *Oct/4 spliced variants are differentially expressed in human pluripotent and non-pluripotent cells*. Stem cell 2008.

Aurora R. and Herr W., *Segments of the POU domain influence one another's DNA-binding specificity*. Mol. Cell. Biol. 1992; 12:455-467.

Avilion A.A., Nicolis S.K., Pevny L.H., Perez L., Vivian N. and Lovell-Badge R., *Multipotent cell lineages in early mouse development depend on Sox2 function*. Genes & Development 2003; 17:126-140.

Babaie J., Herwig R., Greber B., Brink T.C., Wruck W., Groth D., Lehrach H., Burdon T. and Adjaye J., *Analysis of Oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cell*. Stem Cell. 2007; 25: 500-510.

Banerjee-Basu S. and Baxevanis A.D., *Molecular evolution of the homeodomain family of transcription factors*. Nucleic Acids Res. 2001; 29(15): 3258-3269.

Beachy P.A., Varkey J., Young K.E., Von Kessler D.P., Sun B.I. and Ekker S.C., *Cooperative binding of an Ultrabithorax homeodomain protein to nearby and distant DNA sites*. Mol Cell

Biol. 1993; 13(11): 6941-6956.

Ben-Shushan E., Pikarsky E., Klar A. and Bergman Y., *Extinction of Oct3/4 gene expression in embryonal carcinoma x fibroblast somatic cell hybrids is accompanied by changes in the methylation status, chromatin structure, and transcriptional activity of the Oct3/4 upstream region.* Mol Cell Biol. 1993; 13: 891-910.

Blau H.M., Brazelton T.R. and Weimann J.M., *The evolving concept of stem cell: entity or function?* Cell 2001; 105: 829-841.

Botfield M.C., Jancso A. and Weiss M.A., *Biochemical characterization of the Oct-2 POU domain with implications for bipartite DNA recognition.* Biochemistry 1992; 31:5841-5848.

Boyer L.A., Lee T.I., Cole M.F. et al., *Core transcriptional regulatory circuitry in human embryonic stem cells.* Cell 2005; 122: 947-956.

Brehm A., Ohbo K. and Schöler H.R., *The carboxy-terminal transactivation domain of Oct4 acquires cell specificity through the POU domain.* Mol Cell Biol. 1997; 17: 154-162.

Brehm A., Ohbo K., Zwerschke W., Botquin V., Jansen-Durr P. and Schöler H.R., *Synergism with germ line transcription factor Oct4: viral oncoproteins share the ability to mimic a stem cell-specific activity.* Mol Cell Biol. 1999; 19: 2635-2643.

Brittan M. and Wright N.A., *Stem cell in gastrointestinal structure and neoplastic development.* Gut. 2004; 53: 899-910.

Butteroni C., De F.M., Schöler H.R. and Pesce M., *Phage display screening reveals an association between germline- specific transcription factor Oct4 and multiple cellular proteins.* J Mol Biol. 2000; 304: 529-540.

Catena R., Tiveron C., Ronchi A., Porta S., Ferri A., Tatangelo L., Cavallaro M., Favaro R.,

Reinbold R., Schöler H. and Nicolis S.K., *Conserved POU binding DNA sites in the Sox 2 upstream enhancer regulate gene expression in embryonic and neural stem cells*. J Biol Chem 2004; 279(40): 41846-41857.

Cauffman G., Liebaers I., Van Steirteghem A. and Van De Velde H., *POU5F1 isoforms show different expression patterns in human embryonic stem cells and preimplantation embryos*. Stem Cell 2006; 24: 2685-2691.

Chambers I., Colby D., Robertson M., Nichols J., Lee S., Tweedie S. and Smith A., *Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells*. Cell 2003; 113:643-655.

Crouau-Roy B., Amadou C., Bouissou C., Clayton J., Vernet C., Ribouchon M.T. and Pontarotti P., *Localization of the OTF3 gene within the human MHC class I region by physical and meiotic mapping*. Genomics. 1994; 21: 241-243.

Curatola A.M. and Basilico C., *Expression of the K-fgf protooncogene is controlled by 3' regulatory elements which are specific for embryonal carcinoma cells*. Mol Cell Biol. 1990; 10: 2475-2484.

Deans R.J. and Moseley A., *Mesenchymal stem cells: Biology and potential clinical uses*. Exp Hematol. 2000; 28: 875.

Fenyo D., Qin J. and Chait B.T., *Protein identification using mass spectrometric informations*. Electrophoresis 1998; 19: 998-1005.

Fuhrmann G., Chung A.C., Jackson K.J., Hummelke G., Baniahmad A., Sutter J., Sylvester I., Schöler H.R. and Cooney A.J., *Mouse germline restriction of Oct4 expression by germ cell nuclear factor*. Dev Cell. 2001; 1: 377-387.

Furukubo-Tokunaga K., Flister S. and Gehring W.J., *Functional specificity of the Antennapedia homeodomains*. Proc Natl Acad Sci U S A. 1993; 90(13): 6360-6364.

Tesi di Dottorato di Renza Spelat discussa presso l'Università degli Studi di Udine

Gage F.H., *Mammalian neural stem cells*. Science 2000; 287: 1433-1438.

Gehring W.J., Quian Y.Q., Billeter M., Furukubo-Tokunaga K., Schier A.F., Resendez-Perez D., Affolter M., Otting G. and Wüthrich K., *Homeodomain-DNA recognition*. Cell. 1994; 78(2): 211-223.

Gidekel S, Pizov G, Bergman Y et al. *Oct3/4 is a dose-dependent oncogenic fate determinant*. Cancer Cell. 2003; 4: 361-370.

Goto T., Adjaye J., Rodeck C. and Monk M., *Identification of genes expressed in human primordial germ cells at the time of entry of the female germ line into meiosis*. Mol Hum Reprod. 1999; 5: 851-860.

Gronthos S., Mankani M., Brahimi J., Robey P.G. and Shi S., *Postnatal human dental pulp stem cells (DPSCs) in vitro and in vivo*. Proc Natl Acad Sci U S A. 2000; 97: 13625-13630.

Gronthos S., Zannettino A.C., Hay S.J., Shi S., Graves S.E., Kortessidis A. et al. *Molecular and cellular characterization of highly purified stromal stem cells derived from human bone marrow*. J Cell Sci. 2003; 116: 1827-1835.

Guillaudoux T., Mattei M.G., Depetris D., Le Bouteiller P., Pontarotti P., *In situ hybridization localizes the human OTF3 to chromosome 6p21.3-p22 and OTF3L to 12p13*. Cytogenet Cell Genet. 1993; 63: 212-214.

Guo Y., Costa R., Ramsey H., Starnes T., Vance G., Robertson K., Kelley M., Reinbold R., Schöler H.R. and Hromas R., *The embryonic stem cell transcription factors Oct4 and FoxD3 interact to regulate endodermal-specific promoter expression*. Proc Natl Acad Sci U S A. 2002; 99: 3663-3667.

Hansis C., Grifo J. and Krey L., *Oct4 expression in inner cell mass and trophectoderm of human blastocysts*. Mol Hum Reprod. 2000; 6: 999-1004.

Hansis C, Tang Y, Grifo J and Krey L., *Analysis of Oct4 expression and ploidy in individual human blastomeres*. Mol Hum Reprod. 2001;7: 155-161.

Hay D.C., Sutherland L., Clark J. et al., *Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells*. Stem Cell 2004; 22: 225-235.

Herr W. and Cleary M.A., *The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain*. Genes & Development 1995; 9:1679-1693.

Herr W., Sturm R.A., Clerc R.G., Corcoran L.M., Baltimore D., Sharp P.A., Ingraham H.A., Rosenfeld M.G., Finney M., Ruvkun G. and Horvitz H.R., *The POU domain: a large conserved region in the mammalian Pit-1, Oct-1, Oct-2, and Caenorhabditis Elegans Unc-86 gene products*. Genes & Development 1988; 2:1513-1516.

Holland P.W.H., Booth H.A.F. and Bruford E.A., *Classification and nomenclature of all human homeobox genes*. BMC Biology. 2007; 5(47).

Hromas R. and Costa R., *The hepatocyte nuclear factor-3/forkhead transcription regulatory family in development, inflammation and neoplasia*. Crit Rev Oncol Hematol 1995; 20: 12940.

Hummelke G.C. and Cooney A.J., *Germ cell nuclear factor is a transcriptional repressor essential for embryonic development*. Front Biosci. 2001; 6: 1186-1191.

Jacob F. and Monod J., *Genetic regulatory mechanism in the synthesis of proteins*. J Mol Biol. 1961; 3: 318-356.

Jaenisch R., *DNA methylation and imprinting: why bother?* Trends Genet.1997; 13: 323-329.

Keller G., *Embryonic stem cell differentiation: emergence of a new era in biology and medicine*. Genes & Development 2005; 19(10): 1129-1155.

Tesi di Dottorato di Renza Spelat discussa presso l'Università degli Studi di Udine

Kerkis I., Kerkis A., Dozortsev D., Stukart-Parsons G.C., Massironi S.M.G., Pereira L.V., Caplan A.I. and Cerruti H.F., *Isolation and characterization of a population of immature dental pulp stem cells expressing OCT-4 and other embryonic stem cell markers*. Cells Tissues Organs 2006; 184: 105–116.

Klemm J.D., Rould M.A., Aurora R., Herr W. and Pabo C.O., *Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules*. Cell 1994; 77:21-32.

Krishnan B.R., Jamry I. and Chaplin D.D., *Feature Mapping of the HLA Class I Region: Localization of the POU5F1 and TCF19 Genes*. Genomics. 1995; 30: 53-58.

Kristie T.M. and Sharp P.A., *Interactions of the Oct-1 POU subdomains with specific DNA sequences and the HSV α -trans-activator protein*. Genes & Development 1990; 4: 2383-2396.

Kuroda T., Tada M., Kubota H. et al., *Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression*. Mol Cell Bio 2005; 25: 2475-2485

Lamb K.A. and Rizzino A., *Effects of differentiation on the transcriptional regulation of the FGF-4 gene: critical roles played by a distal enhancer*. Mol. Repro. Dev. 1998; 51: 218-224.

Laughon A., *DNA binding specific of homeodomains*. Biochemistry. 1991; 30(48): 11357-11367.

Leblond, *Classification of cell populations on the basis of their proliferative behavior*. National Cancer Institute 1964; 14: 119-150.

Lee J., Kim H.K., Rho J.-Y., Han Y.-M. and Kim J., *The human OCT-4 isoforms differ in their ability to confer self-renewal*. The Journal of Biological Chemistry 2006; 281(44): 33554-33565.

Lenardo M.J., Staudt L., Robbins P. et al., *Repression of the Igh enhancer in teratocarcinoma*
Tesi di Dottorato di Renza Spelat discussa presso l'Università degli Studi di Udine

cells associated with a novel octamer factor. Science. 1989; 243: 544-546.

Liang J., Wan M., Zhang Y., Gu P., Xin H., Yung S.Y., Quin J., Wong J., Cooney A.J., Liu D., Songyang Z., *Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells.* Nat Cell Biol. 2008; 10(6): 731-739.

Liedtke S., Stephan M. and Kögler G. *Oct4 expression revisited: potential pitfalls for data misinterpretation in stem cell research.* Biol Chem. 2008; 389: 845-850.

Liu L. and Roberts R., *Silencing of the gene for the beta subunit of human chorionic gonadotropin by the embryonic transcription factor Oct3/4.* J Biol Chem. 1996; 271: 16683-16689.

Liu L., Leaman D., Villalta M. and Roberts R. *Silencing of the gene for the alpha-subunit of human chorionic gonadotropin by the embryonic transcription factor Oct3/4.* Mol Endocrinol. 1997; 11: 1651-1658.

Lovell-Badge R., *The future for stem cell research.* Nature 2001; 414:88-91.

Mann M. and Wilm M., *Error-tollerant identification of peptides in sequence database by peptide sequence tags.* Anal. Chem. 1994; 66: 4390-99.

Merika M. and Thanos D., *Enhanceosomes.* Curr Opin Genet Dev. 2001; 11(2): 205-208. Review.

Monk M. and Holding C., *Human embryonic genes re-expressed in cancer cells.* Oncogene. 2001; 20: 8085-8091.

Morrison S.J. and Kimble J., *Asymmetric and symmetric stem-cell divisions in development and cancer.* Nature 2006; 441: 1068-1074.

Miyazono M., Lee V.M. and Trojanowski. *Proliferation, cell death, and neuronal differentiation in transplanted human embryonal carcinoma (NTera2) cells depend of the graft*

Tesi di Dottorato di Renza Spelat discussa presso l'Università degli Studi di Udine

site in nude and severe combined immunodeficient mice. Lab Invest. 1995; 73: 273-283.

Nichols J, Zevnik B, Anastassiadis K et al., *Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4.* Cell. 1998; 95:379-391.

Nischimoto M., Fukushima A., Okuda A. and Muramatsu M., *The gene for embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct3/4 and Sox-2.* Mol Cell Biol 1999; 19: 5453-65.

Nischimoto M., Miyagi S., Katayanagi T., Tomioka M., Muramatsu M. and Okuda A., *The embryonic Octamer factor 3/4 displays distinct DNA binding specificity from those of other octamer factors.* Biochem Biophys Res Commun, 2003; 302(3): 581-586.

Niwa H., Miyazaki J. and Smith A.G., *Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells.* Nat. Genet. 2000; 24:372-376.

Niwa H., *Molecular mechanism to maintain stem cell renewal of ES cells.* Cell Str & Func. 2001; 26: 137-148.

Nordhoff V., Hübner K., Bauer A., Orlava I., Malapetsa A. and Schöler H.R., *Comparative analysis of human, bovine, and murine Oct4 upstream promoter sequence.* Mamm Genome. 2001; 12(4): 309-317.

Oakley B.R., Kirsch D.R., Morris N.R., *A simplified ultrasensitive silver stain for detecting proteins in polyacrylamide gels.* Anal Biochem. 1980; 105(2): 361-363.

Okamoto K., Okazawa H., Okuda A. et al., *A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells.* Cell. 1990; 60: 461-472.

Okazawa H., Okamoto K., Ischino F. et al., *The oct3 gene, a gene for an embryonic transcription factor, is controlled by a retinoic acid repressible enhancer.* EMBO J. 1991; 10: 2997-3005.

Ovitt C.E. and Schöler H.R., *The molecular biology of Oct4 in the early mouse embryo*. Mol Hum Reprod. 1998; 4(11): 1021-1031.

Pabo C.O. and Sauer R.T., *Transcription Factors: Structural Families and Principles of DNA Recognition*. Annual Review of Biochemistry 1992; 61:1053-1095.

Palmieri S., Peter W., Hess H., Schöler H.R., *Oct4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation*. Dev Biol. 1994; 166: 259-267.

Pan G., Qin B., Liu N. et al., *Identification of a nuclear localization signal in OCT4 and generation of a dominant negative mutant by its ablation*. J Biol Chem 2004; 279: 37013-37020.

Pan G. and Thomson J.A., *Nanog and transcriptional networks in embryonic stem cell pluripotency*. Cell Research 2007; 17: 42-49.

Pappin D.J., Hojrup P. and BleasPesceby A.J., *Rapid identification of proteins by peptide-mass fingerprinting*. Curr Biol. 1993; 3(6): 327-32.

Pesce M., Wang X., Wolgemuth D. and Schöler H.R., *Differential expression of the Oct4 transcription factor during mouse germ cell differentiation*. Mech Dev. 1998a; 71: 89-98.

Pesce M., Gross M., Schöler H.R., *In line with our ancestors: Oct4 and the mammalian germ*. Bioessays. 1998b; 20: 722-732.

Pesce M., Anastasiadis K. and Schöler H.R., *Oct-4: lessons of totipotency from embryonic stem cells*. Cells Tissues Organs 1999; 165(3-4): 144-152.

Pesce M and Schöler HR., *Oct4: control of totipotency and germline determination*. Mol Reprod Dev. 2000; 55: 452-457.

Pesce M. and Schöler H.R., Oct-4: *Gatekeeper in the beginning of mamalian development*. Stem Cells 2001; 19: 271-278.

Petruz M.F., *Mechanisms of cooperativity and allosteric regulation in proteins*. Cambridge University Press, Cambridge (1990).

Phillips K. and Luisi B., *The virtuoso of versatility: POU proteins that Flex to Fit*. J. Mol. Biol. 2000; 302:1023-1039.

Pierdomenico L., Bonsi L., Calvitti M. et al., *Multipotent Mesenchymal Stem Cells with Immunosoppressive Activity Can Be Easily Isolated from Dental Pulp*. Transplantation. 2005; 80(6): 836-842.

Pleasure S.J. and LeeV.M., *Ntera 2 vells: a human cell line which displays characteristics expected of a human committed neuronal progenitor cell*. J Neurosci Res. 1993; 35: 585-602.

Quian Y.Q., Billeter G., Gehring W.J. and Wüthrich K., *The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: Comparison with prokaryotic repressor*. Cell. 1989; 59(3): 573-580.

Quian Y.Q., Otting G., Furukubo-Tokunaga K., Affolter M, Gehring WJ and Wüthrich., *NMR structure determination reveals that the homeodomain is connected through a flexible linker to the main body in the Drosophila Antennapedia protein*. Proc Natl Acad Sci U S A. 1992; 89(22): 10738-10742.

Reubinoff B.E., Pera M.F., Fong C.Y., Trounson A., Bongso A., *Embryonic stem cell lines from human blastocysts: somatic differentiatyion in vitro*. Nature Biotechnol. 2000; 18: 399-404.

Rosner M.H., Viganao M.A., Ozato K. et al., *A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo*. Nature. 1990; 345: 686-692.

Roux P. P. and Blenis J., *ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions*. Microbiol. Mol. Bio. 2004; 68: 320-344.

Ruvkin G. and Finney M., *Regulation of transcription and cell identity by POU domain proteins*. Cell. 1991; 64: 475-478.

Ryan A.K. and Rosenfeld M.G., *POU domain family values: flexibility, partnership and developmental codes*. Genes & Development 1997; 11:1207-1225.

Salazar-Ciudad I. and Jernvall J., *A gene network model accounting for development and evolution of mammalian teeth*. Proc Natl Acad Sci U S A. 2002; 99: 8116-8120.

Sauvageau G., Lansdorp P.M., Eaves C.J., Hogge D.E., Dragowska W.H., Reid D.S., Largman C., Lawrence H.J. and Humphries R.K., *Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells*. Proc Natl Acad Sci U S A. 1994; 91(25): 12223-12227.

Schöler H.R., Hatzopoulos A.K., Balling R. et al., *A family of octamer-specific proteins present during mouse embryogenesis: evidence of germline-specific expression of an Oct factor*. EMBO J. 1989; 8: 2543-2550.

Schöler H.R., Ruppert S., Suzuki N. et al., *New type of POU domain in germ line-specific protein Oct4*. Nature. 1990a; 334: 435-439.

Schöler H.R., Dressler G.R., Balling R. et al., *Oct4: a germline-specific transcription factor mapping to the mouse t-complex*. EMBO J. 1990b; 9: 2185-2195.

Schöler H.R., *Octamania: the POU factors in murine development*. Trends Genet. 1991A; 7: 323-329.

Schöler H.R., Ciesiolka T., Gruss P., *A nexus between Oct4 and EAI: implications for gene regulation in embryonic stem cells*. Cell. 1991b; 66:291-304.

Schultz W.A. and Hoffmann M.J., *Transcription Factor Networks in Embryonic Stem Cells and Testicular Cancer and the Definition of Epigenetics*. Epigenetics 2007; 2(1): 37:42.

Seo B.M., Miura M., Gronthos S., Bartold P.M., Batouli S., Brahim J. et al., *Investigation of multipotent postnatal stem cells from human periodontal ligament*. Lancet. 2004; 364: 149-155.

Shi S. and Gronthos S., *Perivascular niche of postnatal mesenchymal stem cells in human bone marrow and dental pulp*. J Bone Miner Res. 2003; 18: 696-704.

Smith A., *A glossary for stem cell biology*. Nature 2006; 441: 1060.

Sturm R.A., Das W. and Herr W., *The ubiquitous octamer binding protein Oct-1 contains a POU domain with a homeo box subdomain*. Genes & Development 1988; 2: 1582-1599.

Switzer R.C., Merril C.R. and Shifrin S., *A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels*. Anal Biochem. 1979; 98(1): 231-237.

Takeda J., Seino S. and Bell G.I., *Human Oct3 gene family: cDNA sequences, alternative splicing, gene organization, chromosomal location and expression at low levels in adult tissues*. Nucleic Acids Research 1992; 20(17): 4613-4620.

Taylor H.S., *The role of HOX genes in human implantation*. Hum Reprod Update. 2000; 6(1): 75-79.

Thomson A., Itskovitz-Eldor J., Shapiro S.S., Waknitz M.A., Swiergiel J.J., Marshall V.S. and Jones J.M., *Embryonic stem cell lines derived from Human blastocysts*. Science. 1998; 282: 1145-1147.

Tomioka M., Nishimoto M., Miyagi S., Katayanagi T., Fukui N., Niwa H., Muramatsu M. and Okuda A., *Identification of Sox2 regulatory region which is under the control of Oct3/4-Sox2 complex*. Nucleic Acids Res. 2002; 30: 3202-3213.

Verfaillie C.M., Pera M.F. and Lansdorp P.M., *Stem cell: hype and reality*. American Society of Hematology 2002; 369-391.

Verrijzer C.P, Alkema W.W., Van Weperen W.W., Van Leeuwen H.C., Strating M.J.J. and Van der Vliet P.C., *The DNA binding specificity of the bipartite POU domain and its subdomain*. The Embo Journal 1992; 11: 4993-5003.

Vigano MA and Staudt LM., *Transcriptional activation by Oct3: evidence for a role of the POU-specific domain in mediating functional interaction with Oct1*. Nucleic Acids Res. 1996; 24: 2112-2118.

Wagers A.J. and Weissman I.L., *Plasticity of adult stem cells*. Cell 2004; 116: 639-648.

Weaver RF., *Biologia molecolare*. Mc Graw Hill, New York, USA: 2005.

Weissman I.L., Anderson D.J. and Gage F., *Stem and progenitor cells:origins, phenotypes, lineage commitments and transdifferentiations*. Annu Rev Cell Dev Biol. 2001; 17:387-403.

Winship H. and Michele A.C., *The POU domain: versatility in trascriptional regulation by a flexible two-in-one DNA-binding domain*. Genes & Development 1995; 9:1679-1693.

Xu H., Wang W., Li C., Yu H., Yang A., Wang B., Jin Y. *WWP2 promotes degradation of transcription factor OCT4 in human embryonic stem cells*. Cell Research 2009; 19:561-573.

Yeom Y.I., Ha H.S., Balling R. et al., *Structure, expression and chromosomal location of the Oct4 gene*. Mech Dev. 1991; 35: 171-179.

Yeom Y.I., Fuhrmann G, Ovitt C.E. et al., *Germline regulatory element of Oct4 specific for the totipotent cycle of embryonal cells*. Development. 1996; 122: 881-894.

Young H.E., Duplaa C., Romero-Ramos M. et al., *Adult Reserve Stem Cells and Their Potential for Tissue Engineering*. Cell Biochem Biophys. 2004a; 40: 1-80.

Tesi di Dottorato di Renza Spelat discussa presso l'Università degli Studi di Udine

Young H.E. and Black A.C. JR., *Adult Stem Cells*. Anat Rec. 2004b; 276A: 75-102.

Yu J., Vodyanik M.A., Smuga-Otto K., Antosiewicz-Bourget J., Frane J.L., Tian S., Nie J., Jonsdottir G.A., Ruotti V., Stewart R., Slukvin I.I. and Thomson J.A., *Induced pluripotent stem cell lines derived from human somatic cells*. Science. 2007; 318: 1917-1920.

Yuan H., Corbi N., Basilico C. and Dailey L., *Developmental-specific activity of the FGF4 enhancer requires the synergistic action of Sox2 and Oct3*. Genes Dev. 1995; 9: 2635-2645.