

2. SCOPO E STRATEGIE ADOTTATE

Lo scopo del lavoro è la derivazione di fibroblasti bovini con modificazioni mirate, in modo da ottenere integrazioni di DNA esogeno in regioni specifiche del genoma. Questo è importante per ottenere livelli d'espressione più alti e prevedibili rispetto a quelli ottenuti con integrazioni casuali.

La scelta dei fibroblasti, e più in generale di cellule somatiche, è dettata dal fatto che non esistono ancora cellule embrionali staminali nel bovino, cellule necessarie cioè per realizzare la fusione in blastocisti allo scopo di ottenere animali transgenici chimere per la mutazione introdotta trasferendo le ES di partenza. Abbiamo scelto quindi di mettere a punto tali tecniche di modificazioni mirate ("gene-targeting") in cellule somatiche da utilizzare successivamente come donatrici per il trasferimento nucleare per ottenere eventualmente bovini transgenici.

L'approccio utilizzato per mutare i fibroblasti è quello della ricombinazione bifasica, suggerita da Kolb nel suo lavoro del 2001.

Il gene bersaglio della nostra ricerca è la β -lattoglobulina bovina (BLG).

La prima fase consiste essenzialmente in un passaggio di ricombinazione omologa per l'inserimento all'interno della BLG dei siti *loxP* e *lox2272*, siti di riconoscimento per la ricombinasi Cre del fago P1. A causa tuttavia della bassa frequenza dell'evento di HR nelle cellule somatiche, è necessaria una strategia di arricchimento. Abbiamo scelto la "positive-negative selection", in quanto la tecnica "promoterless", sebbene dia livelli maggiori di arricchimento, può essere utilizzata solo per geni attivamente trascritti nelle cellule usate. Come abbiamo già accennato sopra, il limite maggiore della strategia di selezione positivo-negativa è essenzialmente la bassa efficienza della cassetta di selezione negativa, di solito HSV-tk (herpes simplex virus-thymidine kinase): esso arricchisce gli eventi di HR in presenza di gancyclovir, che viene convertito da HSV-Tk in un nucleoside tossico che inibisce la crescita di cellule che hanno subito un evento d'integrazione casuale. Infatti questo schema di selezione con due agenti selettivi, per esempio G418 per la selezione positiva e gancyclovir per quella negativa, determina solo un parziale arricchimento dei ricombinanti omologhi e un effetto tossico generale sulla popolazione cellulare.

A questo proposito, abbiamo focalizzato la nostra attenzione sulla ricerca di un marcatore ottimale come selettore negativo situato all'esterno della regione di omologia per la β -lattoglobulina. I costrutti allestiti presentano, come selettori negativi, una cassetta di espressione per la p53, proteina chiave nel processo apoptotico, una per siRNA, piccole molecole di RNA interferente, e due tipi differenti per la EGFP (enhanced green fluorescent protein). Tali selettori negativi vengono persi nel caso di avvenuta ricombinazione omologa.

Il nostro lavoro prevede inoltre, una volta inseriti nella β -lattoglobulina bovina i siti *lox* mediante HR, una seconda fase secondo il modello di Kolb. I fibroblasti bovini saranno cotrasfettati con un plasmide che contiene il transgene fiancheggiato dai siti *loxP* e *lox2272* e con un altro che esprime la ricombinasi Cre. Ciò dovrebbe permetterci di inserire un gene esogeno esattamente nel locus della β -lattoglobulina (RMCE: Cre Recombinase-Mediated Cassette Exchange): in questo modo l'unica sequenza che viene inserita è il transgene, senza porzioni plasmidiche e ciò dovrebbe garantirci livelli di espressione prevedibili.

2.1 Proteina p53

La proteina p53 è considerata “il guardiano del genoma”, proteggendo il DNA: essa infatti è coinvolta nei processi di regolazione del ciclo cellulare e dell'apoptosi. Il gene p53 è un gene soppressore di tumori e agisce come fattore trascrizionale. È localizzato sul braccio corto del cromosoma 17 umano. La maggior parte delle attività di p53 sono state studiate mediante topi ottenuti da cellule embrionali staminali con il corrispondente gene inattivato mediante ricombinazione omologa (knock-out). Questi topi mostrano un'alta incidenza di tumori, sebbene lo sviluppo risulti normale. Timociti (linfociti che derivano dal timo, precursori dei linfociti T) prelevati da questi topi sono completamente resistenti a fenomeni di apoptosi indotti mediante radiazioni ionizzanti. Tra le funzioni di p53 vi è l'arresto del ciclo cellulare nel passaggio G1-G2 in risposta a danni del DNA, a seguito del quale inizia il processo di apoptosi qualora tale danno risulti irreparabile o dopo segnali proliferativi anomali. La p53 umana contiene 393 aminoacidi e funziona come un tetramero. Il polipeptide possiede tre domini strutturali e funzionali: dominio di transattivazione amino-terminale, quello di legame al DNA o “core” e quello di regolazione carbossi-terminale. La regione N-terminale è in grado di legare componenti del complesso trascrizionale come le proteine di legame ai TATA box (TBP), ma anche l'oncoproteina Mdm2; il core permette alla p53 di legarsi al DNA in maniera sequenza-specifica; la regione C-terminale aiuta la p53 a legarsi alle sequenze di DNA mediante il suo core. In condizioni normali, la p53 è espressa a basso livello all'interno della cellula, grazie alla presenza di un'altra proteina (mdm2) che si lega ad essa, inducendone la degradazione attraverso una proteolisi ubiquitina-dipendente o regolandone l'esportazione nucleare. In questo modo la p53 attiva direttamente l'espressione del suo regolatore, innescando un potente processo di “feedback” negativo. In caso di danno al DNA, le proteine attivate da processi stressori proteggono la p53 dalla degradazione (fosforilazione/defosforilazione) e attivano la sua funzione come fattore di trascrizione.

Abbiamo dunque ipotizzato che la sovraespressione di p53 potesse essere utile come selettore negativo per l'arricchimento cellulare per il “gene-targeting”, innescando processi apoptotici nelle cellule che avessero subito un evento di ricombinazione casuale e non omologa.

2.2 Short Interfering RNAs (siRNA)

In questi ultimi anni è stato dimostrato che piccole molecole di RNA a doppio filamento, chiamate siRNA (Short Interfering RNAs), possono interferire con l'espressione di geni altamente omologhi. Data l'efficacia e la specificità del meccanismo d'interferenza mediato da RNA, abbiamo realizzato vettori che portassero come selettore negativo una cassetta di espressione codificante siRNA in grado di catalizzare la degradazione del trascritto del gene di fusione selettore positivo HyTk (resistenza all'igromicina/HSV-thymidine kinase) o di interferire con il processo di traduzione dello stesso. In questo modo, cellule rese resistenti all'igromicina grazie all'integrazione nel genoma del gene HyTk, ritornano ad essere sensibili all'antibiotico a causa della presenza del siRNA, nel caso in cui non sia avvenuto l'evento di ricombinazione omologa. Da ciò ne dovrebbe conseguire un arricchimento in cloni ricombinanti omologhi.

In questi ultimi anni molti gruppi di ricerca si sono dedicati allo studio e allo sviluppo di metodiche di silenziamento genico basate sulla tecnica dell'RNA interferente (RNAi). Si tratta di un meccanismo di silenziamento genico post-trascrizionale sequenza specifico mediato dall'introduzione di RNA a doppio filamento (dsRNA) con sequenza omologa a quella del gene da silenziare. L'RNA interferente in natura potrebbe essere un'arma contro le infezioni virali oppure essere coinvolto nella degradazione di dsRNA che derivano dalla trascrizione di trasposoni o di sequenze ripetute.

Il fenomeno dell'RNAi fu scoperto per la prima volta nelle piante (de Lange et al., 1995), ma il termine fu coniato durante gli studi successivi in *C. elegans*.

Nel 1998 Fire e colleghi scoprirono che l'iniezione di RNA a doppio filamento in *C. elegans* portava ad un efficiente silenziamento genico specifico per una sequenza bersaglio desiderata (Fire et al., 1998). Essi dimostrarono che l'introduzione in *C. elegans* di basse concentrazioni di RNA "senso" e "antisenso" che si appaiavano formando piccole molecole di RNA a doppio filamento causava l'attivazione di meccanismi cellulari di soppressione dell'espressione genica più efficiente rispetto a quella generata dall'RNA a singolo filamento ("senso" o "antisenso").

Questo meccanismo fu poi scoperto in diverse specie tra cui funghi, protozoi, insetti (*Drosophila melanogaster*) (Tuschl et al., 1999) e vertebrati. In *Drosophila* si osservò che molecole di dsRNA di circa 70 paia di basi venivano tagliate in piccole molecole di dsRNA di circa 22 nucleotidi, definiti successivamente siRNA (short interfering RNAs), e che l'introduzione di tali molecole sintetizzate chimicamente negli estratti di embrioni portava ad una degradazione parziale dell'RNA omologo. Il silenziamento mediato da RNA è dunque un meccanismo di regolazione genica che si è conservato durante l'evoluzione con molte varianti specie-specifiche. Inoltre il fatto che non sia stato trovato negli organismi procarioti indica che si tratti di un'innovazione da parte dei sistemi eucarioti.

Il dsRNA è riconosciuto dall'enzima Dicer, che fa parte della famiglia delle ribonucleasi Rnasi III. I Dicer si sono conservati durante l'evoluzione e sono stati trovati in *D. melanogaster*, in *C. elegans*, nella pianta del tabacco e nei mammiferi. L'uomo e *C. elegans* hanno un solo Dicer, *Drosophila* due e *Arabidopsis* quattro. Il Dicer contiene diversi domini: un dominio N-terminale con attività elicastica (il cosiddetto dominio PAZ coinvolto nel controllo dello sviluppo), un dominio catalitico con due regioni omologhe a quelle delle proteine della famiglia delle Rnasi III e un dominio di legame per il dsRNA a livello della porzione C-terminale. L'attività del

Dicer è ATP dipendente: da una parte si pensa che sia legata all'attività ATPasica dell'elicasi, dall'altra sembra che l'ATP possa controllare il legame tra Dicer e RNA a doppio filamento o regolare l'attività delle Rnasi.

Il Dicer processa l'RNA a doppio filamento in siRNA di 21-25 nucleotidi a seconda della specie. I siRNA possiedono un'estremità 3' idrossilica con 2 nucleotidi non appaiati e un'estremità 5' fosforilata. Tali caratteristiche sono importanti per il meccanismo dell'interferenza mediata da RNA. I siRNA prodotti dal Dicer vengono incorporati in un complesso di nucleasi, detto RISC (RNA-induced silencing complex) che deve essere convertito da una forma latente, contenente il siRNA a doppio filamento, ad una forma attiva svolgendo il siRNA tramite un'elicasi con un processo ATP-dipendente (RNA elicasi ATP-dipendente). Il RISC contiene anche un'endonucleasi che, utilizzando la sequenza codificata dal filamento di siRNA antisense, trova e rompe la sequenza complementare di mRNA: in particolare essa catalizza l'idrolisi del legame fosfodiesterico, lasciando un'estremità 5' fosfato e un'estremità 3' idrossilica. Il sito di taglio si trova nel mezzo della regione complementare, dieci nucleotidi a monte dell'estremità 5' del siRNA. La reazione richiede ioni magnesio e non è ATP dipendente, mentre l'assemblaggio del RISC con i siRNA necessita di ATP. Nei RISC sono inoltre stati identificati i DEAD-box tipici delle RNA-elicasi: sembra infatti che la presenza di un'elicasi in tale complesso favorisca il rilascio dell'mRNA e del filamento singolo di siRNA che può successivamente essere riciclato. Il RISC contiene anche una proteina della famiglia delle Argonate (Ago). Le proteine Ago presentano due domini conservati: il PAZ (piwi-argonate-zwille) e il PIWI, che probabilmente partecipano all'interazione col Dicer. Il dominio PAZ forse funziona come modulo di legame per Ago, riconoscendo e legando i due caratteristici nucleotidi non appaiati dell'estremità 3' dei siRNA. Questo riconoscimento del 3' terminale è importante per i Dicer e per le Ago per distinguere i siRNA dalle altre molecole di RNA.

L'mRNA bersaglio viene quindi tagliato in frammenti di circa 22 nucleotidi. Quando il taglio è completato, il RISC si allontana e il siRNA può essere utilizzato per un nuovo ciclo di taglio e riconoscimento dell'mRNA. Tale processo avviene nel citoplasma.

L'RNA a doppio filamento viene sintetizzato a livello intracellulare, ma può anche essere introdotto dall'esterno direttamente nelle cellule. Scoperto questo fenomeno, si è pensato di sfruttarlo per il silenziamento genico. La lunghezza minima dell'RNA a doppio filamento necessaria per indurre il meccanismo dell'RNAi è di 26 nucleotidi, ma molecole più lunghe risultano essere più efficienti. Tuttavia la trasfezione di lunghi RNA a doppio filamento in colture cellulari di mammifero induce un potente sistema antivirale, in grado di generare una risposta da parte dell'interferone: esso causa l'attivazione di enzimi che degradano in modo aspecifico l'mRNA intracellulare e inibiscono in modo globale la sua traduzione (Harborth et al., 2001). L'interferone in particolare induce un enzima, la 2'-5' Oligoadenilato Sintasi, il cui prodotto è un essenziale cofattore per una ribonucleasi sequenza-non-specifica, l'Rnasi L (enzima che presenta specificità per RNA di specie diverse incluso l'RNA ribosomale). Inoltre l'interferone induce la PKR, una protein-chinasi RNA dipendente, che se attivata da lunghe molecole di dsRNA, fosforila e inattiva la subunità α del fattore eIF-2 α (eukaryotic initiation factor 2 α) implicato nell'inizio della traduzione dell'RNA messaggero, sopprimendo così la sintesi proteica e portando alla morte cellulare con meccanismi apoptotici e non.

È stato necessario quindi escogitare sistemi per superare i meccanismi di difesa tipici dei mammiferi per introdurre molecole esogene di RNA. La sintesi chimica di siRNA è quello più utilizzato per generare RNAi. Anche i siRNA trascritti *in vitro* dal promotore fagico T7 (siRNA-based hairpin RNA) (Donzé et al., 2002), così come i siRNA isolati da estratti proteici di embrioni di *D. melanogaster* sembrano essere efficaci. Tra le maggiori difficoltà vi è quella di introdurre le piccole molecole di RNA a doppio filamento direttamente nelle cellule. Un'altra possibilità è la produzione di siRNA a partire da lunghe molecole di RNA che vengono tagliate dall'Rnasi III di *E. coli* (esiRNA – endoribonuclease-prepared siRNA) (Yang et al., 2002) e che possono riconoscere più siti all'interno dello stesso mRNA. Clarke inoltre ha inattivato gli enzimi che mediano il sistema di difesa antivirale indotto dall'interferone (Clarke et al., 1995).

Tuttavia in questi casi non si ottengono organismi “knock-out” stabili: per risolvere questo problema si sono creati nuovi tipi di vettori di espressione. Questi vettori presentano inserti codificanti per la sequenza bersaglio del gene d'interesse posti sotto il controllo di promotori di RNA polimerasi III (il promotore murino U6 e l'umano H1 - che normalmente trascrivono gli small nuclear RNA - e quello per il tRNA) che permettono la trascrizione di siRNA funzionali o di loro precursori.

Zheng e colleghi hanno sviluppato un sistema di espressione per siRNA con doppio promotore (pDual) in cui il DNA sintetico codificante una sequenza di siRNA genespecifica è inserita tra i due diversi promotori di polimerasi III, U6 e H1. Dopo la trasfezione in cellule di mammifero, i filamenti senso e antisense vengono trascritti da questi due promotori a partire dal medesimo template: si crea così un RNA a doppio filamento con un'uridina non appaiata in ciascuna estremità 3', cioè con una struttura simile al siRNA generato dal Dicer (Zheng et al., 2004). Un altro sistema prevede invece l'utilizzo di un solo promotore davanti all'inserto oligonucleotidico costituito da due sequenze di circa 21 paia di basi, di cui una invertita, separate da una sequenza spaziatrice di 6-9 paia di basi: in questo caso il trascritto va a formare una struttura a forcina (short hairpin RNA – shRNA) che verrà convertita in siRNA vero e proprio dal Dicer (Hutvagner et al., 2002). Un'altra categoria di vettori di espressione è rappresentata dai vettori virali: i più utilizzati sono i retrovirus, i vettori retrovirali oncogeni basati sul virus della leucemia murina di Moloney (MoMuLV), il virus delle cellule staminali murine (MSCV) e i vettori lentivirali derivati dal virus dell'immunodeficienza umana tipo-1 (HIV-1).

Sono stati studiati anche vettori che contengono promotori di polimerasi II per esprimere lunghi dsRNA. Dato che sono disponibili diversi promotori della polimerasi II tessuto-specifici, questo metodo permette un silenziamento genico mirato a particolari tessuti.

Le possibili applicazioni dell'RNAi sono svariate: dalla lotta contro malattie, quali il cancro e le infezioni da virus e parassiti, all'analisi di problemi legati alla biologia cellulare e dello sviluppo. Esso inoltre può essere usato per lo studio di funzionalità e interazione tra geni. L'RNAi risulta importante anche per l'identificazione di nuovi bersagli per i farmaci. I siRNA potrebbero infine essere utili in varie forme di terapia genica.

In conclusione l'importanza dei siRNA è essenzialmente data dal fatto che si tratta di prodotti cellulari naturali, non producono metaboliti tossici, hanno una lunga durata nelle colture cellulari e sono efficaci anche a basse concentrazioni. In più, a differenza dell'effetto non specifico dei lunghi dsRNA, i siRNA generano un silenziamento genico selettivo nei mammiferi.

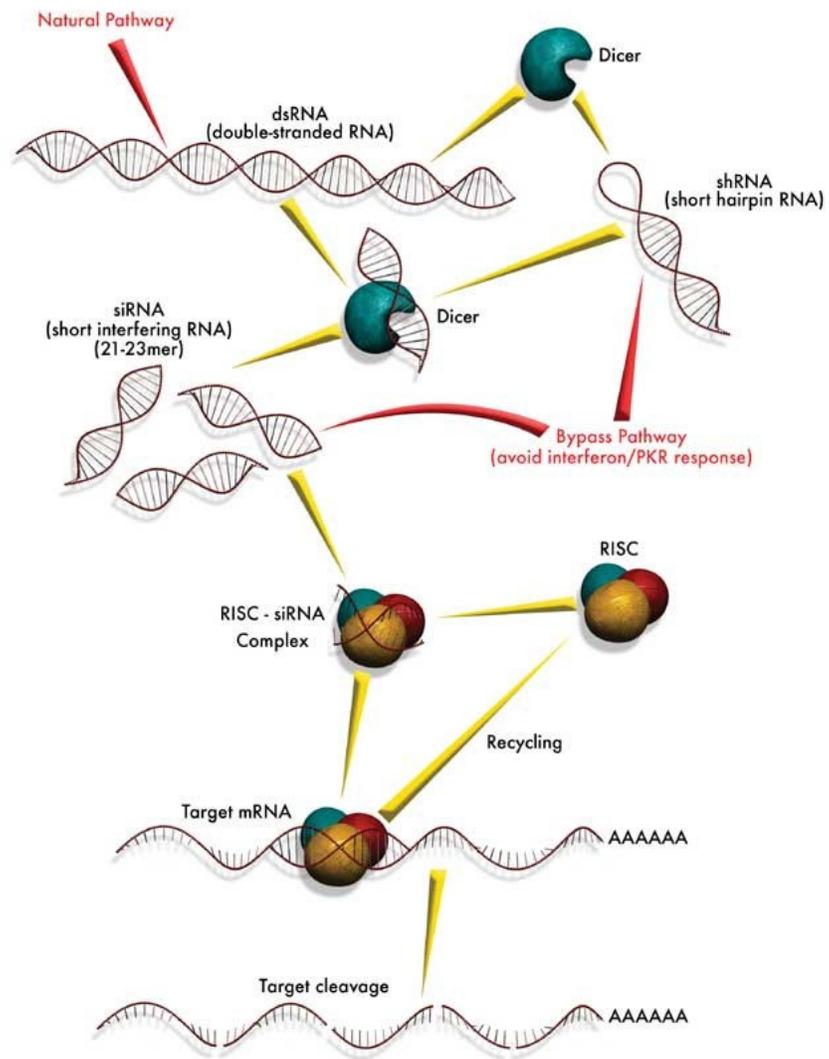


Fig. 1.14
Rappresentazione schematica del meccanismo d'interferenza mediato da RNA.

2.3 Green fluorescent protein (GFP)

Tra i numerosi esseri viventi bioluminescenti vi è la medusa *Aequorea victoria*. Questo organismo del nord Atlantico è in grado di emettere una luce verde fluorescente dai fotociti localizzati alle estremità del suo ombrello. La fluorescenza è generata dall'attivazione sequenziale di due fotoproteine, l'aequorina e la proteina verde fluorescente (GFP). La GFP è una proteina estremamente stabile di 238 amminoacidi. In seguito a legame con calcio, l'aequorina emette una luce blu che eccita la GFP, la quale a sua volta emette fluorescenza verde. Sebbene le caratteristiche di questa proteina fossero conosciute da anni, è stato solo con il clonaggio del cDNA della GFP nel 1992 (Prasher et al., 1992) e la sua successiva espressione eterologa in *E. coli* e *C. elegans* (Chalfie et al., 1994) che i ricercatori si sono resi conto delle reali potenzialità di questa molecola.

La GFP eccitata da luce blu o UV e in presenza di ossigeno, emette fluorescenza verde e non necessita di ulteriori substrati o cofattori esogeni, in quanto il cromoforo è intrinseco alla GFP stessa (Chalfie et al., 1994). L'immediata visualizzazione di questa proteina in colture cellulari è utilizzata come reporter per l'espressione genica, come marcatore per il differenziamento durante le varie fasi di sviluppo, come coda per la localizzazione di proteine (Wang e Hazelrigg, 1994) e potrebbe essere utile per l'arricchimento cellulare. In più la fluorescenza della GFP è specie-indipendente e può essere monitorata in modo non invasivo utilizzando tecniche di microscopia a fluorescenza, citofluorimetria e visualizzazione macroscopica. Essa rappresenta uno strumento importante per la misurazione diretta dell'efficienza di trasfezione. Tuttavia le GFP wild-type presentano varie caratteristiche indesiderate, compresa una bassa intensità di fluorescenza quando eccitata da luce blu, un ritardo nello sviluppo della fluorescenza in seguito alla sintesi proteica e una bassa espressione in vari tipi di cellule di mammifero. In questo modo, la sensibilità della GFP wild-type risulta inferiore a quella di altre proteine reporter, come β -gal, che utilizza un'amplificazione di tipo enzimatico. Per aumentare la rilevazione della GFP in cellule di mammifero trasfettate, sono state create diverse varianti, che contengono mutazioni nel cromoforo (EGFP – enhanced green fluorescent protein).

Nel nostro lavoro la prima cassetta d'espressione per la GFP utilizzata deriva dal plasmide pBSCGFPns, che presenta la sequenza codificante sotto il controllo del promotore del cytomegalovirus (CMV); la seconda invece proviene dal plasmide pCX-EGFP e ha il cDNA della EGFP sotto il controllo del promotore della β -actina di pollo, l'"enhancer" del CMV, l'introne della β -actina e il segnale di poliadenilazione della β -globina bovina (Okabe et al., 1997). Utilizzando la GFP come selettore negativo nei vettori di ricombinazione, ci aspettiamo che risultino verdi le cellule che non sono andate incontro ad HR, in modo così da ridurre il numero di cloni da analizzare per l'evento di ricombinazione omologa.