

UNIVERSITÀ DEGLI STUDI DI NAPOLI
" FEDERICO II "



DOTTORATO DI RICERCA IN
GENETICA
E
MEDICINA MOLECOLARE
XVIII CICLO

**Caratterizzazione molecolare del gene *minifly* di
*Drosophila melanogaster***

Tutore

Ch.ma Prof.ssa
Maria furia

Candidato

Dr.ssa
Riccardo Sara

Coordinatore:

Ch.mo Prof.
Carmelo Bruno Bruni

Anno Accademico 2005/2006

Indice

Indice

Premessa
..... pag.1

Introduzione
.....pag.6

Risultati
.....pag.33

Discussione
.....pag.107

Materiali e Metodi
.....pag.118

Bibliografia
.....pag.137

Premessa

PREMESSA

Da quando il gruppo di Thomas Hunt Morgan utilizzò la *Drosophila melanogaster* per verificare la teoria cromosomica dell'eredità quest'organismo, comunemente noto come moscerino della frutta, ha rappresentato di fatto una guida nel campo degli studi genetici, diventando uno dei sistemi modello più utilizzati.

All'inizio la *Drosophila* fu probabilmente preferita per motivi di carattere essenzialmente pratico, come la notevole facilità di allevamento, il breve tempo di generazione e la progenie molto numerosa. In seguito, però, quest'organismo mostrò ulteriori caratteristiche che risultarono estremamente vantaggiose per le analisi genetiche, quali l'assenza di ricombinazione meiotica nei maschi, la presenza di solo quattro cromosomi e la possibilità di visualizzare facilmente i cromosomi politenici giganti nelle ghiandole salivari ed altri tessuti delle larve.

Per quanto riguarda gli studi dei meccanismi di sviluppo, ulteriori vantaggi furono forniti dall'osservazione che l'esoscheletro, sia nelle larve che negli adulti, fornisce un'ampia varietà di strutture che possono diventare bersaglio di mutazioni, così che molti fenotipi mutanti potevano essere facilmente individuati attraverso una semplice osservazione allo stereomicroscopio. Anche la struttura complessa dell'occhio della *Drosophila*, con i suoi 800 ommatidi, ha fornito un valido sistema sperimentale particolarmente adatto allo studio di mutazioni "da guadagno di funzione" che agiscono da dominanti negative.

Questi vantaggi sono stati rafforzati nel tempo dai risultati ottenuti da generazioni di *Drosophilisti* che hanno sviluppato non solo un'ampia

collezione di mutanti, ma anche un repertorio di metodiche che rende la *Drosophila* uno degli organismi multicellulari per cui è oggi disponibile il più ampio spettro di sofisticate tecniche di transgenesi e di manipolazione genetica. Pochi organismi modello sono riusciti a mantenere per così lungo tempo una posizione di supremazia nello studio dei meccanismi che regolano i processi biologici essenziali, e ciò è stato in gran parte dovuto alla possibilità di effettuare accurate ed esaurienti analisi *in vivo*. Un ulteriore fattore che ha contribuito al successo ed alla diffusione del moscerino della frutta come sistema modello, e che da sempre rappresenta un punto di forza e orgoglio per tutta la comunità dei *Drosophilisti*, è la libera circolazione delle informazioni, la produzione di materiale scientifico su larga scala e la condivisione delle risorse. Fin dall'inizio i *Drosophilisti* stabilirono immediatamente il "principio di apertura" come mezzo utile per un rapido progresso delle conoscenze ed il frutto di questa tradizione è ancora oggi evidente nella ricchezza di materiale e di informazioni messi gratuitamente a servizio della comunità scientifica. L'avvento dell'era genomica, con il sequenziamento dei genomi di molti organismi, ha fortemente accelerato i processi di conoscenza fornendo un impressionante numero di dati e di informazioni a cui è possibile accedere facilmente. Questi dati oggi promuovono una visione più dinamica e globale dei processi biologici, consentendo analisi su scala genomica. Oggi l'insieme di questa enorme mole di dati genetici e molecolari è raccolta nella più ampia e dettagliata banca dati biologica disponibile in rete: la Flybase (<http://flybase.bio.indiana.edu/>). La possibilità di effettuare in maniera quasi immediata analisi comparative tra genomi di diversi organismi ha reso ancora

più evidenti le similitudini esistenti tra vertebrati ed invertebrati, rivelando un notevole grado di conservazione non solo tra i prodotti proteici ma anche tra i processi biologici fondamentali. L'elevato grado di conservazione dei meccanismi molecolari che sono alla base dei processi di sviluppo e differenziamento, nonché del comportamento e dell'apprendimento, conferma la *Drosophila melanogaster*, con il suo genoma di dimensioni venti volte inferiore a quello dei mammiferi, un valido sistema sperimentale per lo studio dei pathway cellulari che portano allo sviluppo dei tessuti e degli organi. L'elevato grado di omologia spesso esistente tra geni di *Drosophila* e di mammifero ha fornito anche l'opportunità di analizzare la conservazione evolutiva dei geni responsabili delle patologie umane di origine genetica. Sorprendentemente, quasi un terzo di tutti i geni che provocano malattie umane note hanno mostrato di possedere una controparte nel genoma di *Drosophila*, così che alcuni di questi sono già stati caratterizzati con successo a livello funzionale proprio partendo da quest'organismo (per una review vedi Bier, 2005). La *Drosophila* continua quindi ad essere una fonte d'informazione utile anche per la biologia umana ed è ragionevole prevedere che gli studi condotti in quest'organismo possano contribuire significativamente alla comprensione delle malattie umane a base genetica. Non è retorica dunque affermare che, per certi aspetti, le ricerche sulla *Drosophila melanogaster* possono essere considerate oggi rilevanti per la biologia degli organismi superiori quanto lo furono ai tempi di Morgan per dimostrare la teoria cromosomica dell'ereditarietà.

La mia tesi di Dottorato si inserisce in un progetto di ricerca più ampio che mira a chiarire i meccanismi molecolari che sono alla base dell'insorgenza di una patologia umana nota come Discheratosi congenita (DC) X-linked, utilizzando la *Drosophila melanogaster* come organismo modello.

La Discheratosi è una rara malattia genetica che si manifesta con una frequenza di 1/1.000.000 ed è caratterizzata da un ampio spettro di manifestazioni cliniche, tra cui ritardo mentale, vita media breve, alterazioni della pelle e delle mucose, danni del midollo osseo e lesioni displastiche della bocca o del tratto gastro-intestinale che talvolta vanno incontro a trasformazioni maligne.

L'insufficienza midollare è presente nell'80% dei pazienti e rappresenta la principale causa di morte. La Discheratosi X-linked è causata da mutazioni del gene *DKCI* (Heiss et al., 1998; Knight et al, 1999), un gene che appartiene ad una famiglia altamente conservata dal lievito all'uomo ed i cui membri codificano proteine ubiquitarie localizzate prevalentemente nel nucleolo (Heiss et al., 1998). L'ortologo di *Drosophila* del gene *DKCI* il gene *minifly* (*mfl*), è altamente correlato al gene umano sia a livello strutturale che a livello funzionale, così che i mutanti da perdita di funzione del gene *mfl* possono rappresentare un buon sistema modello per lo studio di questa malattia genetica legata all'X. I geni *mfl* e *DKCI* appartengono infatti ad una famiglia che è altamente conservata dagli Archaea ai mammiferi, i cui membri codificano proteine coinvolte in numerosi processi essenziali per tutti i tipi di cellule. Questi processi includono la sintesi e la modificazione dello RNA ribosomale (rRNA) e, nei mammiferi, anche il mantenimento delle estremità telomeriche.

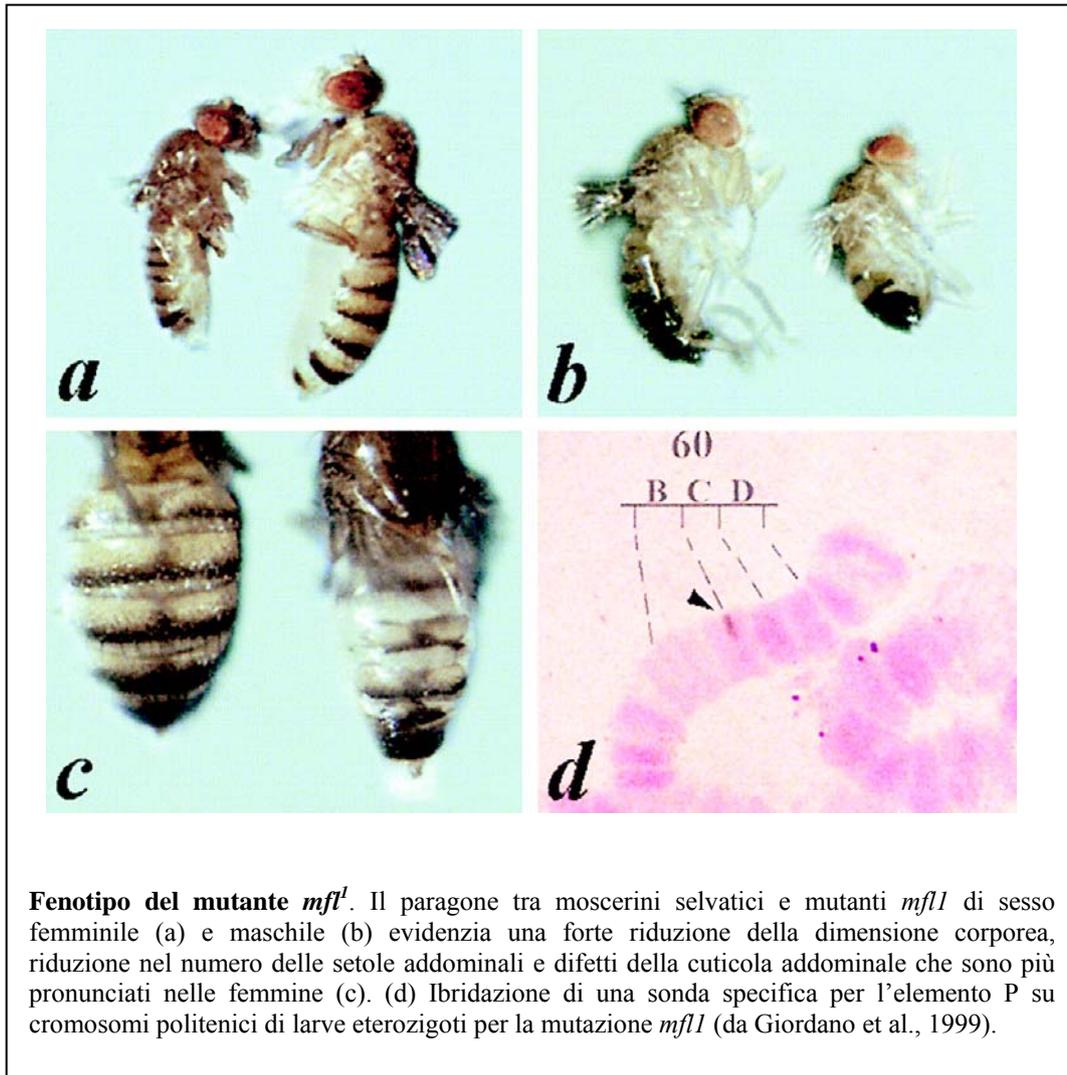
Un'ulteriore funzione di queste proteine, denominate MFL in *Drosophila*, Cbf5p in lievito, Nap57 in ratto e discherina in topo e uomo, è che esse costituiscono uno dei quattro componenti fondamentali dei complessi ribonucleoproteici che si associano specificamente ai piccoli RNA nucleolari (snoRNA) della classe H/ACA, formando i complessi snoRNP che dirigono i processi di maturazione e pseudouridilazione degli RNA bersaglio. All'interno di questi complessi, le proteine della famiglia MFL/discherina agiscono da pseudo-uridina sintetasi, mentre le molecole di snoRNA di tipo H/ACA selezionano gli specifici residui che verranno modificati (vedi Bachellerie et al., 2002).

Nel corso del mio dottorato mi sono essenzialmente occupata della caratterizzazione molecolare del gene *mfl*, analizzandone in dettaglio l'organizzazione e l'attività trascrizionale nel tentativo di chiarire le complesse funzioni che sono state attribuite a questa famiglia genica e fornire informazioni utili per la comprensione dei meccanismi molecolari che sono alla base della discheratosi congenita X-linked.

Introduzione

Il gene *minifly* di *Drosophila*

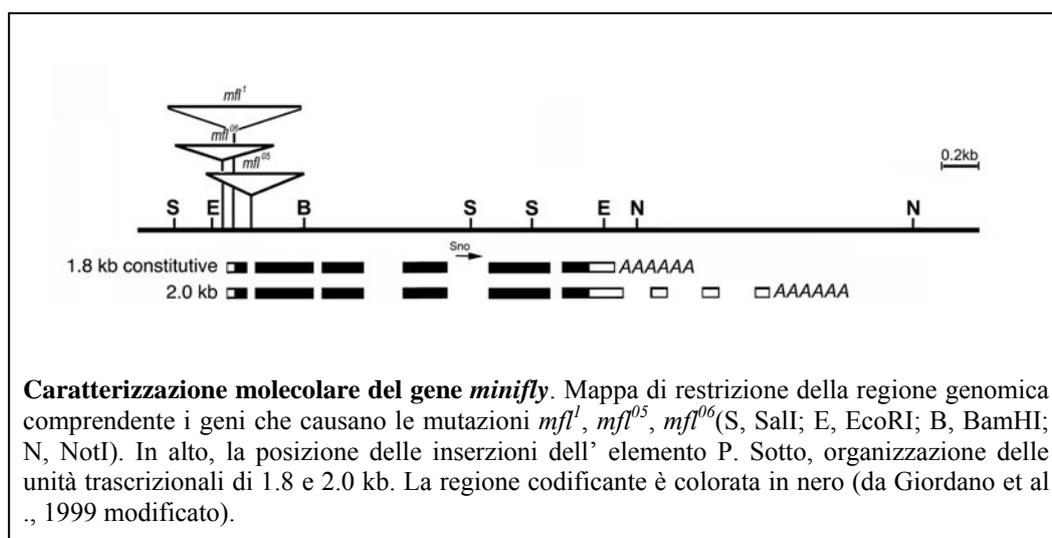
In *Drosophila melanogaster* il gene *mfl*, detto anche *Nop60B* (Philips et al., 1998) mappa sul secondo cromosoma nel braccio R, al limite delle bande politeniche 60B-60C. La caratterizzazione di mutazioni indotte dall'inserzione dell'elemento trasponibile P ha indicato che *mfl* è essenziale per la vita e la fertilità della *Drosophila*. Queste mutazioni, ottenute nel nostro laboratorio alcuni anni fa, sono tutte caratterizzate dalla presenza di un'unica copia dell'elemento P inserito a monte della sequenza codificante, ed è stato osservato che quest'inserzione provoca una significativa riduzione dell'espressione del gene. Successive analisi funzionali evidenziarono che i livelli di trascrizione di *mfl* erano critici, così un effetto dose può controllare la gravità del fenotipo mutante (Giordano et al. 1999). Mutazioni da forte o totale perdita di funzione causavano infatti letalità larvale, mentre mutazioni da parziale perdita di funzione provocavano un fenotipo ipomorfo, caratterizzato da difetti pleiotropici che includevano un'estrema riduzione della dimensione corporea, ritardo dello sviluppo (essenzialmente dovuto al prolungamento della vita larvale di 4-5 giorni), difetti nella cuticola addominale, forte riduzione della lunghezza e dello spessore delle setole addominali e ridotta fertilità femminile (Giordano et al., 1999). Le principali caratteristiche fenotipiche del mutante ipomorfo *mfl*^l risultarono essenzialmente simili a quelle causate dalle mutazioni *Minute*, *mini* o *bobbed* della *Drosophila* che sono noti alterare rispettivamente, la sintesi delle proteine ribosomali, degli rRNA 5S e degli rRNA 18S e 28S. Questa omologia fornì il primo indizio che il gene *mfl* potesse essere coinvolto nella biogenesi dei ribosomi.



Successivi esperimenti confermarono che i mutanti presentavano alterazioni nel meccanismo di maturazione del pre-rRNA, provocando una significativa riduzione degli rRNA maturi 18S e 28S ed un accumulo di molecole di pre-rRNA non maturate (Giordano et al., 1999).

Una caratteristica distintiva del mutante ipomorfo *mfl¹* risultò essere una ridotta fertilità femminile, correlata ad alterazioni morfologiche delle camere ovariche. Colorazioni con il fluorocromo DAPI, che lega specificamente il DNA, evidenziarono che nelle femmine mutanti il 10% delle camere ovariche

presentava un tipico aspetto degenerativo, caratterizzato dalla presenza di nuclei picnotici, condensati ed irregolari e quindi essenzialmente ascrivibile a fenomeni apoptotici. Successivamente, il coinvolgimento dell'apoptosi nella definizione del fenotipo mutante fu confermato mediante RNAi *in vivo*, attivando l'espressione di transgeni capaci di silenziare specificamente il gene *mfl* in tessuti specifici. Con questo approccio fu dimostrato che l'inattivazione del gene era sufficiente per indurre la morte cellulare localizzata (Giordano et al., 2001). Tuttavia, la sterilità femminile causata dall'allele *mfl^l* è attribuibile in parte anche al mancato sviluppo degli embrioni. Infatti, circa il 15% degli embrioni deposti vanno incontro a degenerazione, indicando che la mutazione mostra un effetto materno. Una prima caratterizzazione molecolare del gene *mfl* rivelò che esso produceva due trascritti, lunghi rispettivamente 1.8 e 2.0 kb. Mentre la specie di 1.8 kb era espressa costitutivamente durante il ciclo vitale, quella di 2.0 kb fu osservata solo nelle preparazioni di RNA ottenute da femmine adulte e da embrioni (Giordano et al., 1999).



I due trascritti erano entrambi presenti in embrioni precoci (0-2h), ma mentre l'mRNA zigotico di 1.8 kb persisteva negli stadi successivi, quello di 2.0 kb era presente solo nei primi stadi dell'embriogenesi (0-6h), mostrando il profilo di espressione tipico degli mRNA di derivazione materna. I due trascritti, pur mostrando un differente profilo di espressione, condividevano gran parte della loro sequenza e differivano esclusivamente nella loro regione 3' non tradotta (3' UTR) che nel trascritto di 2.0 kb presentava due esoni addizionali. In esperimenti di ibridazione *in situ* su camere ovariche, l'uso di sonde derivate dalla regione 3' UTR specifica per il trascritto da 2.0 kb dimostrarono con chiarezza che questo trascritto era prodotto durante l'oogenesi e trasmesso per via materna all'oocita, per essere poi ereditato dall'embrione maturo. Nel loro insieme, questi dati suggerirono che il trascritto materno di 2.0 kb suppliva all'assenza del trascritto zigotico durante le prime fasi dello sviluppo embrionale, finchè la sua sintesi non veniva attivata. Questo tipo di situazione è molto frequente per i trascritti che svolgono un ruolo essenziale nei processi di sviluppo e differenziamento, la cui carenza si manifesta spesso con sterilità ad effetto materno. Come già detto, i due trascritti da 1.8 e 2.0 kb contengono la stessa ORF (Open Reading Frame) che è capace di codificare per la sintesi di una proteina di 508 amminoacidi, con una massa molecolare di 56 kD. La proteina MFL risulta ubiquitaria; la sua localizzazione subcellulare è essenzialmente nucleolare, anche se occasionalmente essa è stata osservata diffondere nel citoplasma (Giordano et al.,1999). Questo dato farebbe supporre che la proteina, oltre a svolgere il ruolo di pseudouridina sintasi all'interno dei

complessi snoRNP di tipo H/ACA, possa essere implicata in processi di “shuttling” fra nucleo e citoplasma. Un aspetto descritto per la proteina ortologa Nap57 di ratto (Meier and Blobel, 1994). Come molti geni che codificano proteine coinvolte nella sintesi, struttura o funzione dell'apparato traduzionale, il gene *mfl* ospita nel suo quarto introne un gene che codifica un piccolo RNA funzionale. La caratterizzazione di questo trascritto permise di stabilire che si trattava di un piccolo RNA nucleolare (snoRNA) della classe H/ACA. Questo snoRNA è capace di dirigere la pseudouridilazione del residuo U1820 sull' rRNA 18S di *Drosophila*, un residuo equivalente allo U1689 dell' rRNA umano, la cui modificazione è diretta dallo snoRNA U70. Essendo il primo esempio di snoRNA della classe H/ACA identificato in *Drosophila*, lo snoRNA derivato dal IV introne del gene *mfl* fu denominato snoRNA H1.

Correlazioni strutturali e funzionali fra il gene *DKC1* dei mammiferi ed il gene *mfl* di *Drosophila*

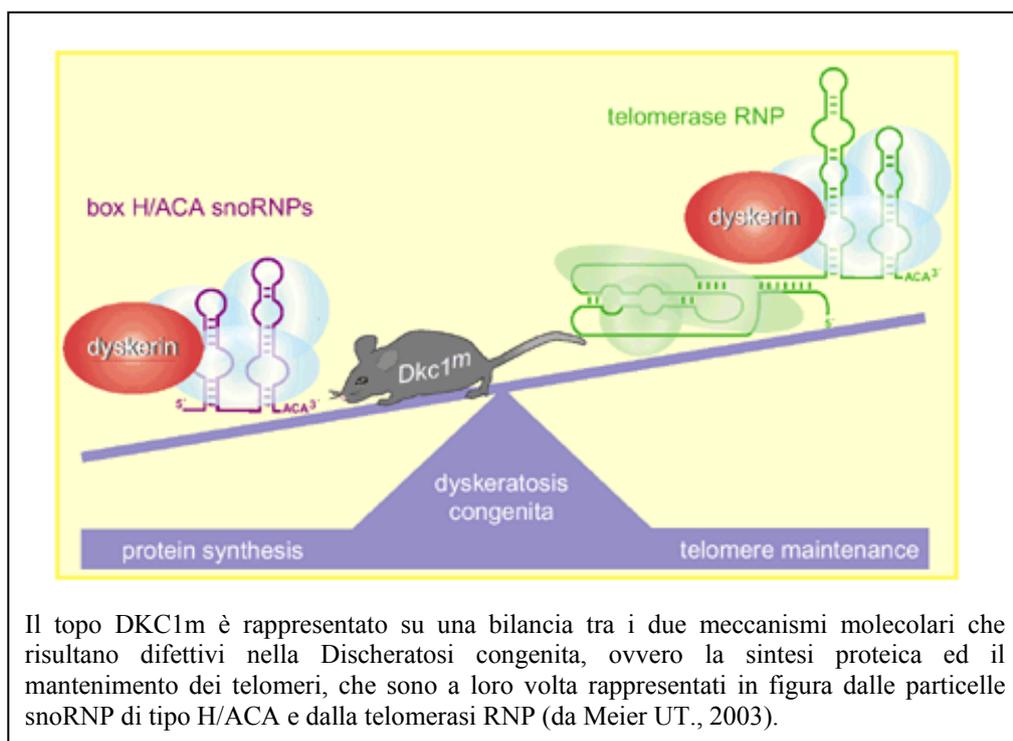
La discheratosi congenita (DC) è una malattia multisistemica che presenta eterogeneità clinica e genetica: sono note infatti forme recessive legate all'X e forme autosomiche, sia dominanti sia recessive. La malattia è caratterizzata da una triade di sintomi muco cutanei che comprende anomalie della pigmentazione cutanea, distrofia ungueale e leucoplachia delle mucose.



Anomalie fenotipiche in individui affetti da discheratosi congenita legata all'X.

Possono essere presenti anche altre anomalie, come corta statura, perdita di denti e capelli, difficoltà nell'apprendimento e lesioni del tratto gastrointestinale. La rilevante insufficienza del midollo osseo, l'invecchiamento prematuro e l'aumentata suscettibilità all'insorgenza di tumori causata da questa malattia possono condurre a morte prematura.

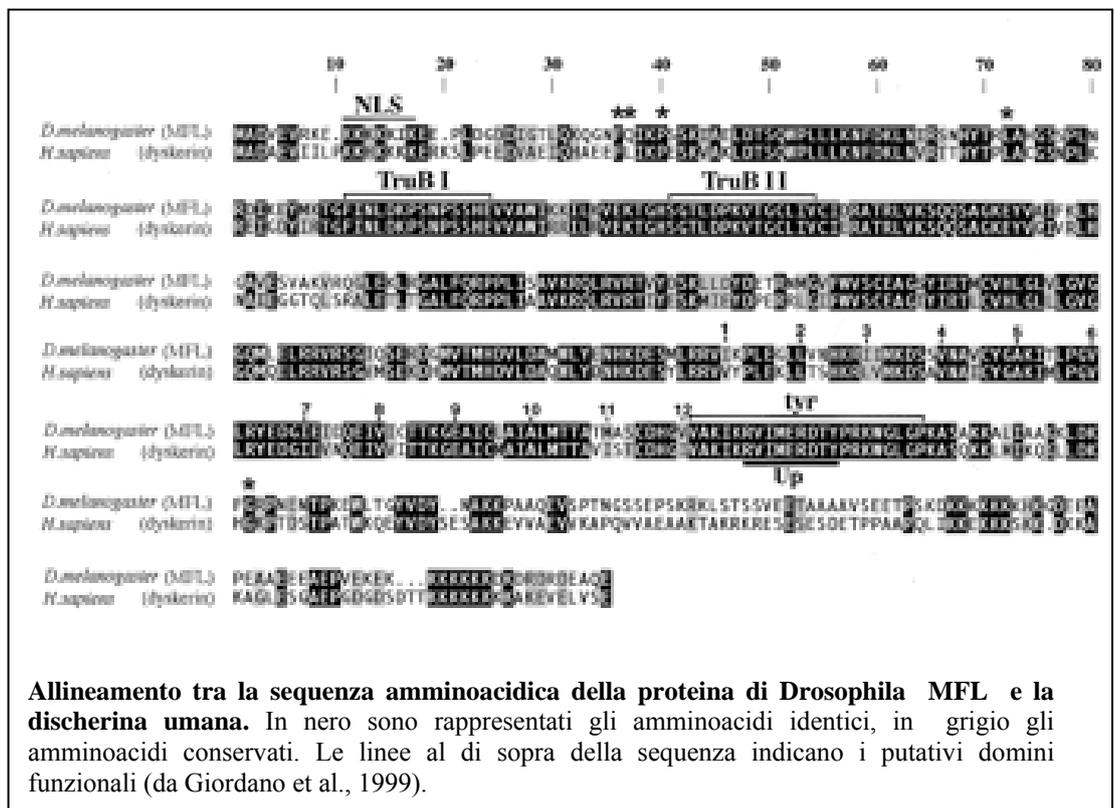
E' stato dimostrato che una forma autosomica dominante della discheratosi è dovuta a mutazioni della componente ad RNA della telomerasi (hTR) (Vulliamy et al, 2001; per una review, vedi Mason, 2003). I pazienti portatori di mutazioni nel gene *hTR* hanno infatti i telomeri significativamente più corti degli individui sani della stessa età e la gravità della malattia aumenta nel corso delle generazioni, un fenomeno di anticipazione probabilmente dovuto all'ereditarietà dei telomeri più corti (Wong et al., 2004). Queste osservazioni hanno rafforzato l'ipotesi, sostenuta da alcuni autori, che anche la forma X-linked della malattia potesse essere correlata essenzialmente alla riduzione della lunghezza dei telomeri (Mitchell et al, 1999). Nell'uomo la discherina, prodotta dal gene *DKCI*, svolge infatti almeno due tipi di diverse funzioni: una prima funzione, come già detto, è quella di dirigere la maturazione dell'rRNA e la modificazione dell'rRNA e di altri RNA cellulari, quali gli snRNA, mentre l'altra è quella di stabilizzare il complesso enzimatico della telomerasi, legando una specifica sequenza di hTR, che nei mammiferi presenta un dominio H/ACA. Mitchell e collaboratori (1999) osservarono infatti che i fibroblasti primari di maschi colpiti da discheratosi congenita X-linked accumulavano livelli ridotti di hTR; in queste cellule l'attività telomerasica appariva diminuita ed i telomeri erano più corti del normale, suggerendo che quest'aspetto potesse essere determinante per l'insorgenza della patologia. Successivamente, differenti mutazioni della discherina sono state associate alla ridotta stabilità di differenti sottoinsiemi degli snoRNA di tipo H/ACA (Mochizuki et al., 2004). Sebbene le conseguenze fisiologiche di questi effetti siano attualmente sconosciute, è probabile che alcune caratteristiche della malattia possano essere



correlate ai ruoli funzionali svolti dalle specifiche molecole di snoRNA la cui stabilità risulta più drasticamente colpita. Infatti, l'interazione con tutti i piccoli RNA nucleolari (snoRNA) della classe H/ACA rappresenta una proprietà rilevante della discherina le cui implicazioni non sono state finora completamente ben definite.

E' interessante notare che la discherina appartiene ad una famiglia proteica altamente conservata, i cui membri sono presenti in tutti gli organismi eucariotici e perfino negli Archaea. Queste proteine anche se presenti in tutti gli eucarioti, vengono spesso indicate con sigle diverse: Cbf5p in lievito, tripanosoma e Archaea, MFL in *Drosophila melanogaster*, Nap57 in ratto e discherina in topo e uomo. Data la stretta correlazione funzionale e strutturale esistente tra il gene di *Drosophila* e quello umano, l'isolamento dei mutanti *mfl* fornì il primo sistema modello animale per lo studio delle basi molecolari della discheratosi congenita X-linked. Il gene *mfl* è infatti altamente correlato alla sua controparte umana, il gene *DKC1*, sia a livello strutturale che funzionale.

La proteina MFL condivide infatti il 66% di identità e il 79% di similarità con la discherina umana (Giordano ed altri,1999), e ciò rafforza la validità di questo sistema modello. E' da notare che totale identità esiste tra le proteine di *Drosophila*, di topo e uomo all'interno dei domini TruBI e TruBII che sono direttamente coinvolti nel processo di pseudouridilazione. Un dominio idrofobico ripetuto, probabilmente coinvolto nel trasporto nucleocitoplasmatico della proteina del ratto (Meier e Blobel, 1998), è anche altamente conservato fra *Drosophila* e mammiferi.



Questo dominio è seguito da un tratto di più di 20 amminoacidi aventi un residuo di tirosina centrale che è identico nella proteina di *Drosophila* e nella discherina di ratto, topo e uomo e che, data la sua conservazione, potrebbe

svolgere un ruolo funzionale rilevante. Un'ulteriore aspetto è rappresentato dalla presenza, sia in mammifero che in *Drosophila*, di due segnali di localizzazione nucleare (NLS), uno in posizione N-terminale e l'altro in posizione carbossi-terminale .

È interessante notare che le cinque mutazioni missenso più frequenti nei malati di discheratosi congenita X-linked cadono nelle regioni conservate tra il gene umano e quello della *Drosophila*, confermando ulteriormente la validità dei mutanti *mfl* quali sistema modello per lo studio di questa malattia umana. Mutazioni missenso nel gene *DKCI* sono state identificate anche nei malati affetti dalla sindrome di Hoyeraal-Hereidarsson (Knight et al., 1999; Yaghmai ed altri, 2000), la quale è ora riconosciuta come una grave variante allelica della DC. Anche se le mutazioni più frequenti del gene *DKCI* sono principalmente di tipo missenso, sono state trovate anche mutazioni in due ipotetici siti di splicing (uno dei quali si trova nell'ultimo esone) e una mutazione del promotore che dovrebbe agire sui livelli di espressione della discherina piuttosto che sulla sua sequenza amminoacidica. Quest' ultimo dato suggerisce che anche una diminuzione della quantità della proteina normale possa causare la malattia, come già osservato in *Drosophila* (per una review, vedi Salowsky et al, 2002; Vulliamy et al, 2006). L'analisi genetica e funzionale dei mutanti *mfl* ha in effetti concretamente anticipato molte delle osservazioni successivamente riportate per i topi mutanti nel gene *DKCI*. Infatti, un ruolo analogo a quello svolto dalla proteina MFL nella maturazione dell'rRNA e nella pseudouridilazione (Giordano et al. 1999) è stato successivamente dimostrato anche per la discherina umana e di topo (He et al.,

2002; Ruggero et al., 2003), indicando che l'alterazione della biogenesi dei ribosomi costituisce sicuramente un aspetto rilevante per la patogenesi di questa malattia. Poiché la proteina MFL può occasionalmente diffondere nel citoplasma (Giordano et al., 1999), come osservato precedentemente per la proteina NAP57 in cellule di ratto (Meier e Blobel, 1998), è stato ipotizzato che queste proteine possano trasportare i complessi rRNP nel citoplasma. Appare quindi molto probabile che le proteine di questa famiglia possano essere coinvolte in differenti aspetti della biogenesi del ribosoma. Per quanto riguarda il gene *mfl*, poiché nessuna delle mutazioni isolate finora colpisce la regione codificante del gene, i mutanti ottenuti mostrano alterazioni quantitative, ma non qualitative del prodotto genico. Il livello della proteina MFL è risultato essere critico e mostra un effetto-dose: quando esso si abbassa al di sotto di un valore soglia l'effetto è di letalità, mentre se il livello è ridotto, ma al di sopra di un valore soglia, si raggiunge il fenotipo vitale e ipomorfo vitale tipico della mutazione *mfl^l* (Giordano et al, 1999). È possibile che anche nell'uomo il livello di attività della discherina potrebbe rappresentare uno dei parametri critici capaci di generare la malattia. L'osservazione che le mutazioni che causano questa patologia finora identificate sono quasi tutte localizzate nelle regioni codificanti è solo in apparente contrasto con quanto osservato nella *Drosophila*, dato che le mutazioni da totale o grave perdita di funzione potrebbero risultare non vitali. Le mutazioni più frequenti nei malati potrebbero essere quelle che causano una parziale perdita di funzione, così che il livello di attività della discherina sarebbe ancora compatibile con la sopravvivenza. In accordo con questa ipotesi, i pazienti potrebbero portare

mutazioni ipomorfe, equivalenti al fenotipo vitale *mfl*^l. Quest'ipotesi è stata rafforzata dai risultati di esperimenti effettuati in topo. Infatti, utilizzando la ricombinazione omologa ed il sistema inducibile Cre/loxP sono state prodotte due tipi di delezioni del gene *DKCI*: una prima delezione (DC1) rimuoveva gli esoni 12-15, che codificano un tratto di 124 amminoacidi localizzato all'estremità C-terminale della discherina, mentre un'altra delezione (DC7) eliminava solo l'ultimo esone che codifica 21 amminoacidi in posizione C-terminale, riproducendo una delezione osservata in uno dei pazienti esaminati (He et al., 2002). Sorprendentemente, entrambe le delezioni mostrarono un effetto materno, dato che veniva osservato il 100% di letalità embrionale solo quando la delezione era presente sull'allele *DKCI* fornito dalla madre (He et al., 2002). Le femmine eterozigoti per le delezioni del gene *DKCI* erano incapaci di trasmettere l'allele alle generazioni successive, perché esso risultava letale sia negli embrioni maschili che femminili. Quest'effetto materno fu osservato sia per i ceppi portatori della delezione DC1 che per quelli portatori della delezione DC7. Nessun embrione vitale nei giorni E 7.5 ed E 9.5 era di sesso maschile, mentre quelli di sesso femminile morivano più tardi, all'incirca nel giorno E 9.5 e mostravano una degenerazione dei tessuti extraembrionali, nei quali è noto che è inattivato il cromosoma X paterno. Femmine portatrici della delezione del gene *DKCI* di derivazione paterna mostrarono un'estrema tendenza all'inattivazione dell'X mutante ed il cromosoma X selvatico risultava attivo in pressochè tutte le cellule. Poiché l'inattivazione casuale del cromosoma X avviene ad uno stadio precoce dell'embriogenesi, le cellule nelle quali il cromosoma X attivo portava l'allele

mutante diventavano più piccole di quelle che esprimevano l'allele normale e venivano progressivamente eliminate per competizione clonale. L'osservazione che anche la piccola delezione che copriva solo l'ultimo esone (DC7) provocava letalità (He et al., 2002), sottolinea che la porzione 3' del gene ha un'alta rilevanza biologica anche nei mammiferi e non solo in *Drosophila* dove, come già detto, i due trascritti di 1.8 e 2.0 kb prodotti dal gene *mfl* differiscono solo per la loro sequenza 3'UTR. E' poi interessante notare che il trascritto di 2.0 kb, presente nelle preparazioni di RNA di embrioni e di femmine adulte, fornisce una possibile base molecolare per l'effetto materno osservato in *Drosophila*, mentre i meccanismi che causano l'effetto materno osservato in topo per il gene *DKCI* restano da definire.

In conclusione, l'insieme dei dati ottenuti in topo suggerisce che la discheratosi X-linked possa essere correlata ad una ridotta attività della discherina, mentre la completa inattivazione del gene *DKCI* provoca letalità embrionale. Un altro dato degno di nota è che, sebbene la discherina sia una proteina ubiquitaria, gli aspetti patologici osservati nei pazienti si verificano principalmente in alcuni tessuti. Poiché l'attività della discherina è necessaria per la sintesi proteica in ogni cellula, quest'osservazione è abbastanza sorprendente. Se però si assume che il livello dell'attività proteica sia un parametro critico, la quantità di rRNA correttamente maturato o pseudouridilato potrebbe essere sufficiente nelle cellule a crescita lenta, mentre i tessuti ad alta proliferazione, dove è richiesto un più elevato tasso di sintesi proteica, accumulerebbero difetti degenerativi. E' interessante poi notare che l'inibizione della sintesi proteica è uno degli stimoli capaci di indurre la morte della cellula per apoptosi, probabilmente attraverso

la diminuzione dei livelli di proteine essenziali o inibendo la sintesi delle proteine che normalmente sopprimono l'attivazione di questo processo. Analogamente la riduzione dell'attività della discherina potrebbe indurre l'apoptosi, così come è stato osservato in *Drosophila* per le mutazioni del gene *mfl*. Un'ulteriore interessante correlazione fra i geni ortologhi di *mfl* e *DKC1* e uomo è basata sui risultati di esperimenti di Northern blot che hanno portato ad identificare, sia in topo che in uomo, un raro trascritto alternativo di circa 4.5 kb che si aggiunge al trascritto principale di 2.6 kb di tipo ubiquitario (Heiss et al, 1998; Heiss et al, 2000). In topo, questo lungo trascritto è stato osservato nel fegato, nel rene e negli embrioni di 7 giorni, mentre nell'uomo è stato osservato nei linfociti del sangue periferico e, a livelli molto bassi, nei testicoli e nel cervello. Questo trascritto, la cui struttura e la cui funzione sono finora ignote, potrebbe corrispondere ad uno di dimensioni simili (4.4 kb) osservato in *Drosophila* nelle preparazioni di RNA ottenute da femmine adulte e da embrioni (Giordano et al, 1999). I geni *DKC1* e *mfl* condividono un'altra caratteristica, cioè entrambi ospitano geni che codificano snoRNA nei loro introni (Giordano et al, 1999; Huttenhofer et al., 2001; He et al., 2002; Accardo et al., 2004; Riccardo et al., manoscritto in preparazione). Come già descritto, *mfl* ospita nel suo quarto introne un gene che codifica per uno snoRNA di tipo H/ACA, chiamato *snoHI*, che è risultato funzionalmente equivalente allo snoRNA umano *U70* (Ganot et al, 1997). Anche il gene umano *DKC1* ospita nei suoi introni geni che codificano snoRNA di tipo H/ACA. Nell'ottavo introne è infatti localizzato lo snoRNA ACA 36 (indicato come snoRNA 36), uno snoRNA bifunzionale che guida la pseudouridilazione dello rRNA18S

nella posizione U105 ed U1244, mentre nel dodicesimo introne è invece ospitato il gene ACA 56, responsabile della pseudouridilazione dello rRNA 28S in posizione U1664 (Huttenhofer et al., 2001; He et al., 2002; Kiss et al., 2004). I geni che codificano snoRNA sono spesso localizzati in introni di geni che codificano proteine coinvolte nel processo di traduzione o nella biogenesi del ribosoma (Smith e Steitz, 1997). Gli snoRNA intronici generalmente sono cotrascritti con il pre-mRNA del gene ospite e la loro biosintesi è per lo più dipendente dallo splicing che libera l'introne, che viene poi maturato mediante digestione esonucleolitica (Tycowski *et al.*, 1993; Kiss and Filipowicz, 1995).

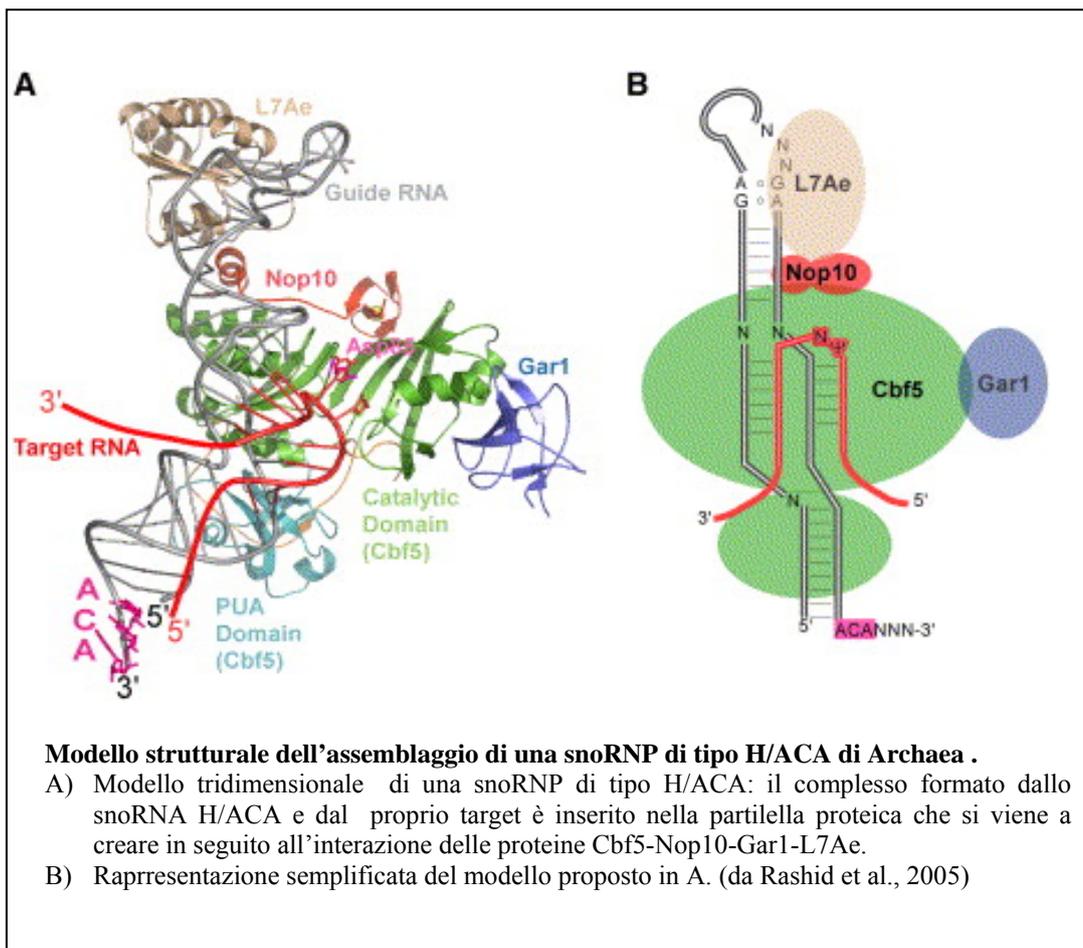
Struttura e funzione dei complessi di tipo snoRNP H/ACA

La modificazione post-trascrizionale delle molecole di RNA è una caratteristica comune a tutti gli organismi. La metilazione del ribosio sul gruppo 2'-ossidrilico e la conversione di uridine in pseudouridine rappresentano i due tipi di modificazioni più frequentemente riscontrati a carico degli RNA cellulari, primo fra tutti l'RNA ribosomale.

Dagli Archaea agli eucarioti, questi due tipi di modificazione, sono mediate da complessi ribonucleoproteici in cui un piccolo RNA "guida" riconosce il sito da modificare mediante complementarità. Le piccole molecole di RNA che "guidano" questo tipo di modificazioni presentano localizzazione nucleolare e si dividono in due grandi classi: la classe degli snoRNA di tipo C/D e quella di tipo H/ACA. Questa suddivisione è basata sulla presenza di distinti motivi di sequenza (chiamati box C, D, H ed ACA) che sono evolutivamente conservati in tutte le specie studiate (Kiss, 2001). Gli snoRNA (small nucleolar RNA) della classe C/D sono responsabili della metilazione dei gruppi 2'-*O*-idrossilici del ribosio, mentre gli snoRNA della classe H/ACA dirigono la conversione delle uridine in pseudouridine nelle molecole degli RNA bersaglio (Nicoloso *et al.*, 1996; Ni *et al.*, 1997; Weinstein and Steitz, 1999). Sebbene sia stata evidenziata una certa variabilità nel numero e nella composizione del repertorio degli RNA guida nei diversi organismi, la natura delle proteine con cui le due famiglie di snoRNA si associano ed i meccanismi attraverso cui dirigono la specifica modificazione del nucleotide bersaglio sono risultati molto conservati durante l'evoluzione.

La pseudouridilazione dell'RNA consiste nella conversione delle uridine in pseudouridine ed aumenta la stabilità termica della molecola di RNA modificata, il repertorio di basi che possono interagire per formare strutture secondarie ed i possibili siti di legame per proteine (Arnez e Steitz, 1994; Newby e Greenbaum, 2002; Yarian et al, 1999). Negli eucarioti, l'importanza funzionale di questa modificazione è stata ampiamente dimostrata in vari processi cellulari, quali la biogenesi dei ribosomi, l'assemblaggio e la funzionalità dei complessi snRNP. L' isomero rotazionale dell'uridina, è spesso presente in regioni conservate dei tRNA, degli rRNA e degli snRNA. Nei batteri, la conversione delle uridine in pseudouridine è catalizzata da molte differenti molecole di RNA pseudouridina sintasi, ciascuna formata da un unico polipeptide, che assicura il riconoscimento dell'RNA substrato e forniscono l'attività catalitica (Motorin Y. et al., 1998; Nurse K. et al., 1995; Wrzesinski J. et al., 1995). Negli Archaea e negli eucarioti la reazione di pseudouridilazione (così come la 2'-O metilazione) è invece catalizzata da particelle ribonucleoproteiche (dette snoRNP negli eucarioti e sRNP negli Archaea; per una review vedi Henras et al, 2004; Bachellerie et al, 2002) formate da una molecola di RNA guida di tipo H/ACA e da un set di 4 proteine tutte evolutivamente molto conservate. L'RNA guida presente in un singolo complesso H/ACA snoRNP è responsabile della specificità della reazione perché contiene la regione complementare all'RNA target. Negli eucarioti, questa complementarietà è distribuita lungo due corte (5-9nt) sequenze localizzate in un loop interno allo snoRNA di tipo H/ACA che prende il nome di "tasca di pseudouridilazione". I due elementi di sequenza conservati, noti

come box ACA o box H (ANANNA), che sono caratteristici di questa famiglia di snoRNA sono localizzati alla base della forcina che ospita la tasca di pseudouridilazione e sono essenziali per l'associazione dello snoRNA alle 4 proteine sempre presenti nelle snoRNP di questa classe, cioè le proteine Nop10p, Nhp2p (L7Ae in Archaea), Gar1p e Cbf5p (Watkins et al., 1998; Bousquet-Antonelli et al., 1997; Dragon et al., 2000; Henras et al., 1998; Lafontaine et al., 1998; Wang and Meier, 2004). Le proteine sopracitate, altamente conservate anche negli Archaea, vanno quindi a costituire il “core” di questo tipo di particelle ribonucleoproteiche.



In lievito, tutte e quattro queste proteine sono risultate essenziali per la vitalità e tutte, tranne Gar1p, sono importanti per la stabilità delle altre componenti del complesso. La proteina Cbf5p di lievito, inizialmente isolata per la sua capacità di legare con bassa affinità i centromeri *in vitro*, è l'enzima che catalizza l'isomerizzazione delle uridine in pseudouridine e, come già accennato appartiene alla famiglia proteica di cui fanno parte anche la proteina MFL di *Drosophila* e la discherina umana.

I primi dati funzionali sui membri di questa famiglia genica derivano infatti soprattutto da studi effettuati in lievito (Jiang et al, 1993), *Drosophila* (Giordano et al, 1999) e ratto (Meier e Blobel, 1994). Dall'insieme di questi studi emerse chiaramente che esse erano essenziali per la biosintesi, la maturazione ed il trasporto dell'RNA ribosomale (Lafontaine et al, 1998; Giordano et al., 1999). Successivamente, anche in topo e uomo fu osservato che l'inattivazione del gene *DKC1* influenzava non solo il meccanismo di maturazione dell'rRNA ma anche la sua pseudouridilazione (Lafontaine et al, 1998; Ruggero et al., 2003; He et al., 2002), sottolineando l'esistenza di uno stretto legame tra questi due processi.

L'inefficiente maturazione dell'rRNA causata dalle mutazioni nei geni *Cbf5/mfl/DKC1* potrebbe in parte derivare dalla diminuita pseudouridilazione degli rRNA precursori o dal ridotto accumulo degli snoRNA che mediano gli eventi di taglio durante la maturazione del pre-rRNA. È stato suggerito che la pseudouridilazione può contribuire anche al ripiegamento dell'rRNA, all'assemblaggio delle rRNP e a quello delle subunità ribosomali. Un'ulteriore ipotesi è che la pseudouridilazione possa influenzare la capacità di

riconoscimento del codone da parte del ribosoma. Quest'ipotesi è stata recentemente avvalorata da dati ottenuti da linee cellulari dei pazienti con discheratosi congenita X-linked e cellule staminali derivate da embrioni di topo portanti mutazioni puntiformi del gene *DKC1* (*DKC1^m*). In questi esperimenti le cellule mutanti hanno mostrato difetti nella traduzione di un sottoinsieme di mRNA che condivide la modalità di inizio della traduzione a partire da un elemento IRES (IRES: Internal Ribosomal Entry Site). È stato infatti osservato che tre mRNA, prodotti rispettivamente dai geni p27(kip1), XIAP e Bcl-xL, erano scarsamente tradotti nelle cellule mutanti, e tutti presentavano un elemento IRES che media l'attacco interno al ribosoma (Yoon et al., 2006). L'elemento IRES è importante per l'accurata espressione non solo dei tre mRNA sopracitati ma anche di molti mRNA cellulari e virali (Holcik et al., 1999; Kullmann et al., 2002; Yoon et al., 2006). Sebbene gli elementi IRES siano attivi in condizioni di crescita normale, la traduzione dipendente da questi elementi è favorita da specifici stimoli cellulari, quali l'induzione della morte cellulare programmata (indotta per esempio da radiazioni gamma) o durante fasi distinte del ciclo cellulare, quando è diminuita la traduzione cap-dipendente (le fasi G₀/G₁ del ciclo cellulare, Holcik et al., 2005). Questo difetto traduzionale potrebbe spiegare due specifiche caratteristiche patologiche dell'X-DC: la suscettibilità all'insorgenza del cancro e l'insufficienza midollare, che si estrinseca in una ridotta produzione delle cellule della linea ematopoietica.

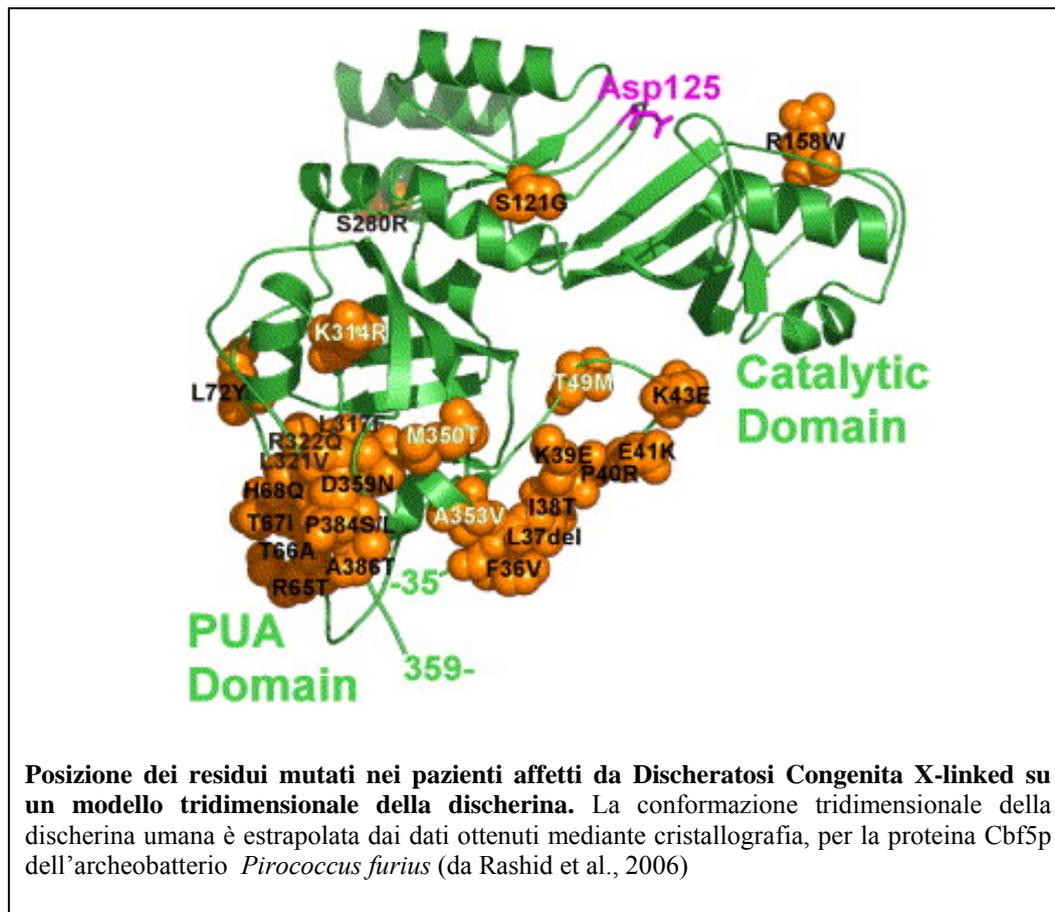
Negli ultimi anni numerose ricerche sono state focalizzate sulla struttura delle componenti delle snoRNP di tipo H/ACA e sulle loro modalità di assemblaggio

per la formazione di complessi funzionalmente attivi. Negli eucarioti, la difficoltà di ottenere forme ricombinanti solubili delle proteine Cbf5/Nap57 e Gar1 ha finora reso notevolmente complicati gli studi di tipo strutturale e funzionale, e le informazioni attualmente a disposizione possono descrivere l'interazione solo tra alcune delle componenti di questi complessi. Molto di quello che attualmente è noto sull'assemblaggio di queste snoRNP deriva infatti da studi condotti negli Archaea, organismi in cui le componenti proteiche risultano più solubili. Data l'elevata conservazione di queste componenti è però possibile estrapolare agli eucarioti molte delle informazioni ottenute in questi semplici organismi. Negli Archaea è stato recentemente possibile ricostituire un complesso sRNP H/ACA attivo a partire da una molecola di RNA trascritta *in vitro* e proteine ricombinanti espresse in *E.coli*. Questo studio ha evidenziato che aCbf5p riconosce e lega specificamente gli RNA guida della classe H/ACA e che questa interazione dipende dal trinucleotide conservato ACA e dalla tasca di pseudouridilazione (Baker et al., 2005; Charpentier et al., 2005). Risultati analoghi sono stati ottenuti recentemente anche in lievito, dove è stato riportato il primo studio *in vitro* che descrive l'interazione tra una proteina Cbf5p eucariotica e uno snoRNA H/ACA (Normand et al., 2006).

Nel lievito e negli Archaea, la proteina Cbf5p è stata dimostrata responsabile anche del reclutamento delle altre due proteine essenziali per la formazione del complesso guida della pseudouridilazione, cioè le proteine Gar1p e Nop10p. Infatti, saggi di interazione proteina-proteina hanno evidenziato che Gar1p e Nop10p interagiscono direttamente con Cbf5p anche in assenza dell'RNA

guida e dell'altra proteina essenziale L7Ae di Archaea ortologo a Nhp2p di lievito e mammiferi.

La struttura cristallina della proteina aCbf5p degli Archaea ha recentemente fornito l'opportunità di studiare l'arrangiamento spaziale degli amminoacidi che sono alterati nei pazienti affetti da discheratosi congenita X-linked (Rashid et al., 2006).



La struttura di aCbf5 è risultata simile a quella dell'enzima batterico TruB (Hoang and Ferre-D'Amare, 2001), ma un' importante differenza che caratterizza Cbf5p è la presenza di una sequenza di 33 residui aggiuntivi all'N terminale che avvolge il dominio PUA, localizzato in posizione C-terminale (Manival et al., 2006). Questa sequenza ammino-terminale è stata trovata in tutti gli enzimi aCbf5p caratterizzati negli Archaea, così come nelle proteine

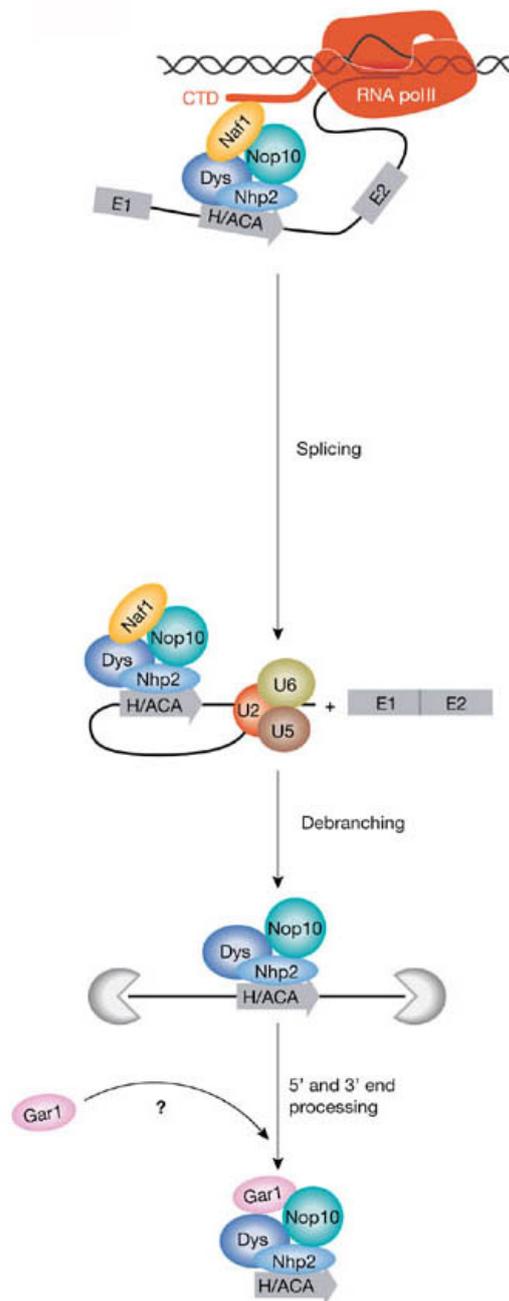
ortologhe degli eucarioti. I 33 residui N terminali della proteina aCbf5p sono risultati importanti per il corretto folding e/o stabilità dell'enzima. Il sito attivo di aCbf5p contiene un residuo Asp, conservato in tutte le RNA pseudouridina sintasi, che è essenziale per l'attività catalitica (Huang et al., 1998; Gu et al., 1999). Nell'enzima batterico TruB due residui di tirosina (Tyr76 e Tyr179) giocano un ruolo importante nell'orientamento dell'uracile che è bersaglio della modificazione (Hoang and Ferre-D'Amare, 2001) ed entrambi i residui sono conservati nella proteina aCbf5p degli Archaea. Uno dei ruoli giocati dalla proteina Nop10p sembra essere quello di mascherare le cariche negative presenti sulla superficie di aCbf5p, favorendo l'interazione tra l'RNA guida e l'RNA substrato (Manival et al., 2006). Il dominio PUA della proteina aCbf5p probabilmente interagisce direttamente con la sequenza ACA degli snoRNA, che è localizzata a una distanza costante dall'uridina bersaglio nel complesso sRNA-RNA target (da 14 a 16 bp fiancheggiando il motivo ACA) (Manival et al., 2006). Questa distanza costante può rappresentare un parametro determinante per un buon posizionamento dell'uridina che deve essere modificata all'interno del sito attivo. Le mutazioni più frequentemente presenti nei pazienti affetti da discheratosi X-linked sono localizzate nel dominio PUA (per esempio le mutazioni A353V e G402E) e nell'espansione dell'N terminale che avvolge il dominio PUA. Questi dati hanno sottolineato l'importanza del dominio PUA nell'assemblaggio degli snoRNP di tipo H/ACA (Mitchell et al., 1999; Mochizuchi et al., 2004). Negli Archaea, la proteina aCbf5p, come già detto, media l'associazione delle proteine aGar1p e aNop10p dello snoRNA di tipo H/ACA. Queste tre proteine possono formare un complesso eterotrimerico

aCbf5/aGar1/aNop10 che può interagire con l'RNA guida (Baker et al, 2005; Charpentier et al, 2005). Dai dati emersi dagli studi sulla proteina aCbf5 degli Archaea appare prevedibile che le mutazioni interne al dominio PUA dovrebbero indebolire specificamente l'interazione tra la proteina Cbf5 e gli snoRNA, inducendo un ridotto accumulo di queste molecole.

L'organizzazione dei complessi RNP di tipo H/ACA nel lievito assomiglia a quella degli Archaea; infatti, Cbf5p, Gar1p e Nop10p possono formare un complesso indipendentemente dal quarto componente essenziale, la proteina Nhp2p (L7Ae di Archaea), e dall'RNA guida. In lievito, Cbf5p è risultata necessaria per la stabilità delle altre componenti delle snoRNPs, come Gar1p e per quella degli snoRNA del tipo H/ACA (Lafontaine et al, 1998). Un'importante dato emerso recentemente è che, negli Archaea ed in lievito, la proteina Cbf5p interagisce indipendentemente con Gar1p e con Nop10p e lega gli RNA guida anche in assenza delle altre proteine, mentre nei mammiferi il riconoscimento specifico degli snoRNA H/ACA richiede tutte e tre le componenti del complesso trimerico. Da esperimenti eseguiti da Wang e Meier (2004) sembra infatti che nei mammiferi la proteina Nop10 giochi un ruolo centrale nella formazione del complesso eterotrimerico Cbf5p/Nop10p/Gar1p. Attualmente, non si sa se le discrepanze osservate tra il sistema dei mammiferi e quello degli Archaea e del lievito riflettono fondamentali differenze nelle snoRNPs. Al di là di questa differenza, la conservazione funzionale delle componenti principali degli snoRNP della classe H/ACA è così notevole che la proteina aCbf5 degli Archaea è stata recentemente mostrata assemblarsi efficientemente con la proteina Nop10p di lievito e con l'RNA della telomerasi

umana (Hamma et al., 2005), che è noto contenere la box H/ACA. Nell'uomo, il ruolo giocato dalla discherina come componente degli snoRNP H/ACA potrebbe essere cruciale nel determinare la capacità proliferativa delle cellule. E' importante sottolineare che possono essere pseudouridilati, oltre all'rRNA, anche i tRNA, gli snRNA spliceosomali (Donmez et al, 2004; Valadkhan and Manley, 2003, Yang et al, 2005; Zhao and Yu, 2004) e probabilmente altri tipi di RNA cellulari (per una review, vedi Bachellerie et al., 2002). Per esempio, risultati molto interessanti sono stati ottenuti recentemente nei tripanosomi. In questi organismi, il silenziamento del gene *Cbf5* ottenuto mediante RNAi provoca difetti nella maturazione dell'rRNA, principalmente dovuti all'instabilità dello snR30, uno snoRNA della classe H/ACA che è essenziale per la maturazione dell'rRNA. Tuttavia, la pseudouridilazione viene ridotta non solo sull'rRNA, ma anche sulle molecole di snRNA e sull'SLRNA (Splicing Leader RNA) coinvolto nel processo di trans-splicing. Sorprendentemente, il silenziamento di *Cbf5* non solo riduce la pseudouridilazione sull'SL RNA, ma ne impedisce anche la modificazione del nucleotide +4 del cap, che è stato dimostrato importante per l'efficienza del processo di *trans-splicing*. Questi difetti di modificazione si traducono nell'inibizione dell'intero meccanismo di trans-splicing, che contribuisce alla riduzione dei livelli di crescita cellulare (Barth et al., 2005). Funzioni addizionali sono state attribuite anche alla proteina *Cbf5p* di lievito, dove il silenziamento del gene *Cbf5* provoca frammentazione del nucleolo ed impedisce la localizzazione nucleolare dei tRNA (Kendall et al., 2000). Particolarmente interessante è poi il coinvolgimento di questa proteina nella

biosintesi degli snoRNA della classe H/ACA. L'esistenza di una sinergia tra il processamento degli snoRNA intronici e lo splicing del pre-mRNA ospite è un'aspetto ampiamente sviluppato negli ultimi anni. E' già noto da qualche tempo che la proteina Naf1p è un fattore essenziale per la biogenesi delle snoRNP di tipo H/ACA (Dez et al, 2002; Fatica et al, 2002; Yang et al, 2002). In lievito, questa proteina partecipa alle prime fasi della sintesi degli snoRNA H/ACA e si trova associata ai geni (per snoRNA H/ACA) in attiva trascrizione, mentre non mostra nessuna interazione con le particelle snoRNP mature. (Yang *et al*, 2005; Ballarino *et al*, 2005). Risultati analoghi sono stati evidenziati recentemente anche in cellule umane, dove il gruppo di Thomas Meier ha chiarito le modalità di trascrizione dei geni per snoRNA di tipo H/ACA intronici ed ha mostrato che Naf1 insieme alle proteine discherina, Nhp2p e Nop10 si localizzano specificamente sui geni per snoRNA H/ACA trascrizionalmente attivi (Darzacq *et al*, 2006). Lo stesso gruppo ha proposto che, nel corso della maturazione delle snoRNP, Naf1, è rimpiazzata dalla proteina Gar1, e questo fa sì che la particella matura sia rilasciata dal sito di trascrizione dello snoRNA H/ACA e trasportata nel nucleolo. L'insieme dei dati ottenuti in vari organismi dunque, suggerisce che le funzioni svolte da queste essenziali proteine possano essere molteplici.



Modello di assemblaggio dipendente dalla trascrizione di una snoRNP intronica di tipo H/ACA .L'assemblaggio del presnoRNP H/ACA avviene con una modalità co-trascrizionale. Il fattore di assemblaggio nucleare (Naf1), può interagire con il dominio carbossi-terminale (CTD) della RNA polimerasi (pol) II, è può facilitare il richiamo ed il legame delle proteine del core della particella snoRNP di tipo H/ACA alla sequenza dello snoRNA in fase di sintesi. La sostituzione della proteina Naf1 con Gar1 determina il rilascio della particella matura (da Richard and Kiss. 2006)

Risultati

Caratterizzazione molecolare del gene *mfl*

Considerata la varietà dei ruoli biologici svolti dai geni orologi *Cbf5/mfl/DKCl* ho deciso di analizzare in maggiore dettaglio l'attività trascrizionale del gene *mfl* di *Drosophila* (chiamato anche Nop60B), con l'obiettivo di identificare nuovi prodotti che potessero fornire un'adeguata base molecolare alle molteplici funzioni attribuite ai membri di questa famiglia genica. Come precedentemente accennato, dal gene *mfl* derivano due molecole di mRNA, rispettivamente di 1.8 e 2.0 kb (Giordano et al., 1999). Questi due messaggeri presentano un differente profilo di espressione, con il trascritto da 1.8 kb presente in maniera costitutiva in entrambi i sessi e durante tutto il ciclo vitale e quello da 2.0 kb accumulato preferenzialmente in femmine, dove si localizza negli ovari e viene trasmesso all'embrione per via materna. I due mRNA presentano un' identica regione codificante e differiscono solo a livello della loro regione 3' non tradotta (3' UTR). L'identificazione di questi due trascritti tuttavia, non è risultata sufficiente a spiegare i risultati ottenuti in esperimenti di recupero fenotipico (Giordano et al., 1999). Infatti, i tentativi di recuperare la letalità provocata dalle mutazioni *mfl*⁰⁵ e *mfl*⁰⁶ utilizzando un transgene comprendente il cDNA completo rappresentativo del trascritto costitutivo di 1.8 kb sotto il controllo del promotore *hsp70* portarono a risultati negativi (Giordano et al., 1999). Per spiegare questo risultato fu considerata la possibilità che i livelli di espressione ectopica della proteina fossero insufficienti a sostenere la sintesi proteica in alcuni tipi di cellule, soprattutto durante il processo di metamorfosi, oppure che i tassi di sintesi della proteina MFL, essendo sotto il controllo di un promotore inducibile, non fossero

sufficientemente costanti e ciò potesse interferire con la stabilità delle altre componenti dei complessi snoRNP di tipo H/ACA (Giordano et al., 1999). Inoltre, poiché nel quarto introne del gene *mfl* è ospitato il gene per lo snoRNA *H1*, non si poteva escludere che la letalità causata dalle mutazioni *mfl*⁰⁵ ed *mfl*⁰⁶ (i due alleli che presentano il fenotipo più drastico) potesse essere causata non solo dalla riduzione della proteina MFL ma, almeno in parte, anche dal ridotto accumulo di questo snoRNA. Per chiarire quest'aspetto ho quindi preparato un secondo costrutto da utilizzare in esperimenti di transgenesi finalizzati ad ottenere il recupero della letalità indotta dalle mutazioni *mfl*⁰⁵ e *mfl*⁰⁶.

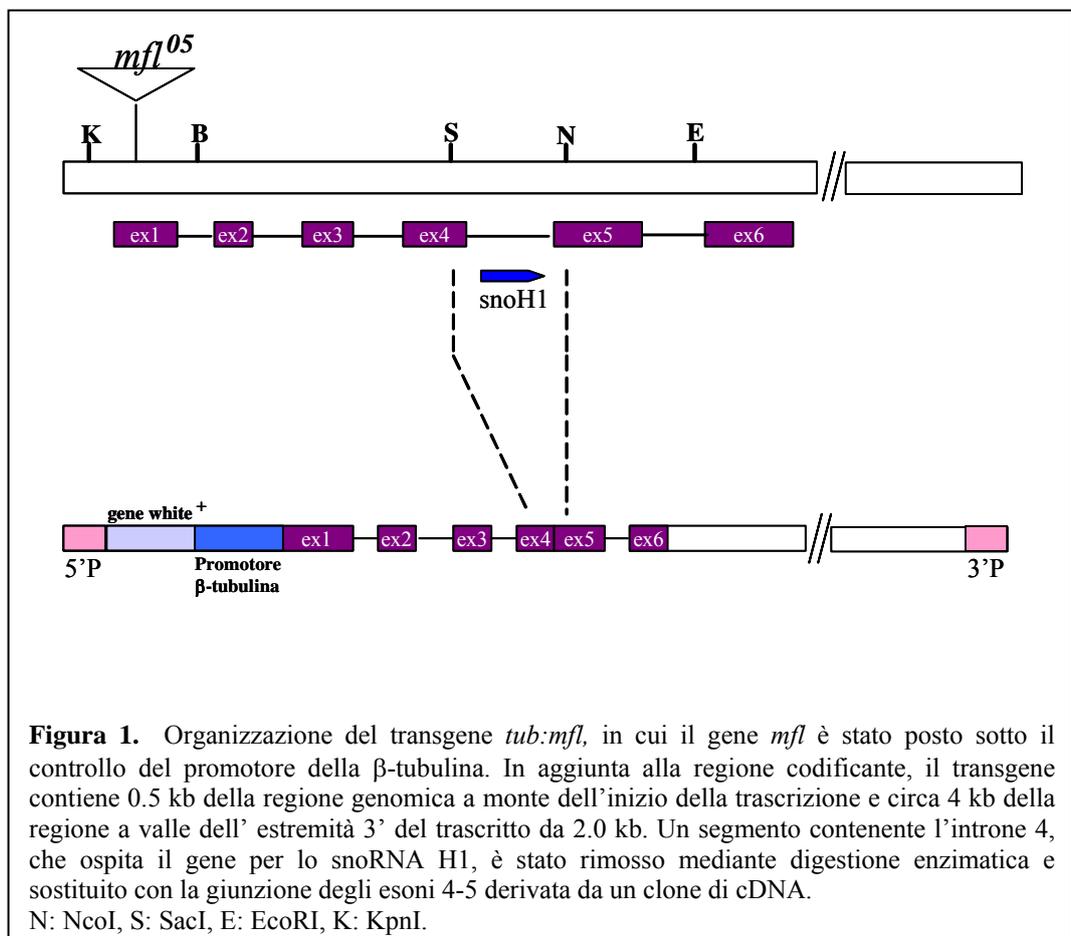
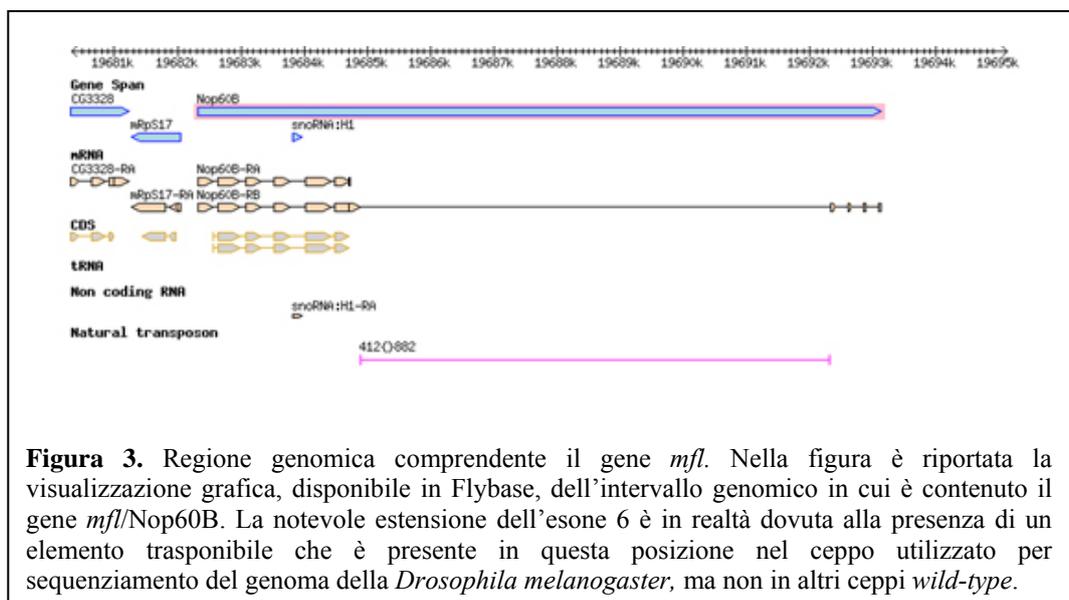


Figura 1. Organizzazione del transgene *tub:mfl*, in cui il gene *mfl* è stato posto sotto il controllo del promotore della β -tubulina. In aggiunta alla regione codificante, il transgene contiene 0.5 kb della regione genomica a monte dell'inizio della trascrizione e circa 4 kb della regione a valle dell'estremità 3' del trascritto da 2.0 kb. Un segmento contenente l'introne 4, che ospita il gene per lo snoRNA H1, è stato rimosso mediante digestione enzimatica e sostituito con la giunzione degli esoni 4-5 derivata da un clone di cDNA.
N: NcoI, S: SacI, E: EcoRI, K: KpnI.

Per la preparazione del transgene (denominato *tub:mfl*) ho utilizzato il vettore di trasformazione pTUBU, derivato dall'elemento trasponibile P, in cui ho inserito un frammento genomico di circa 8 kb comprendente l'intera regione codificante del gene *mfl*, 0.5 kb a monte del sito di inizio della trascrizione e circa 4 kb a valle del sito di poliadenilazione più distale. La trascrizione del gene *mfl* è stata posta sotto il controllo del promotore costitutivo della β -tubulina (fig. 1). Tuttavia, l'inclusione della regione genomica immediatamente fiancheggiante l'inizio della trascrizione del gene *mfl*, che ne contiene il promotore minimo, garantisce la possibilità che la trascrizione possa partire anche dal promotore endogeno, la cui sequenza potrebbe essere necessaria per una corretta espressione del gene *in vivo*. Il segmento genomico è stato poi ulteriormente modificato per verificare l'eventuale coinvolgimento dello snoRNA *H1* nella genesi del fenotipo mutante. A questo scopo, il quarto introne (che contiene il gene *H1*) è stato rimosso per digestione enzimatica, così che nel transgene la sequenza del quarto esone risulta contigua a quella del quinto esone. L'iniezione di questo costrutto nella linea germinale di embrioni del ceppo 3664 (descritto in materiali e metodi) ha portato ad ottenere due linee transgeniche indipendenti, l'una caratterizzata dall'inserzione stabile dell'elemento P sul cromosoma sessuale X e l'altra sul cromosoma 2. Poiché il gene *mfl* mappa sul cromosoma 2, per l'esperimento di recupero fenotipico è risultato vantaggioso utilizzare la linea transgenica in cui l'inserzione del transgene *tub:mfl* era avvenuta sul cromosoma X, dato che l'uso di questa linea consentiva di inserire il transgene in background mutante mediante semplici incroci. Gli incroci necessari per la produzione di ceppi *mfl*⁰⁵ recanti il vettore

con l'inserto genomico *tub:mfl* sono riassunti nello schema riportato in figura 2. L'analisi della fase letale condotta sugli individui omozigoti *mfl*⁰⁵ recanti il transgene *tub:mfl* sul cromosoma X ha successivamente dimostrato che questo costrutto garantisce un efficiente rescue fenotipico, con il 100% degli embrioni transgenici capace di raggiungere lo stadio adulto nei normali tempi di sviluppo (fig. 2). Una prima conclusione che si può trarre da questi esperimenti è che il segmento genomico di circa 8 kb inserito nel transgene *tub:mfl* contiene tutte le sequenze necessarie per un completo recupero fenotipico. Sulla base dei dati ottenuti appare inoltre evidente che lo snoRNAH1 in questi ceppi transgenici rimane molto basso, ma ciò non comporta nessuna anomalia fenotipica evidente, nè tanto meno ritardo di sviluppo o riduzione di fertilità. Questi risultati mi hanno quindi spinto ad analizzare in dettaglio l'attività trascrizionale del segmento genomico che permetteva un pieno recupero fenotipico. Un'analisi bioinformatica di questo segmento genomico utilizzando la banca dati Flybase (<http://flybase.bio.indiana.edu/>) ha portato a stabilire che in esso non erano stati identificati altri geni oltre *mfl*. Infatti, il primo gene fiancheggiante *mfl* in 3' dista circa 20 kb dal sito di poliadenilazione più a valle. Quest'indicazione è stata successivamente anche confermata da me direttamente, mediante esperimenti di Northern Blot in cui ho analizzato RNA poly(A)+ con sonde genomiche derivate da questa regione. Nessun trascritto è stato evidenziato in questi esperimenti, confermando i dati riportati in Flybase.

Per quanto riguarda l'estremità 5' del gene *mfl*, esperimenti di primer-extension avevano precedentemente portato ad identificare un unico sito di inizio della trascrizione (Giordano et al., 1999), e successive analisi nella banca dati Flybase hanno evidenziato che questo sito è localizzato solo 300 bp a monte del gene adiacente *RPS17*, trascritto con polarità opposta (fig. 3). L'insieme delle osservazioni raccolte indicava quindi che il transgene utilizzato non conteneva altri geni oltre *mfl*; tuttavia, il differente risultato ottenuto negli esperimenti di recupero fenotipico in cui era stato utilizzato un transgene contenente il cDNA completo rappresentativo del trascritto costitutivo di 1.8 kb suggeriva che l'attività trascrizionale del gene *mfl* non fosse stata completamente definita nel corso dei precedenti esperimenti (Giordano et al., 1999). L'utilizzo di diversi siti di inizio e termine della trascrizione e/o di eventi di splicing alternativo sono i meccanismi più frequentemente utilizzati per ampliare e diversificare le capacità codificanti dei geni eucariotici e potrebbero svolgere un compito importante anche nel determinare il ruolo biologico svolto dai membri della famiglia genica cui appartiene *mfl*.



Per verificare tale ipotesi ho quindi deciso di riesaminare l'attività trascrizionale di questo gene eseguendo, in primo luogo, esperimenti di 3'-RACE. Infatti, l'osservazione che il transgene capace di recuperare il fenotipo mutante comprendeva solo 500 bp a monte del sito di inizio della trascrizione ed il risultato dei precedenti esperimenti di primer extension rendevano più probabile la possibilità che eventuali nuovi trascritti del gene *mfl* potessero derivare dall'uso di siti alternativi di poliadenilazione. Per saggiare questa ipotesi ho quindi svolto esperimenti di 3' RACE in cui ho utilizzato come stampo preparazioni di RNA poly(A)⁺ ottenute da moscerini adulti separati in base al sesso. Sulla base della struttura degli mRNA di 1.8 e 2.0 kb precedentemente descritti (Giordano et al., 1999; fig. 4) ho selezionato diverse sequenze interne agli esoni 5, 6, 7 e 9 che ho poi utilizzato per dirigere la sintesi di oligonucleotidi (lungi da 23 a 29 nt) da utilizzare come "forward primer". Ogni "forward primer" è stato utilizzato in coppia con un "reverse primer" costituito dal primer oligoAdaptor (per la posizione del "forward primer" vedi fig. 4A; per la sequenza vedi Materiali e Metodi). Infatti negli esperimenti di 3' RACE la reazione di retro trascrizione per convertire l'RNA poly(A)⁺ in cDNA viene eseguita con un particolare oligo dT che è caratterizzato, nella sua estemita 5', dalla presenza di una sequenza di 15 nt che prende il nome di sequenza Adaptor. Questa sequenza, in combinazione con gli oligonucleotidi specifici per il gene d'interesse, permette di ottenere nelle successive reazioni di amplificazione l'estremità 3' del trascritto comprensiva della coda di poliadenilazione.

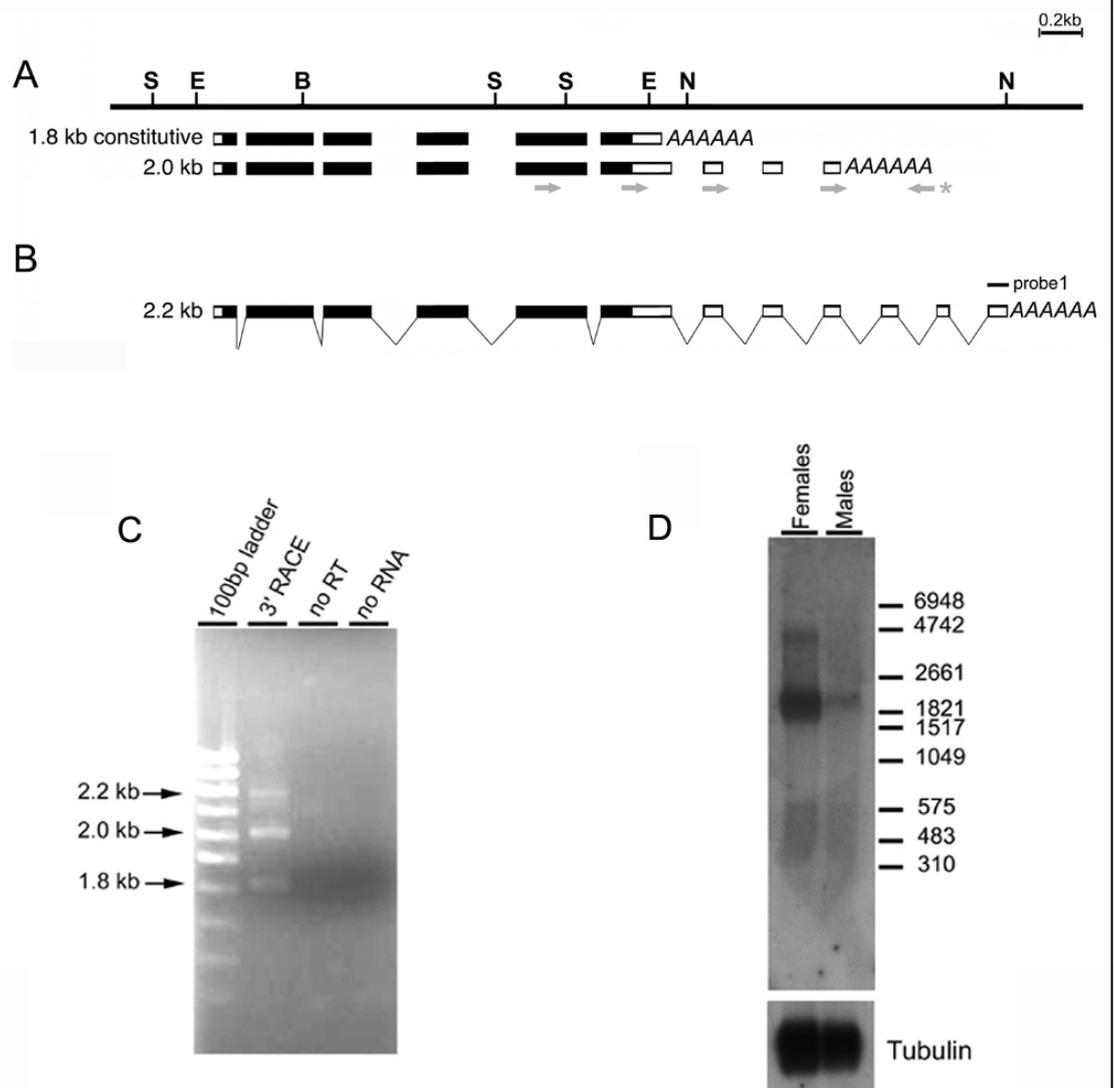


Figura 4. Organizzazione molecolare del gene *minifly*.

A) Organizzazione delle unità trascrizionali di 1.8 e 2.0 kb, precedentemente descritte da Giordano et al.,1999. Le frecce indicano i primer utilizzati negli esperimenti di 3' RACE. **B)** Organizzazione del nuovo trascritto di 2.2 kb: in nero è indicata la posizione della sonda utilizzata nell'esperimento di Northern blot indicato in **(D)**. Le regioni esoniche che compongono la regione codificante sono colorate in nero. **C)** 3'-RACE di preparazioni di RNA poly(A)⁺ da femmine adulte, in cui sono evidenti le tre bande che derivano dai trascritti che contengono differenti sequenze 3' UTR. In quest' esperimento è stato utilizzato un primer derivato dalla sequenza dell'esone 6, in coppia con l'oligo Adaptor (P*). **D)** Esperimento di Northern blot in cui sono state analizzate preparazioni di RNA poly (A)⁺ ottenute da femmine (F) e maschi (M) adulti con una sonda genomica che copre l'esone 12 (probe 1).

In figura 4C è riportato il risultato ottenuto dall'esperimento condotto su RNA poly(A)⁺ utilizzando per l'amplificazione il primer oligo Adaptor ed uno dei forward primer utilizzati, nel caso riportato in figura (4C) il primer è quello posizionato nell'esone 6. In questo set di esperimenti è stata sempre osservata la presenza, in aggiunta ai prodotti attesi sulla base dei due siti di poliadenilazione noti, di una banda addizionale di maggiore lunghezza che risultava marcatamente più abbondante quando veniva utilizzato come stampo una preparazione di RNA poly(A)⁺ ottenuta da femmine adulte (fig. 4C). Le bande addizionali ottenute in ciascuna reazione di 3'RACE sono state quindi clonate nel vettore plasmidico *pMOSBlue* e sequenziate. L'analisi delle sequenze nucleotidiche dei cloni ottenuti ha rivelato la presenza di un nuovo sito di poliadenilazione localizzato circa 0.8 kb più a valle del più distale dei due siti di terminazione precedentemente descritti (fig. 4B, C).

Le sequenze ottenute hanno rivelato che i trascritti che terminano in corrispondenza di questo nuovo sito condividono la stessa sequenza codificante dei due messaggeri descritti in precedenza, differendo quindi solo per la presenza di una regione 3' UTR più lunga che contiene tre esoni aggiuntivi (gli esoni 10, 11 e 12; fig. 4B). Le analisi di Northern blot di preparazioni di RNA poly(A)⁺ con una sonda genomica che copre uno dei nuovi esoni identificati in 3' hanno successivamente confermato la presenza di un trascritto addizionale, lungo 2.2 kb, particolarmente abbondante nelle femmine (fig. 4D). È da notare che in questo stesso esperimento è stato anche osservato un trascritto più lungo, la cui dimensione è circa 4.4 kb. Tale trascritto era già stato notato in precedenti esperimenti di Northern blot in cui erano state utilizzate sonde

corrispondenti ai primi esoni del gene *mfl*; tuttavia, non si riuscì ad isolare alcun clone di cDNA rappresentativo di questa subforma (Giordano et al., 1999). Anche il sequenziamento dei cloni ottenuti negli esperimenti di 3'-RACE o negli esperimenti di RT-PCR (descritti successivamente) non ha contribuito a fornire ulteriori informazioni, così che la struttura di questo RNA resta ancora oggi ambigua e non si può escludere che possa corrispondere ad un precursore non completamente maturato.

L'mRNA da 2.2kb, come quello di 2.0 kb precedentemente descritto, mostra un'espressione differenziale fra i due sessi. E' interessante notare che l'identificazione di questo nuovo messaggero porta a 3 il numero dei diversi trascritti prodotti dal gene *mfl* capaci di codificare per la stessa proteina, suggerendo la possibilità che le differenti regioni 3' UTR che caratterizzano questi trascritti potrebbero svolgere un differente ruolo funzionale.

Poiché è noto che le sequenze 3'UTR contengono spesso elementi di controllo agenti in *cis* che possono influenzare la stabilità, la localizzazione e l'efficienza della traduzione dell'mRNA (per una recente review vedi Hughes, 2006), ho controllato l'eventuale presenza di sequenze funzionali note all'interno delle 3'UTR che distinguono gli mRNA di 1.8, di 2.0 e di 2.2 kb (fig. 5), utilizzando il data base UTRsource (al sito www.ba.itb.cnr.it/UTR), che raccoglie informazioni specifiche per le regioni 3' e 5'UTR. Tuttavia, nessun motivo funzionale noto influenzare la stabilità dei messaggeri è stato identificato all'interno di queste regioni. Anche se non si può escludere la presenza di siti regolativi non ancora identificati, è tuttavia possibile ipotizzare che la diversa lunghezza di queste sequenze (fig. 5) possa di per se stessa rappresentare un

segnale importante per determinare una diversa efficienza traduzionale di questi tre messaggeri. Infatti, è noto che il gene *mfl* appartiene alla classe dei geni 5' TOP (Giordano et al., 1999, Tycowski et al., 1998), cioè dei geni che

Sequenza della regione 3'-UTR del trascritto costitutivo di 1.8 kb (373 nt):

CCAGCGUUGAGGAAACUGCUGCGGGCGGCGGUAUCCGAGGAGACUCCUCCAAGGACAAAAAGAA
 GAAGAAAAAGAAGCACAAAGGGCGACGAGGAAGCCCCAGAAGCCGCUGAGGAGGAGGCAGAGCCA
 GUAGAAAAGGAGAAGAAAAAGAAGAAGAAGGACAAGGAUAGGGACAGAGACGAAGCUCAGG
 AAUAGGAAUCAGAAGUUUCUAUAGGAUUAGGUUUAAAAGAGAAUUCUACGCUUUUUUGUAGAGGC
 ACCUACCGAUUGCUAGUGUUAGGUUUAAGGAGUGUUUUUAUGUGUAGAAGUGCAUACAAUUACA
 AAUGCAUUUAAGUGACAUCGCAAUGCAAUAAAUGAACAUUUUUAUGUGGAAACUCUAUCGCUAG

Sequenza della regione 3'-UTR del trascritto femminile di 2.0 kb (602 nt):

CCAGCGUUGAGGAAACUGCUGCGGGCGGCGGUAUCCGAGGAGACUCCUCCAAGGACAAAAAGAA
 GAAGAAAAAGAAGCACAAAGGGCGACGAGGAAGCCCCAGAAGCCGCUGAGGAGGAGGCAGAGCCA
 GUAGAAAAGGAGAAGAAAAAGAAGAAGAAGGACAAGGAUAGGGACAGAGACGAAGCUCAGG
 AAUAGGAAUCAGAAGUUUCUAUAGGAUUAGGUUUAAAAGAGAAUUCUACGCUUUUUUGUAGAGGC
 ACCUACCGAUUGCUAGUGUUAGGUUUAAGGAGUGUUUUUAUGUGUAGAAGUGCAUACAAUUACA
 AAUGCAUUUAAGUGACAUCGCAAUGCAAUAAAUGAACAUUUUUAUGUGGAAACUCUAUCGCUAG
 UUUCUUAGGUCUUAGCUUUUCCUUAGGUUUUUCAGAUAGAAAGGCAGGGAAUACAGACACAGCAA
 AGGCUUGGAAUGCUGCUCAGAGGACAAAGCGGGAGUGACGGGAGAUACGGGAUCGGCGGAACGA
 AAACUAAAAACGUGCUCGAAUAAAACAUAGGAAUUAAGGUAAGAUUUGUAGUUAAGGUUAUGCAA
 UUAUGGACUAUAACUUUAAU

Sequenza della regione 3'-UTR del nuovo trascritto femminile di 2.2 kb (958 nt)

CCAGCGUUGAGGAAACUGCUGCGGGCGGCGGUAUCCGAGGAGACUCCUCCAAGGACAAAAAGAA
 GAAGAAAAAGAAGCACAAAGGGCGACGAGGAAGCCCCAGAAGCCGCUGAGGAGGAGGCAGAGCCA
 GUAGAAAAGGAGAAGAAAAAGAAGAAGAAGGACAAGGAUAGGGACAGAGACGAAGCUCAGG
 AAUAGGAAUCAGAAGUUUCUAUAGGAUUAGGUUUAAAAGAGAAUUCUACGCUUUUUUGUAGAGGC
 ACCUACCGAUUGCUAGUGUUAGGUUUAAGGAGUGUUUUUAUGUGUAGAAGUGCAUACAAUUACA
 AAUGCAUUUAAGUGACAUCGCAAUGCAAUAAAUGAACAUUUUUAUGUGGAAACUCUAUCGCUAG
 UUUCUUAGGUCUUAGCUUUUCCUUAGGUUUUUCAGAUAGAAAGGCAGGGAAUACAGACACAGCAA
 AGGCUUGGAAUGCUGCUCAGAGGACAAAGCGGGAGUGACGGGAGAUACGGGAUCGGCGGAACGA
 AAACUAAAAACGUGCUCGAAUAAAACAUAGGAAUUAAGGUAAGAUUUGUAGUUAAGGUUAUGCAA
 UUAUGGACUAUAACUUUAAUCUUGGAAUGCUGCUCAGAGGACAAAGCGGGAGUGACGGGAGAU
 ACGGGAUCGGCGGAACGAAAACUAAAAACGUGCUCGAAUAAAACAUAGGAAUUAAGGUAAGAUUU
 GUUGAAUUAAGGUUAUGCAAUUAUGGACUAUAACUUUAAUAGGUUCAAUUUUUUUUUGAAAAUC
 AUUGAAUGCCACAGGAUCUUAAAAAGUUCGCAUGGCAUUUUUAUGCGUUGAGUAAAACAAGUUGUG
 CUUUUGGCGACUUUCCACUGACUGUCCAAUUGAGUUCCAUUUCAUGCCUUUUUUUAGGUACAA
 CACUGCCCAAGACAUCAUAUAUACUUUAUAAGCAAUAAAAGAAACAUAAC

Figura 5. Sequenza delle regioni 3'UTR, parzialmente sovrapposte, che caratterizzano gli mRNA di 1.8, 2.0 e 2.2 kb. La sequenza indicata in rosso rappresenta la 3'UTR del trascritto costitutivo di 1.8 kb, quella in blu la sequenza aggiuntiva presente nel trascritto di 2.0 kb e quella in nero la sequenza presente esclusivamente nella 3' UTR del trascritto di 2.2 kb.

codificano per proteine implicate nella biogenesi del ribosoma o nella funzionalità dell'apparato traduzionale. Dal punto di vista strutturale, gli

mRNA 5'TOP condividono un'organizzazione comune delle loro sequenze 5'-UTR. Queste sequenze sono sempre piuttosto corte ed iniziano invariabilmente con una C seguita da un tratto di 6-10 pirimidine (5'-TOP = 5'-Terminale di OligoPirimidine), da cui deriva la sigla 5'-TOP che identifica questa classe di geni. La sequenza 5'TOP svolge un importante ruolo regolativo a livello traduzionale: infatti, l'efficienza della traduzione degli mRNA 5' TOP si adatta alle diverse condizioni di crescita, così che in cellule in attiva proliferazione questi messaggeri sono localizzati sui polisomi, mentre nelle cellule quiescenti o in arresto della crescita essi sono per lo più sequestrati in complessi mRNP inattivi (Amaldi et Pierandrei-Amaldi, 1997).

Nelle sequenze 5'UTR degli mRNA delle proteine ribosomali (rp-mRNAs), la sequenza 5'TOP è l'elemento in *cis* responsabile della regolazione traduzionale. Tuttavia, è stato recentemente osservato che i messaggeri che codificano le proteine ribosomali sono caratterizzati dall'aver una sequenza 3'UTR insolitamente corta, in genere inferiore ai 100 nucleotidi (Ledda et al., 2005). Ledda e collaboratori (2005) hanno dimostrato che, mentre la 5'UTR contenente la sequenza TOP è sufficiente a conferire una regolazione traduzionale associata alla crescita cellulare, la stringenza di tale regolazione sembra diminuire all'aumento della lunghezza della 3'UTR. La lunghezza della 3'UTR non appare invece influenzare la traduzione degli mRNA che hanno 5'UTR prive della sequenza TOP, che rimangono per lo più associati con i polisomi sia in condizioni di crescita che in quelle di arresto della crescita (Ledda et al., 2005). Sulla base di questi dati, è possibile ipotizzare che per i trascritti di 1.8, 2.0 e 2.2 kb la lunghezza, e non la sequenza della 3'UTR,

potrebbe di per sé risultare un elemento importante per influenzare l'efficienza di traduzione dei rispettivi messaggeri in differenti condizioni di crescita. Ulteriori ipotesi sul ruolo funzionale svolto dalle sequenze 3'UTR che caratterizzano i trascritti da 1.8, 2.0 e 2.2 kb potrebbero anche essere basate sul modello della traduzione a "cappio chiuso", secondo il quale le proteine associate con le estremità 5' e 3' di un mRNA interagiscono fra loro per facilitare l'inizio della traduzione. Un ovvio ruolo per queste interazioni proteina-proteina è che la conseguente connessione tra le estremità 5' e 3' dell'mRNA può selezionare gli mRNA integri come stampo per la traduzione, o proteggerli dalla degradazione. Tuttavia, le sequenze 3' UTR potrebbero anche influenzare l'efficienza di reinizio della traduzione di uno stesso mRNA, poiché le subunità ribosomali potrebbero essere rilasciate più o meno vicino al sito di inizio. Infine, un altro possibile ruolo regolativo svolto dalle differenti 3' UTR che caratterizzano i tre messaggeri del gene *mfl* riguarda la loro eventuale interazione con le molecole di microRNA. I microRNA (miRNA) sono piccole molecole di RNA, lunghe circa 20-22nt, che svolgono un'importante ruolo regolativo. Infatti, legandosi per complementarità anche imperfetta ai trascritti bersaglio, essi ne reprimono la traduzione o ne inducono la degradazione se la complementarità è perfetta (Lai EC, 2005). Negli ultimi anni, le molecole di microRNA sono risultate implicate in numerosi processi cellulari, quali lo sviluppo, il differenziamento, la proliferazione cellulare e l'apoptosi ed è ormai ben accertato che esse svolgono funzioni essenziali nell'organogenesi (per una review, vedi Zhang et al., 2006). Poiché il riconoscimento tra un miRNA ed un mRNA bersaglio prevede solo una corta

sequenza di complementarità, che può anche essere imperfetta, un singolo miRNA può controllare la traduzione di un gran numero di messaggeri. Quindi, anche se i geni per i microRNA rappresentano solo dall'1% al 3% del genoma dei diversi organismi, è stato stimato che questa percentuale potrebbe essere sufficiente per controllare la traduzione di più di un terzo di tutti gli mRNA cellulari (Enright et al., 2003). E' inoltre noto che ogni messaggero può essere bersaglio di uno o più microRNA, rendendo questo meccanismo di regolazione post-trascrizionale molto complesso.

Come primo approccio a questo tipo di problematica ho voluto verificare se le 3'UTR caratteristiche di ciascun trascritto del gene *mfl* (riportate in fig. 5 e 6 rispettivamente nei colori rosso, blu e nero) potessero risultare potenziale bersaglio per i microRNA finora identificati in *Drosophila*. Per questo tipo di analisi ho utilizzato un programma bioinformatico disponibile in rete, chiamato *MicroInspector* (<http://www.imbb.forth.gr/microinspector>), che permette di analizzare una sequenza definita di RNA, quale quella di un RNA messaggero o parti di esso, identificandone all'interno l'esistenza di putativi siti di legame per i miRNA (Rusinov et al., 2005). I risultati di questa ricerca, riportati nella figura 6, fanno riferimento alla temperatura di 22°C, ma sono estendibili a tutte le temperature comprese nell'intervallo consigliato (Rusinov et al., 2005). Come si può notare, per ciascun appaiamento è riportato il valore di energia libera (ΔG). Sebbene non esista un intervallo di valori di ΔG che possa essere considerato di per sé sufficiente per dimostrare la reale validità dell'appaiamento considerato, va sottolineato che i valori riportati per ciascun appaiamento sono dello stesso ordine di grandezza di quelli osservati per molti

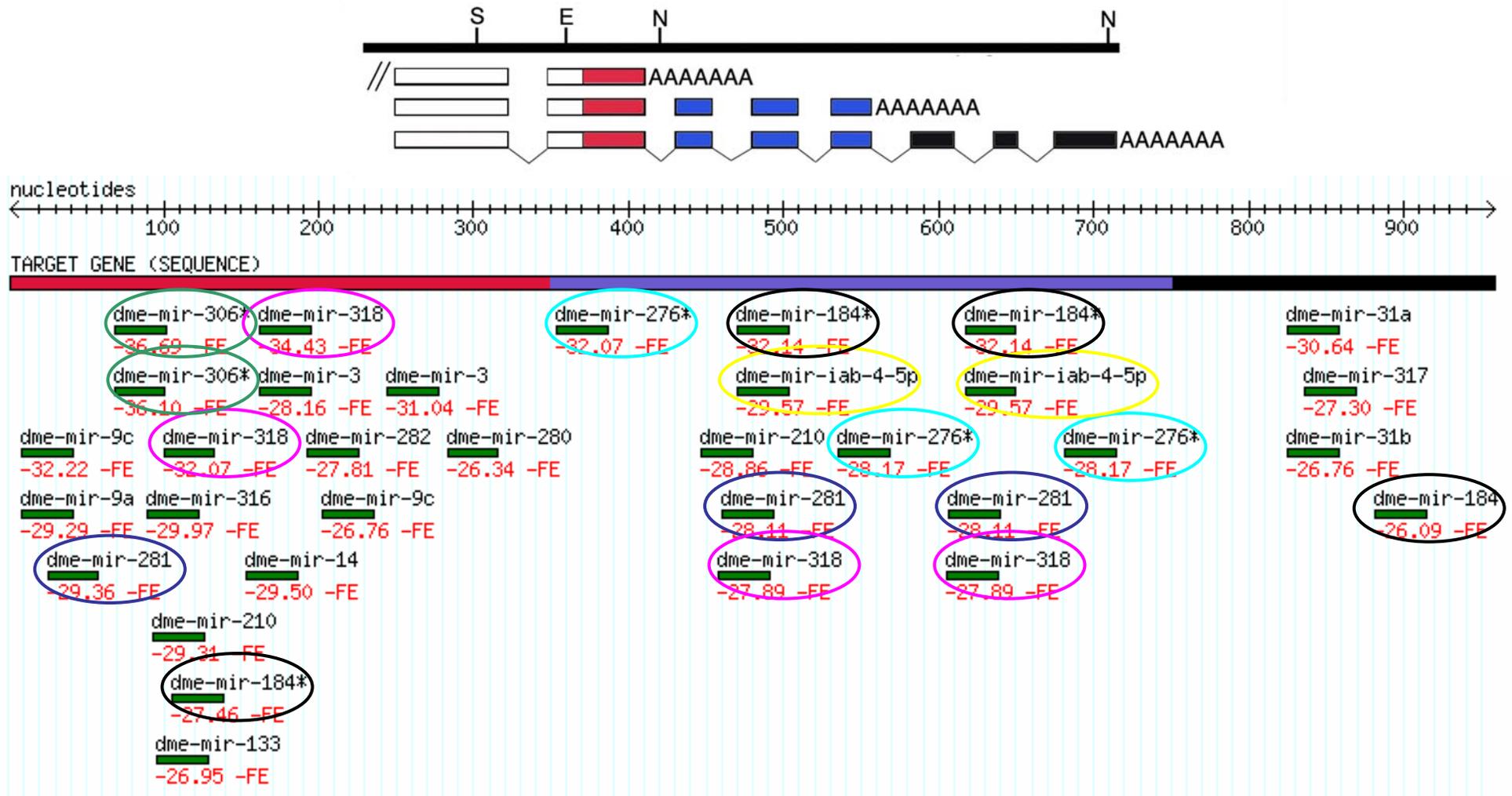
appaiamenti miRNA-target che sono stati validati *in vivo*. Un aspetto interessante dei dati ottenuti è la presenza sulla medesima sequenza 3'UTR di siti multipli di riconoscimento per lo stesso microRNA, un dato che è generalmente considerato indicativo di una reale interazione *in vivo*. Nel caso della regione 3' UTR dell'mRNA di 1.8 kb (riportata in rosso nella figura 6), 2 siti di legame sono stati identificati nella sequenza per il miRNA 318 (cerchiato in rosa) e per il miRNA 306 (cerchiato in verde). A questi siti si aggiungono, all'interno della 3' UTR dell'mRNA di 2.0 kb (in blu), 2 siti di legame per il miRNA 184 (cerchiato in nero), per il miRNA 281 (cerchiato in blu) ed il miRNA iab-4-5p (cerchiato in giallo), nonché 3 siti di legame per il miRNA 276 (cerchiato in azzurro). Inoltre, nella sequenza 3'UTR dell'mRNA di 2.0 kb i siti di legame per il miRNA 184 (cerchiato in nero) salgono a tre e quelli per il miRNA 318 (cerchiato in rosa) salgono a 4. Infatti, all'unico sito per il miRNA 184 presente nella regione condivisa con l' mRNA di 1.8 kb si aggiungono i due siti localizzati all'interno del tratto specifico per questo trascritto, e ai due siti presenti nella 3'UTR dell'mRNA da 1.8 kb per il miR 318 se ne aggiungono altri due. Infine, è interessante notare che nella 3' UTR dell'mRNA da 2.2 kb (riportata in nero nella figura 6) il numero dei siti di legame per il miR 184 sale a 4.

Fra i miRNA potenzialmente capaci di legare in più posizioni le 3'UTR dei trascritti del gene *mfl* risulta evolutivamente conservato anche in mammiferi solo miR 184, mentre gli altri sono stati finora descritti solo in insetti. Fra di essi, il miRNA iab-4-5p, è quello meglio caratterizzato a livello funzionale, ed è stato descritto che la sua espressione ectopica nei dischi immaginali delle

altere inibisce *in vivo* l'attività di *Ultrabitorax* (Ubx) ed induce una classica trasformazione omeotica delle altere in ali (Ronshaugen et al., 2005).

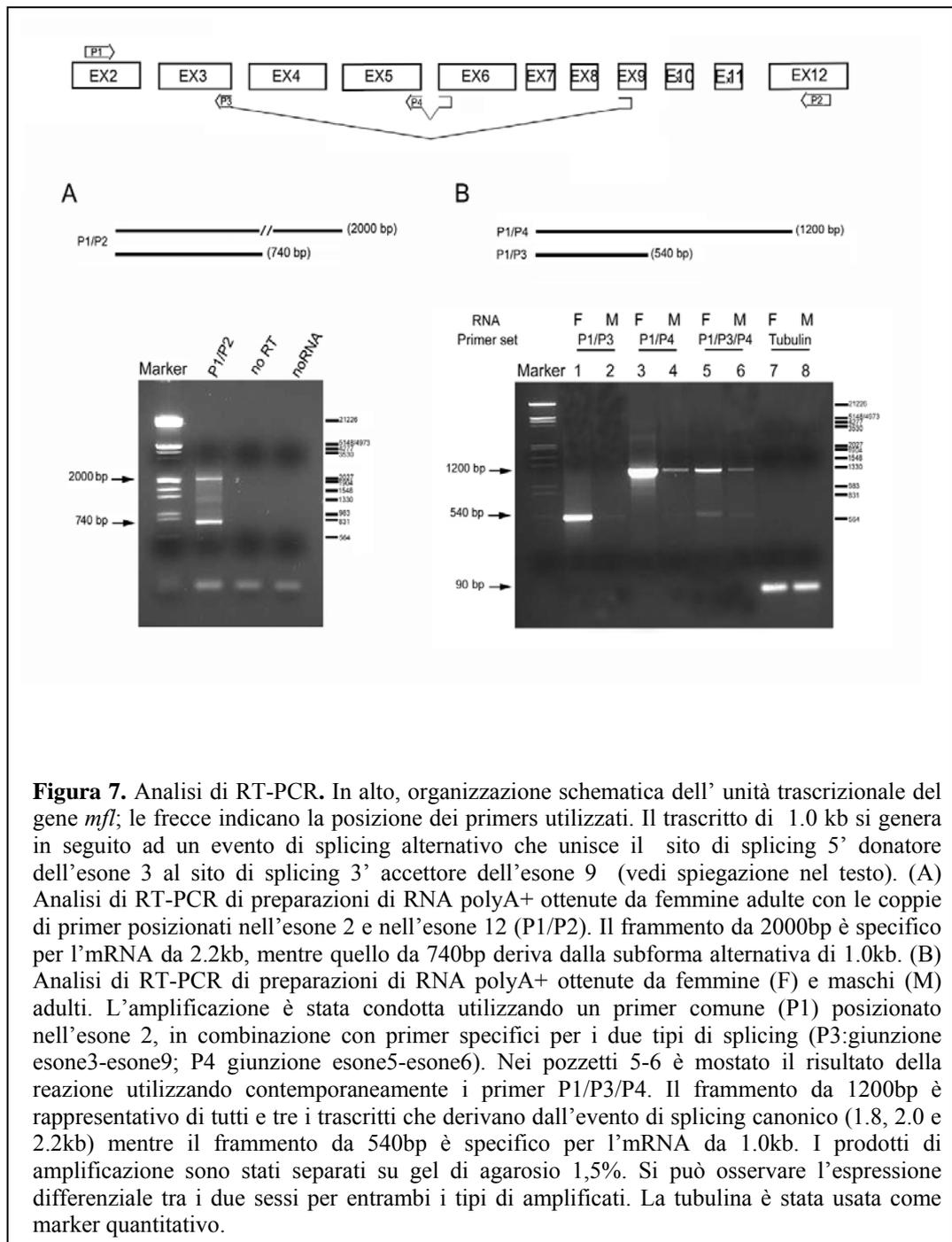
I dati ottenuti in questa ricerca bioinformatica forniscono un interessante punto di partenza per analisi più approfondite e per ispirare successivi esperimenti.

Figura 6. Siti di potenziale legame per microRNA all'interno delle regioni 3'UTR degli mRNA da 1.8, 2.0 e 2.2 kb



Il gene *mfl* può codificare una nuova subforma proteica derivata da un raro pattern di splicing alternativo

Allo scopo di verificare se i trascritti che terminano al nuovo sito di poliadenilazione identificato nel corso degli esperimenti precedenti potessero essere caratterizzati da eventuali eventi di splicing alternativo ho successivamente eseguito esperimenti di RT-PCR utilizzando “reverse primers” derivati dai nuovi esoni identificati in 3’ (gli esoni 10, 11, 12; vedi Materiali e Metodi) e “forward primers” derivati dalla sequenza degli esoni più a monte (gli esoni 2, 5 e 6). In questi esperimenti sono state utilizzate come stampo preparazioni di RNA poly(A)⁺ ottenute da moscerini adulti, separati in base al sesso, e diverse combinazioni di “forward” e “reverse” primers. In questi esperimenti, i primer derivati dagli esoni 5 e 6, in ogni combinazione con i diversi “reverse” primer utilizzati, hanno sempre generato un singolo prodotto di amplificazione, marcatamente più abbondante in femmine. Le analisi di sequenza dei prodotti ottenuti hanno confermato la struttura ottenuta nei precedenti esperimenti di 3’RACE, cioè tipica del trascritto di 2.2 kb. Al contrario, quando è stato usato il primer derivato dall’esone 2 in combinazione con i primer derivati dagli esoni 10, 11 e 12, in tutte le reazioni è stato ottenuto un prodotto addizionale, più abbondante se veniva retrotrascritta ed amplificata la preparazione di RNA poly(A)⁺ ottenuta da femmine (fig. 7A). Il clonaggio ed il sequenziamento dei prodotti addizionali ottenuti nelle varie reazioni di RT-PCR ha portato successivamente ad evidenziare la presenza di un raro evento di splicing alternativo che unisce il sito di splicing 5’ donatore dell’esone 3 al sito di splicing 3’ accettore dell’esone 9 (fig. 7A, B).



Per confermare ulteriormente la presenza di questo nuovo pattern di splicing è stato successivamente utilizzato lo stesso “forward primer” derivato dall'esone 2 (P1) in combinazione o con un “reverse primer” derivato dalla giunzione degli esoni 3-9 (P3), specifico per il nuovo evento di splicing alternativo,

oppure in combinazione con un primer derivato dalla giunzione degli esoni 5-6 (P4), che è caratteristica del pattern di splicing prevalente che è comune all'mRNA costitutivo di 1.8 kb ed ai trascritti di 2.0 e 2.2 kb. Come mostrato in fig. 7B, entrambe le coppie di primer hanno prodotto un'amplificazione positiva dei prodotti attesi, confermando la presenza di un nuovo pattern di splicing alternativo all'interno della regione esaminata. E' da notare che entrambi i prodotti di amplificazione risultano presenti a livelli significativamente più alti nelle femmine, indicando che anche il raro evento di splicing alternativo sembra esibire un profilo di espressione differente nei due sessi, come è stato già osservato per l'mRNA di 2.0 kb (Giordano et al., 1999) e per quello di 2.2 kb. Allo scopo di ottenere maggiori informazioni sulla struttura del trascritto derivato da questo nuovo evento di splicing si è quindi proceduto al clonaggio del cDNA completo, che è stato ottenuto mediante combinazione di esperimenti di 5' e 3' RACE, secondo lo schema descritto in figura 8. L'amplificazione dell'estremità 5' del trascritto è stata condotta utilizzando oligonucleotidi derivati dalla sequenza dell'esone 3, così da ottenere prodotti di amplificazione che potessero sovrapporsi parzialmente a quelli precedentemente ottenuti con gli esperimenti di 3' RACE. Negli esperimenti di 5'RACE è stato ottenuto sempre un unico tipo di amplificato che presentava le dimensioni attese in base al sito di inizio della trascrizione precedentemente mappato, confermando ulteriormente la presenza di un'unica estremità 5' per tutti i trascritti derivati dal gene. La zona di sovrapposizione presente nei prodotti di amplificazione ottenuti negli esperimenti di 5' e 3'RACE comprende una parte del secondo esone ed una parte del terzo esone,

ed è caratterizzata dalla presenza di un sito di restrizione unico riconosciuto dall'enzima batterico BamH1 (fig. 8).

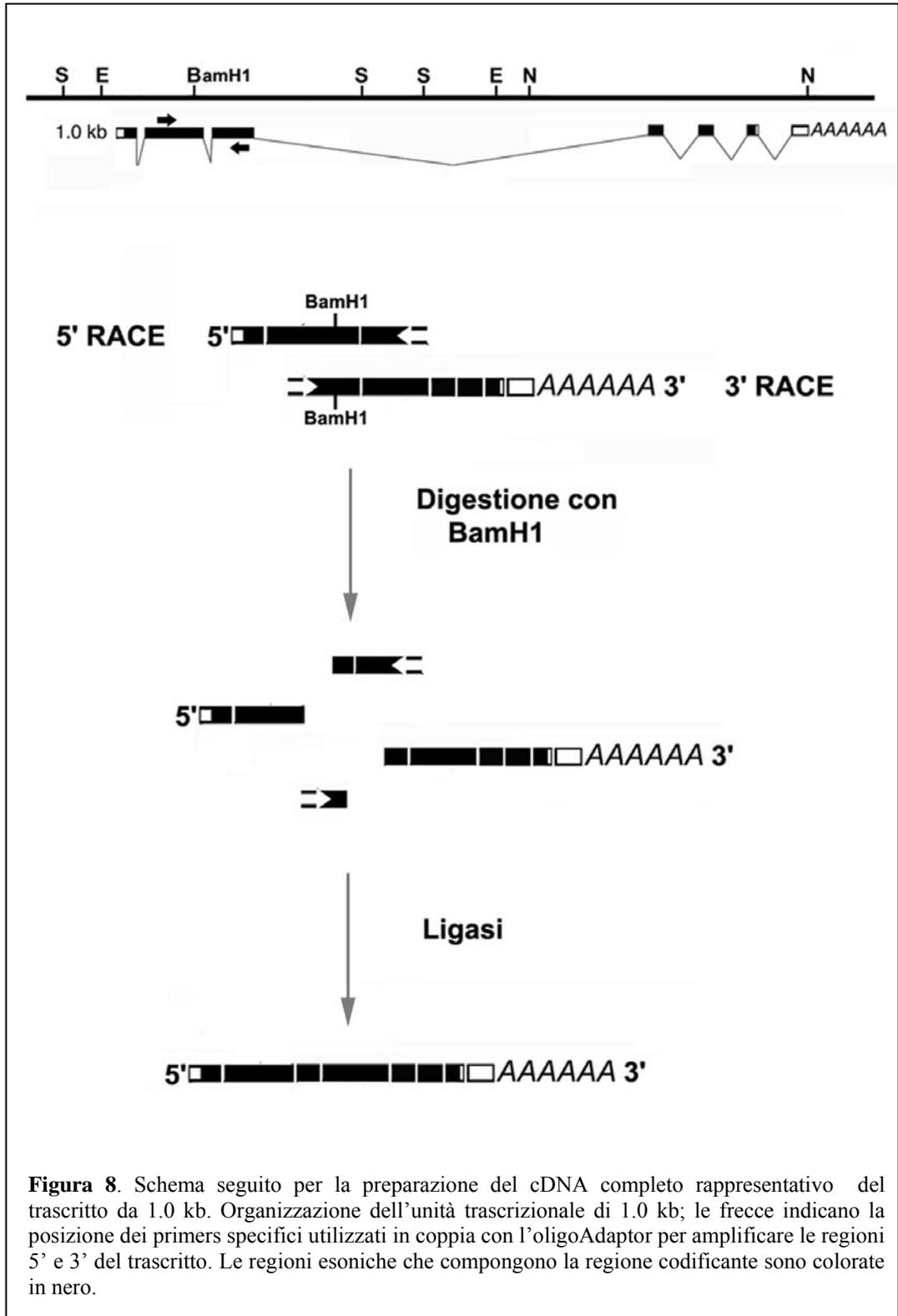


Figura 8. Schema seguito per la preparazione del cDNA completo rappresentativo del trascritto da 1.0 kb. Organizzazione dell'unità trascrizionale di 1.0 kb; le frecce indicano la posizione dei primers specifici utilizzati in coppia con l'oligoAdaptor per amplificare le regioni 5' e 3' del trascritto. Le regioni esoniche che compongono la regione codificante sono colorate in nero.

Mediante digestione con quest'enzima e successiva reazione di ligasi è stato quindi possibile ricostruire un cDNA completo. Il frammento così ottenuto mostrava una lunghezza di circa 1 kb e l'analisi della sequenza nucleotidica ha rivelato che esso era rappresentativo di un trascritto caratterizzato dall'esclusione degli esoni 4, 5, 6, 7 e 8 e potenzialmente capace di codificare una nuova subforma proteica, da me denominata MFL α (discussa in seguito).

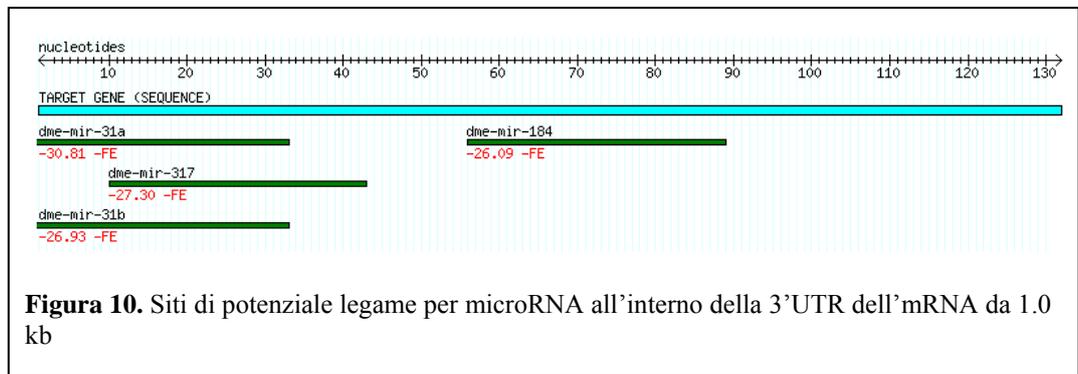
E' interessante notare inoltre che la lunghezza della sequenza 3'UTR presente nel trascritto di 1.0 kb è di soli 131 nucleotidi (fig 9). In accordo con i dati precedentemente discussi di Ledda e collaboratori (2005), la ridotta lunghezza di questa 3'UTR dovrebbe implicare una maggiore stringenza nella regolazione traduzionale di questo messaggero rispetto a quella prevista per gli mRNA di 1.8, 2.0 e 2.2 kb.

```
GUAAACAAGUUGUGCUUUUGGCGACUUCCACUGACUGUCCAAUUGAGUUCC  
AUUUCAUGCCCUUUUUUAGGUACAACACUGCCCAAGACAUCAUAUAAACUUU  
AUAAGCAAUAAAAGAAACAUACAC
```

Figura 9. Sequenza della 3'UTR dell' mRNA di 1.0 kb (lunga 131 nt)

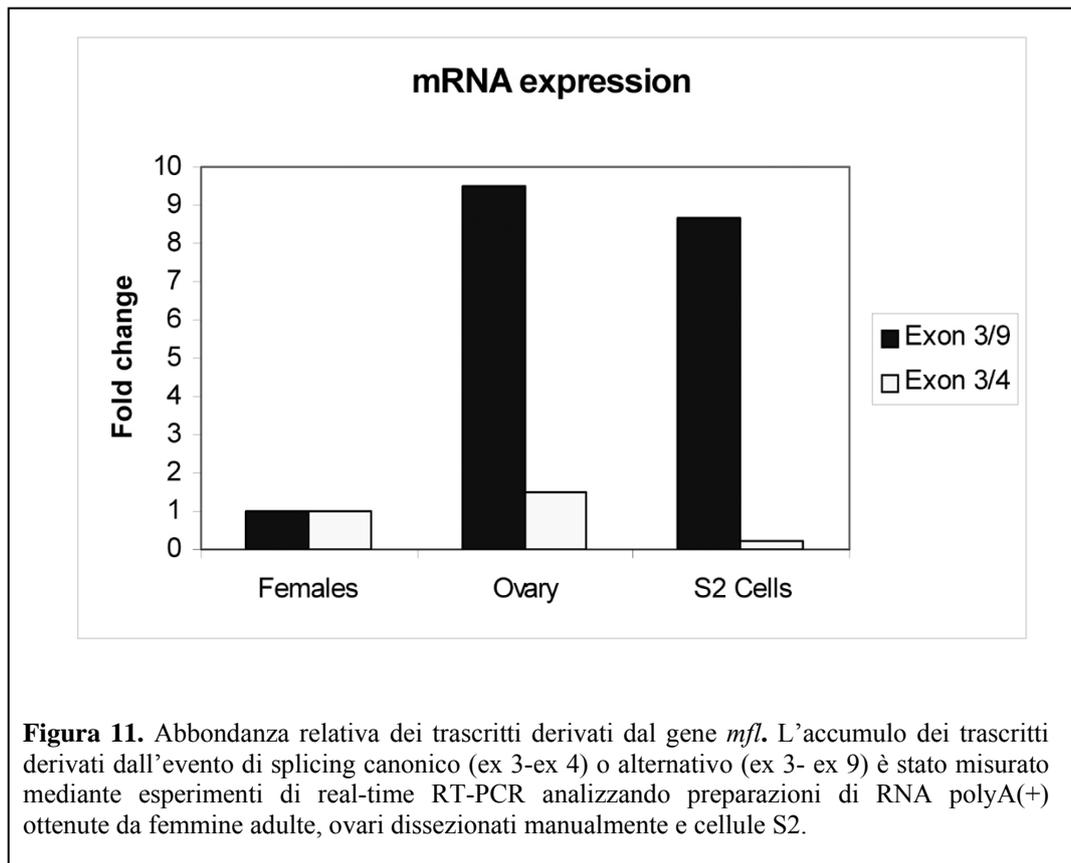
Analogamente a quanto fatto per le 3' UTR dei trascritti da 1.8, 2.0 e 2.2kb, anche questa sequenza è stata analizzata con il programma MicroInspector per la ricerca di putativi siti di legame per molecole di microRNA. Come si può notare dalla rappresentazione grafica riportata in figura 10, nessun sito di legame multiplo è presente su questa corta regione; al suo interno sono però

presenti gli stessi siti di legame precedentemente descritti per la porzione di 3'UTR specifica per il trascritto da 2.2 kb (vedi fig. 6), dato che queste due 3'UTR sono in parte coincidenti.



Il nuovo mRNA identificato negli esperimenti di RT-PCR appare accumularsi a livelli molto bassi, dato che nei precedenti esperimenti di Northern blot, anche analizzando preparazioni di RNA poly(A)+ ottenute da femmine adulte, esso non è stato mai visualizzato. Il reale accumulo di questo trascritto è stato quindi seguito con esperimenti di Real-Time RT-PCR, analizzando preparazioni di RNA poly(A)+ ottenute da cellule di *Drosophila* in coltura della linea S2, da femmine adulte e da ovari di femmine adulte ottenuti mediante dissezione manuale. Tutti gli oligonucleotidi utilizzati come innesco in questo set di esperimenti sono stati progettati con il programma Primer Express 2.0 software (Applied Biosystems) e sottoposti ad allineamento con il BLASTN, per verificare che fossero specifici per il gene d'interesse. Per ridurre al minimo i rischi di contaminazione da DNA tutti gli oligonucleotidi sono stati posizionati o in corrispondenza di differenti esoni o tra le giunzioni esone-esone specifiche per i due tipi di splicing esaminati (3-4 o 3-9). I dati ottenuti in questi esperimenti sono stati normalizzati utilizzando come

controllo l'espressione dell'mRNA del gene dell' α *Tubulina84B*, che è espresso in maniera costitutiva durante tutto il ciclo vitale della *Drosophila*. L'abbondanza relativa del trascritto da 1.0 kb nei tre campioni è mostrata in figura 11.



Considerando come valore di riferimento per entrambi gli eventi di splicing il valore ottenuto nell'analisi delle preparazioni di RNA poly(A)+ da femmine adulte, l'accumulo dell'mRNA da 1.0 kb è risultato circa 9 volte più elevato negli ovari e nelle cellule S2. Al contrario, l'accumulo dei trascritti da 1.8, 2.0 e 2.2kb, che derivano dall'evento di splicing canonico che unisce l'esone 3 al 4, non risulta significativamente aumentato negli ovari, ed appare di poco ridotto nelle cellule S2 in attiva proliferazione.

La subforma proteica MFL α

L'mRNA da 1 kb ha la potenzialità di codificare una nuova subforma proteica, denominata MFL α . Questa nuova subforma è prevista avere una sequenza amminoacidica di 258 residui, con una massa molecolare di 28.9 kDa. La sequenza amminoacidica di questa proteina è riportata in figura 12A, dove è stata allineata alla sequenza della proteina MFL, codificata dai trascritti da 1.8, 2.0 e 2.2 kb, ed alla sequenza della discherina umana. La nuova subforma condivide con la proteina MFL precedentemente identificata un segmento di 211 aa presente nella porzione ammino terminale, in cui sono contenuti un segnale di localizzazione nucleare e due domini altamente conservati, TrubI e TrubII, che hanno una forte omologia con le tRNA pseudouridina sintetasi di lievito e batteri e che sono considerati essere direttamente coinvolti nei processi di pseudouridilazione (Koonin EV, 1996). Tuttavia, MFL α differisce dalla proteina MFL nella sua porzione carbossi-terminale, dove un piccolo tratto di 43 aa, che non presenta nessun motivo funzionalmente noto, rimpiazza una lunga porzione carbossi-terminale della proteina MFL che comprende alcuni importanti domini funzionali, quali il dominio RNA binding denominato "PUA", il dominio "tyr" ed il segnale di localizzazione nucleare bipartito presente in posizione carbossi-terminale (NLS). Sebbene sia stato osservato che l'assenza dell'NLS carbossi-terminale non abolisce la localizzazione nucleare della discherina di mammifero (Heiss et al., 1999), è possibile che l'assenza di questo segnale possa però influire sul tasso di trasporto, sulla ritenzione oppure sulla localizzazione subnucleare di questa proteina in risposta a diversi stimoli cellulari.

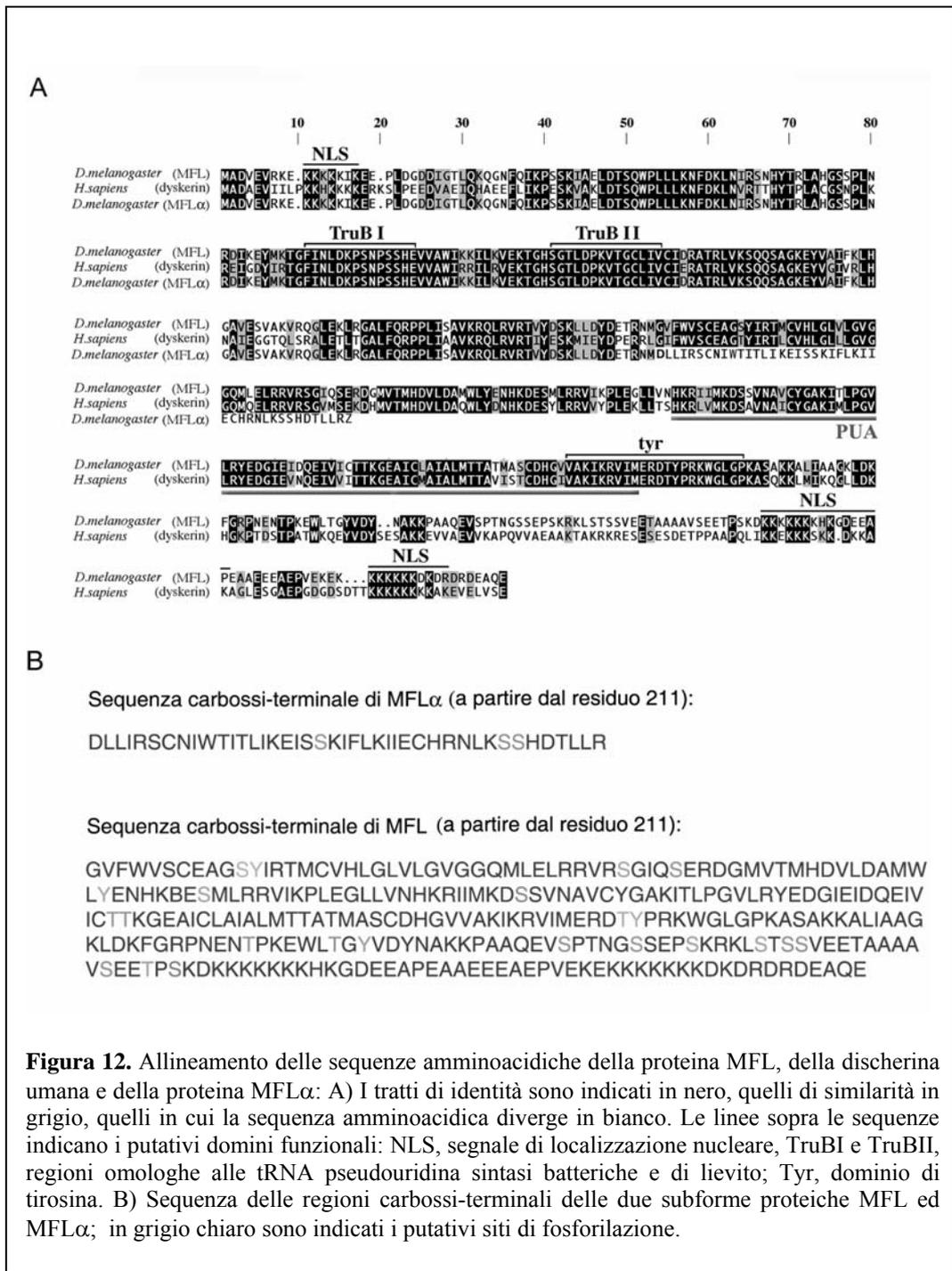


Figura 12. Allineamento delle sequenze amminoacidiche della proteina MFL, della discherina umana e della proteina MFLα: A) I tratti di identità sono indicati in nero, quelli di similarità in grigio, quelli in cui la sequenza amminoacidica diverge in bianco. Le linee sopra le sequenze indicano i putativi domini funzionali: NLS, segnale di localizzazione nucleare, TruBI e TruBII, regioni omologhe alle tRNA pseudouridina sintasi batteriche e di lievito; Tyr, dominio di tirosina. B) Sequenza delle regioni carbossi-terminali delle due subforme proteiche MFL ed MFLα; in grigio chiaro sono indicati i putativi siti di fosforilazione.

Infatti, numerose evidenze nei diversi organismi dimostrano che la porzione C-terminale delle proteine Cbf5p/MFL/discherina ha una notevole rilevanza dal punto di vista funzionale. Ad esempio, è interessante notare che nella discherina umana la delezione di un cluster ricco in lisine al carbossi-terminale influenza il tasso di trasporto della proteina nel nucleo, e l'eliminazione di un dominio basico C-terminale della proteina Cbf5p di lievito porta ad un ritardo

della fase G2/M, legando questa porzione della proteina al controllo del ciclo cellulare. Sembra plausibile, quindi, che la presenza di una porzione carbossi-terminale specifica possa conferire alla proteina MFL α una particolare localizzazione e/o specifiche proprietà funzionali.

Il fatto che la porzione C-terminale assente nella proteina MFL α sia particolarmente ricca in putativi siti di fosforilazione rappresenta un ulteriore carattere distintivo tra le due subforme (fig. 12B). Questo aspetto potrebbe suggerire la possibilità che MFL e MFL α potrebbero essere regolate in maniera differente in risposta a differenti condizioni di crescita, o durante il ciclo cellulare. Infatti, delle 7 treonine e 5 tirosine putativamente fosforilate in MFL solo un residuo è presente in MFL α , che a sua volta presenta nella sua porzione carbossi-terminale specifica tre nuovi putativi siti di fosforilazione (fig. 12B).

Poiché è noto che durante il ciclo cellulare molti processi sono controllati con una modalità dipendente dalla fosforilazione, è probabile che l'attività delle proteine MFL e MFL α possa essere differentemente modulata in risposta a questo tipo di modificazione post traduzionale. È anche possibile che la regione C-terminale che distingue MFL da MFL α possa influenzarne la stabilità, forse determinando una differente vita media alle due proteine. È infine possibile che la specifica regione C-terminale presente in MFL α possa portare questa proteina ad interagire con differente efficienza o con diversa specificità con gli altri componenti degli snoRNP di tipo H/ACA, oppure possa giocare ruoli specifici nell'assemblaggio di questi complessi. In effetti, l'assenza del dominio PUA potrebbe alterare l'interazione con le molecole di snoRNA, o con specifici sottoinsiemi di snoRNA, conferendo a questa sottoforma

differenti proprietà che potrebbero avere importanti implicazioni sulla sua funzione. Negli Archaea ed in lievito, il dominio PUA è stato dimostrato essere importante per l'assemblaggio degli snoRNP di tipo H/ACA, poiché interagisce probabilmente con la sequenza ACA delle molecole di snoRNA (Baker et al., 2005; Manival et al., 2006). Il dominio Tyr è stato, invece, mostrato essere importante per l'orientamento dell'uracile che deve essere pseudouridilato (Hoang e Ferre-D'Amare, 2001).

Allo scopo di verificare l'effettivo accumulo *in vivo* della proteina MFL α è stato quindi preparato un anticorpo policlonale indirizzato contro due peptidi della porzione ammino-terminale comune ad entrambe le proteine MFL/MFL α . Quest'anticorpo, prodotto in coniglio (vedi Materiali e Metodi), è stato saggiato in esperimenti di western blot condotti utilizzando estratti proteici ottenuti da cellule S2 di *Drosophila* (fig13). Poiché MFL α rappresenta il primo prodotto proteico alternativo finora descritto per i membri della famiglia genica *Cbf5/mfl/DKC1*, la sua caratterizzazione risultava particolarmente interessante e richiede ulteriori analisi. Inoltre, il ruolo funzionale di questa subforma potrebbe essere rilevante soprattutto nella definizione del fenotipo mutante *mfl*, dal momento che, come già detto, in precedenti esperimenti di recupero del fenotipo l'espressione di un transgene contenente il cDNA che codificava la proteina costitutiva MFL non era sufficiente per recuperare a livello funzionale le mutazioni *mfl* (Giordano et al., 1999).

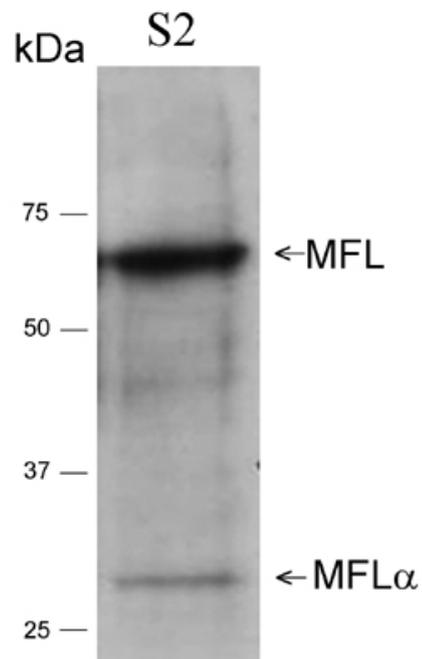


Figura 13. Analisi per western blot di estratti proteici ottenuti da cellule S2 di *Drosophila*. L'anticorpo policlonale diretto contro due peptidi della porzione ammino-terminale, comune ad entrambe le proteine MFL/MFL α , riconosce una banda di 56 kDa, rappresentativa della proteina MFL, ed una banda meno abbondante, della dimensione attesa per la subforma MFL α .

Un cluster di geni che codificano per piccole molecole di ncRNAs è localizzato negli introni 3' UTR del gene *mfl*.

Per analizzare in maggiore dettaglio la trascrizione di *mfl* sono state effettuate analisi di Northern blot anche su preparazioni di RNA totale, utilizzando sonde genomiche derivanti dalla regione 3'.

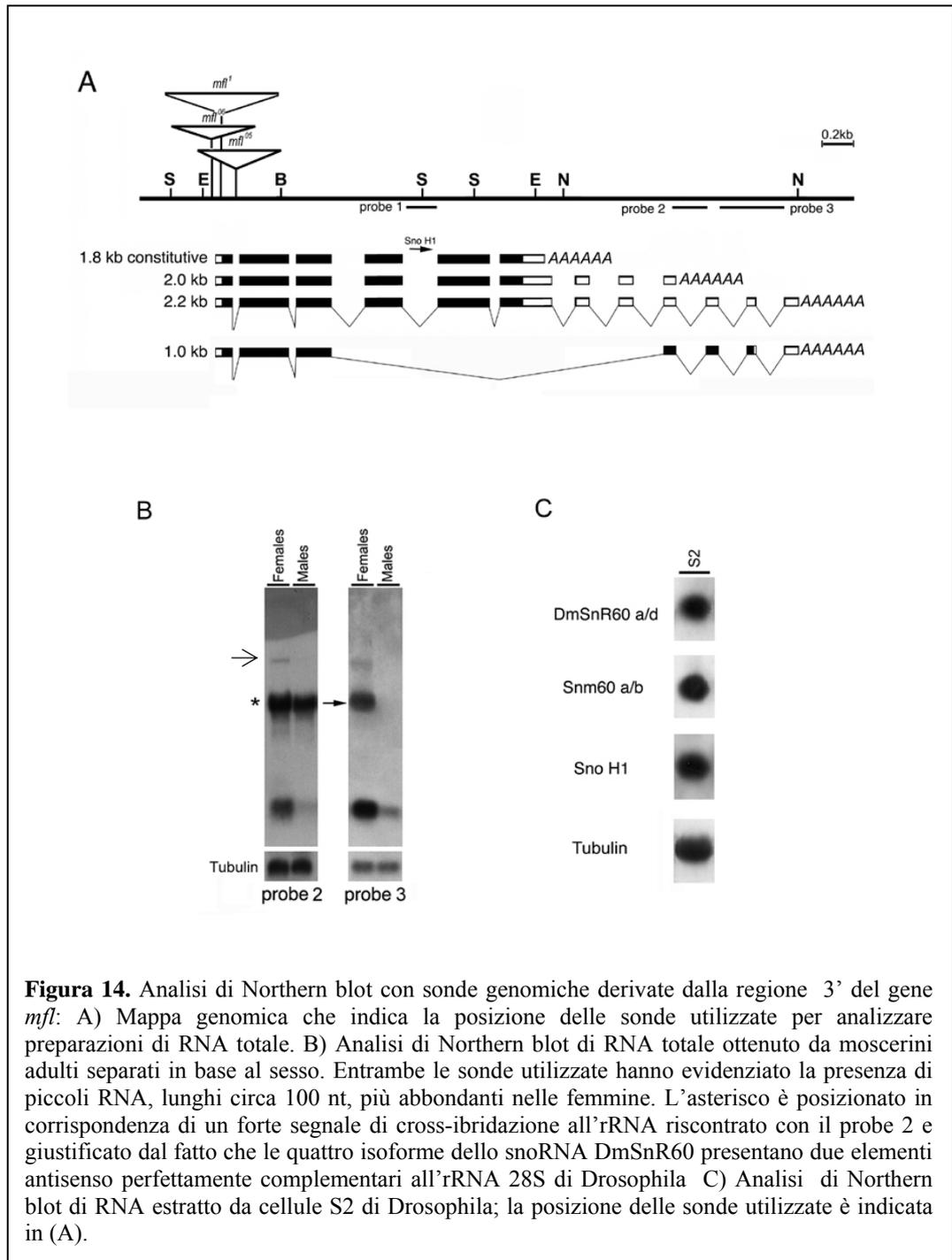
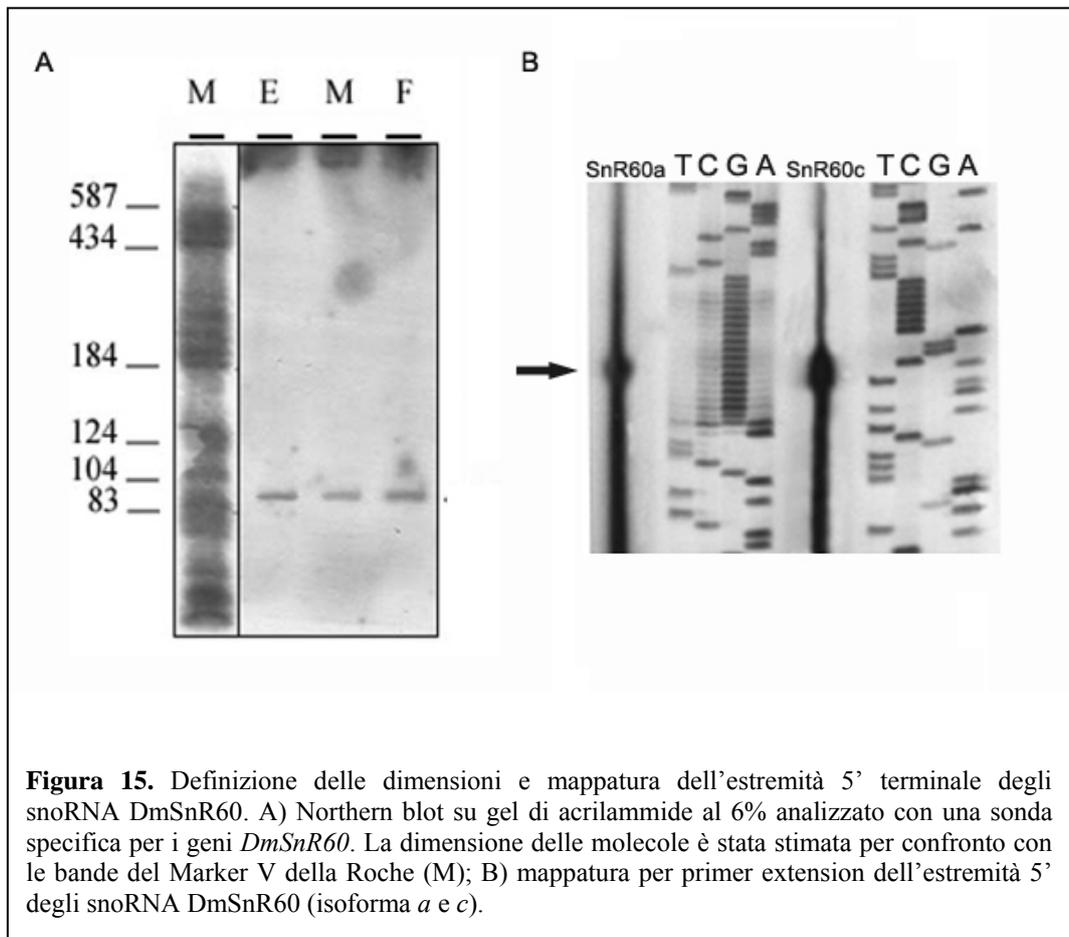


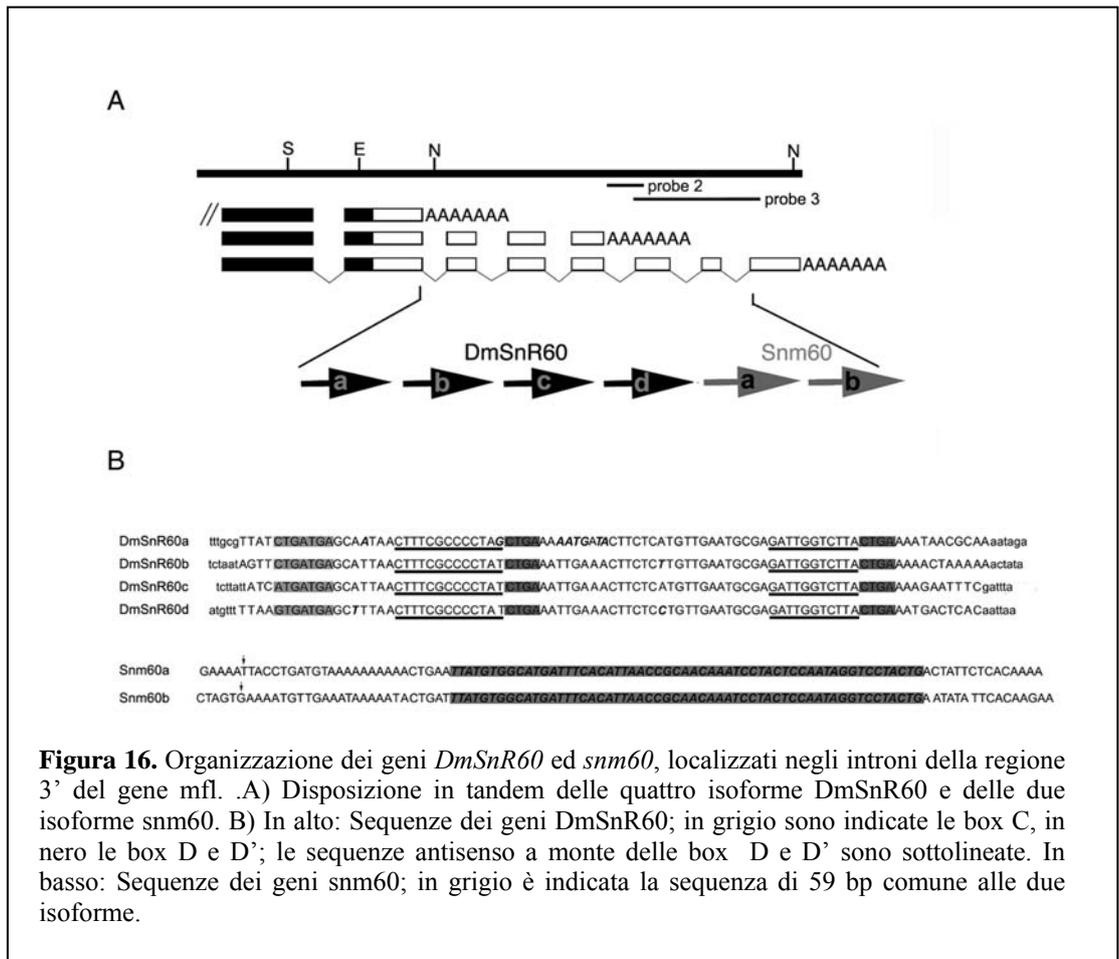
Figura 14. Analisi di Northern blot con sonde genomiche derivate dalla regione 3' del gene *mfl*: A) Mappa genomica che indica la posizione delle sonde utilizzate per analizzare preparazioni di RNA totale. B) Analisi di Northern blot di RNA totale ottenuto da moscerini adulti separati in base al sesso. Entrambe le sonde utilizzate hanno evidenziato la presenza di piccoli RNA, lunghi circa 100 nt, più abbondanti nelle femmine. L'asterisco è posizionato in corrispondenza di un forte segnale di cross-ibridazione all'rRNA riscontrato con il probe 2 e giustificato dal fatto che le quattro isoforme dello snoRNA DmSnR60 presentano due elementi antisense perfettamente complementari all'rRNA 28S di *Drosophila*. C) Analisi di Northern blot di RNA estratto da cellule S2 di *Drosophila*; la posizione delle sonde utilizzate è indicata in (A).

Come mostrato in figura 14, questi esperimenti hanno rivelato che un insieme di piccole molecole di RNA lunghe circa 100 nt deriva da questa regione. È interessante notare che queste molecole mostrano tutte un livello di espressione maggiore nelle femmine, confermando la tendenza generale che vede l'espressione del gene a livelli quantitativamente diversi nei due sessi. Per mappare con maggior precisione le sequenze genomiche da cui si originano queste piccole molecole di RNA, sonde più corte rispetto alle precedenti, a localizzazione sia esonica che intronica, sono state utilizzate per analizzare preparazioni di RNA totale ottenute dalla linea cellulare S2 di *Drosophila*. Le sonde derivate dagli introni 6, 7, 8 e 9 hanno tutte identificato un trascritto di piccole dimensioni (fig. 14), la cui lunghezza è stata stimata essere di circa 85 nt in base a successivi esperimenti di Northern blot su gel di acrilammide al 6% (figura 15 A) e di mappatura dell'estremità 5' mediante primer extension (fig. 15 B). La tecnica della primer extension si basa su una reazione di retro trascrizione e prevede l'utilizzo di un oligonucleotide marcato radiattivamente alla sua estremità 5'. Il risultante segnale di stop identifica l'estremità 5' dei trascritti con un'approssimazione di pochi nucleotidi. L'analisi della sequenza nucleotidica ha rivelato che quattro isoforme di uno snoRNA della famiglia C/D (denominata *DmSnR60 a/d*), precedentemente identificate nel corso di un'analisi genomica (Accardo et al., 2004), sono in realtà localizzate all'interno degli introni presenti nelle diverse regioni 3'UTR del gene *mfl*. Inoltre, come mostrato nella figura 14, anche le sonde derivate dagli introni 10 e 11 identificano piccole molecole di RNA. La dimensione di queste molecole



è stata però stimata essere di circa 100 nt, in base ad esperimenti di primer extension e di Northern blot su gel di acrilammide al 6%. L'analisi delle sequenze degli introni 10 ed 11 ha poi portato a stabilire che questi piccoli trascritti derivano da due copie di un gene che manca dei caratteristici motivi di sequenza e di struttura tipici delle classi di snoRNA di tipo C/D o di tipo H/ACA e che non appartiene a nessun'altra famiglia di piccoli ncRNA a funzione nota (fig. 16). Queste piccole molecole di RNA sono state quindi denominate snm60 a e b, cioè small non messenger RNA derivati dalla regione politenica 60, in accordo con una nomenclatura proposta da Huttenhofer et al. (2002). Le due copie di snm60 (a, b; numeri di accesso di GenBank

DQ142641 e DQ142642) condividono l'85% della sequenza totale e mostrano un segmento interno di 59 bp di perfetta identità (fig. 16 B, in basso). In base ad analisi bioinformatica della sequenza genomica di *Drosophila* è risultato che nessun'altra copia del gene *snm60* è presente nel genoma.



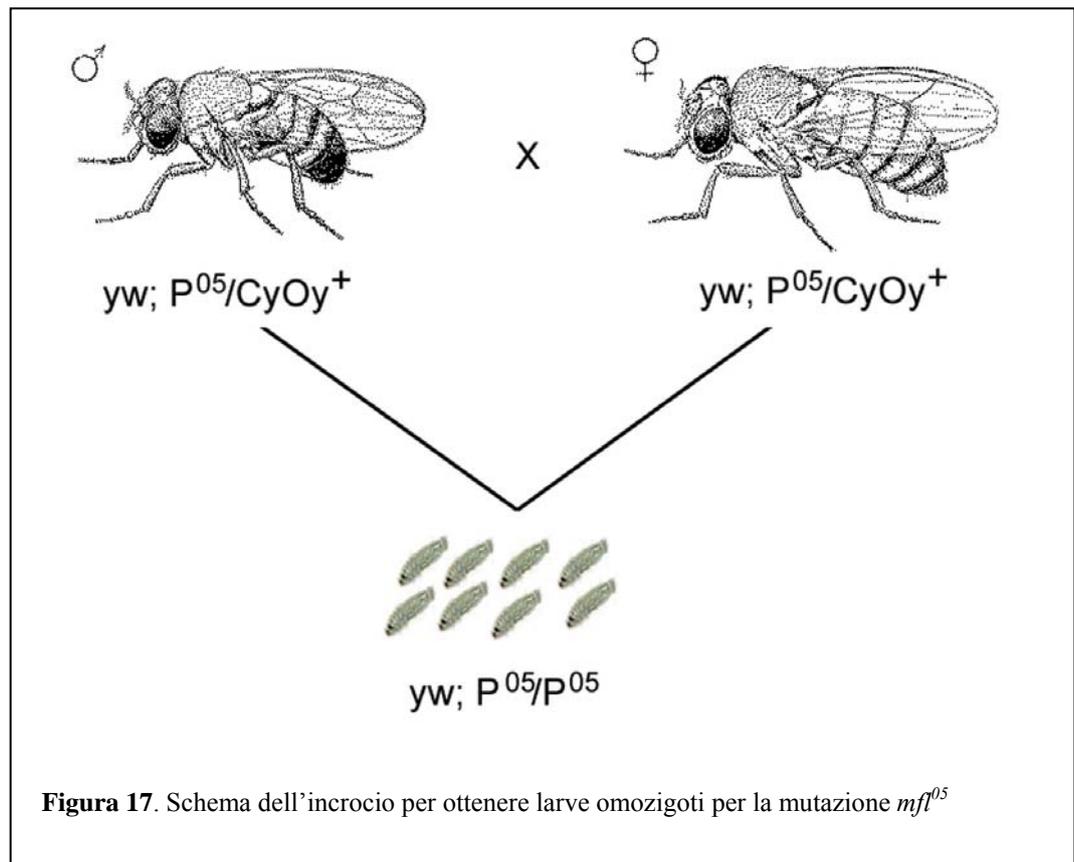
Biogenesi delle molecole DmSnR60 e snm60

E' interessante notare che la struttura del cluster di geni intronici presente nel gene *mfl* appare derivare da fenomeni di duplicazione avvenuti nel corso dell'evoluzione e si conforma alla regola di un gene per introne generalmente osservata nei genomi animali. Un'interessante conseguenza funzionale dell'arrangiamento coding/non coding mostrato dal gene *mfl* è che i suoi trascritti primari che terminano in corrispondenza dei differenti siti alternativi di poliadenilazione possono portare alla produzione di distinti sottoinsiemi di ncRNAs derivati dai diversi introni. Ad esempio, la maturazione di una molecola di pre-mRNA che termina al primo sito di poliadenilazione può portare alla produzione del solo snoRNA H1 (che è di tipo H/ACA), mentre la maturazione di una molecola che termina al sito di poliadenilazione centrale può portare anche alla produzione di tre isoforme dello snoRNA DmSnR60 (a-c), di tipo C/D. Infine, la maturazione dei trascritti primari che terminano al sito di poliadenilazione più a valle può portare alla produzione di tre differenti tipi di piccoli ncRNA, cioè di snoH1 e di tutte le isoforme DmSnR60(a-d) e snm60 (a-b). L'uso differenziale dell'estremità 3' permette quindi al gene *mfl* non solo di produrre due distinte subforme proteiche, ma anche tre differenti tipi di ncRNA che potrebbero potenzialmente contribuire alle funzioni biologiche di questo gene.

Allo scopo di verificare che le piccole molecole di RNA identificate negli introni della regione 3' del gene *mfl* fossero effettivamente prodotte in sincronia con la trascrizione del gene ospite ho voluto analizzare il profilo di espressione dei geni *snoH1*, *DmSnR60 (a-d)* e *snm60 (a-b)* nel ceppo mutante

*mfl*⁰⁵, in cui, come precedentemente descritto, l'inserzione di un singolo elemento trasponibile P nella regione 5' provoca una forte riduzione dei livelli di trascrizione. In accordo con il comune meccanismo di biogenesi dipendente dallo splicing (per una review Kiss T, 2006), l'espressione dei geni intronici è prevista essere coordinata con quella del gene ospite, cosicché l'accumulo di tutte le molecole di ncRNA derivate dalla maturazione degli introni era attesa essere fortemente ridotta in questo ceppo mutante.

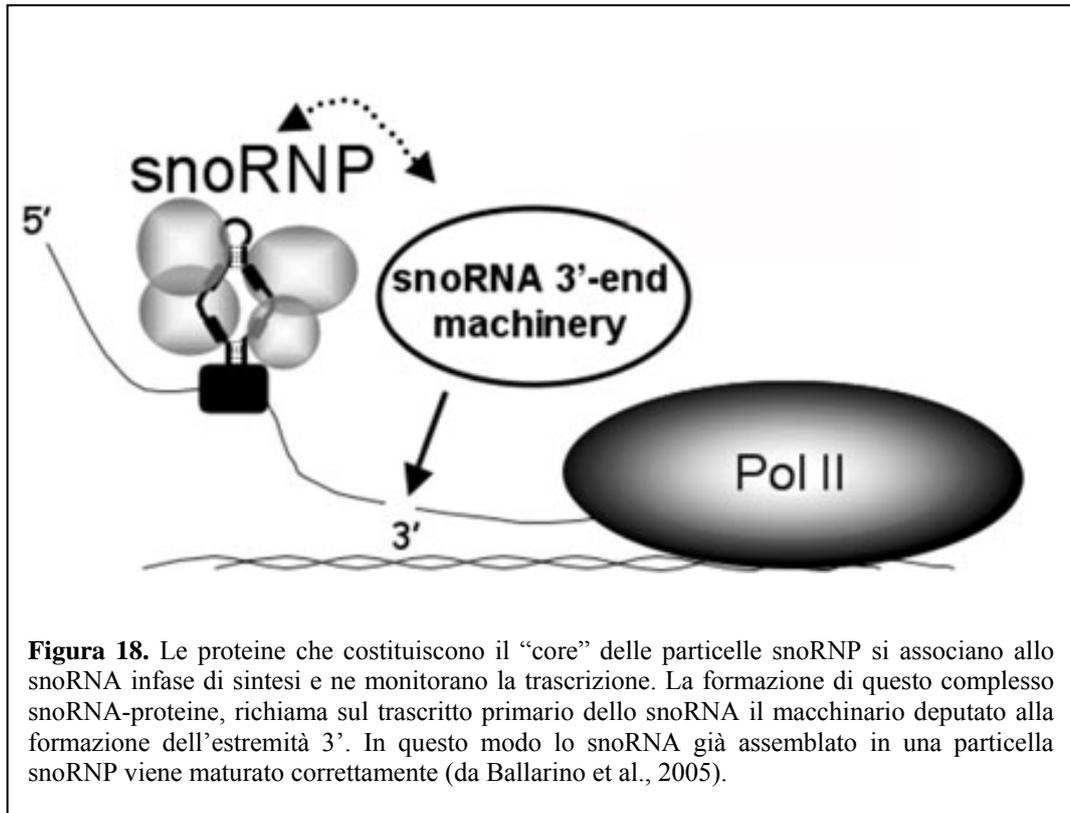
Poiché la mutazione *mfl*⁰⁵ provoca letalità tra il secondo ed il terzo stadio larvale, l'RNA è stato estratto da larve di primo stadio omozigoti per la mutazione, selezionate grazie alla presenza dell'espressione fenotipica del marcatore *yellow* (*y*) (fig. 17).



In successivi esperimenti di Northern blot, la preparazione di RNA così ottenuta è stata analizzata in parallelo con una preparazione di controllo, estratta da larve selvatiche nel medesimo stadio di sviluppo, utilizzando sonde genomiche corrispondenti agli introni 4, 9, 10-12 (fig 14). Le quantità di RNA dei due campioni sono state controllate utilizzando una sonda specifica per il trascritto del gene dell'*αTubulina84B*, che viene espresso in maniera costitutiva durante il ciclo vitale della *Drosophila*. Come mostrato in figura 19, nei mutanti *mfl*⁰⁵ l'accumulo delle molecole DmSnR60 risulta significativamente meno ridotto di quello dello snoH1, che deriva dalla maturazione del quarto introne. Questo risultato era in realtà atteso. Infatti, lo snoH1 appartiene alla classe degli snoRNA di tipo H/ACA i cui membri dipendono per la loro biogenesi e la loro stabilità dalla proteina MFL, la cui sintesi è drasticamente ridotta nel mutante *mfl*⁰⁵ (Giordano et al., 1999). Anche in Archaea, lievito e topo (Lafontaine et al., 1998; Mochizuki et al., 2004; Barth et al., 2005) è ben noto che l'inattivazione dei geni ortologhi, denominati rispettivamente *Cbf5* e *DKC1*, provoca una drastica riduzione dell'accumulo degli snoRNA di tipo H/ACA ma non di quelli di tipo C/D.

In lievito è stato stabilito che la proteina Cbf5p partecipa attivamente alla biosintesi degli snoRNA di tipo H/ACA, innescando l'assemblaggio cotrascrizionale delle snoRNP di tipo H/ACA sullo snoRNA nascente (Ballarino et al., 2005; fig. 18). La stretta correlazione tra l'accumulo degli snoRNA di tipo H/ACA e quello delle proteine della famiglia Cbf5p/MFL/discherina risulta ovvia se si considera la loro coesistenza

all'interno di uno stesso complesso ribonucleoproteico e la reciproca dipendenza per lo svolgimento della funzione comune.



E' inoltre interessante notare che negli Archaea il silenziamento del gene *aCbf5* ha evidenziato come differenti snoRNA della classe H/ACA derivanti dal medesimo RNA policistronico possono accumularsi a livelli differenti (Barth et al., 2005), un effetto che potrebbe essere dovuto a meccanismi di regolazione post-trascrizionali.

Tuttavia, l'esperimento sopra descritto ha dato un risultato inatteso. Infatti, come mostrato in figura 19 l'accumulo delle molecole *snm60* non è risultato affatto ridotto nelle larve mutanti, dove appare essere addirittura più abbondante che in quelle di tipo selvatico.

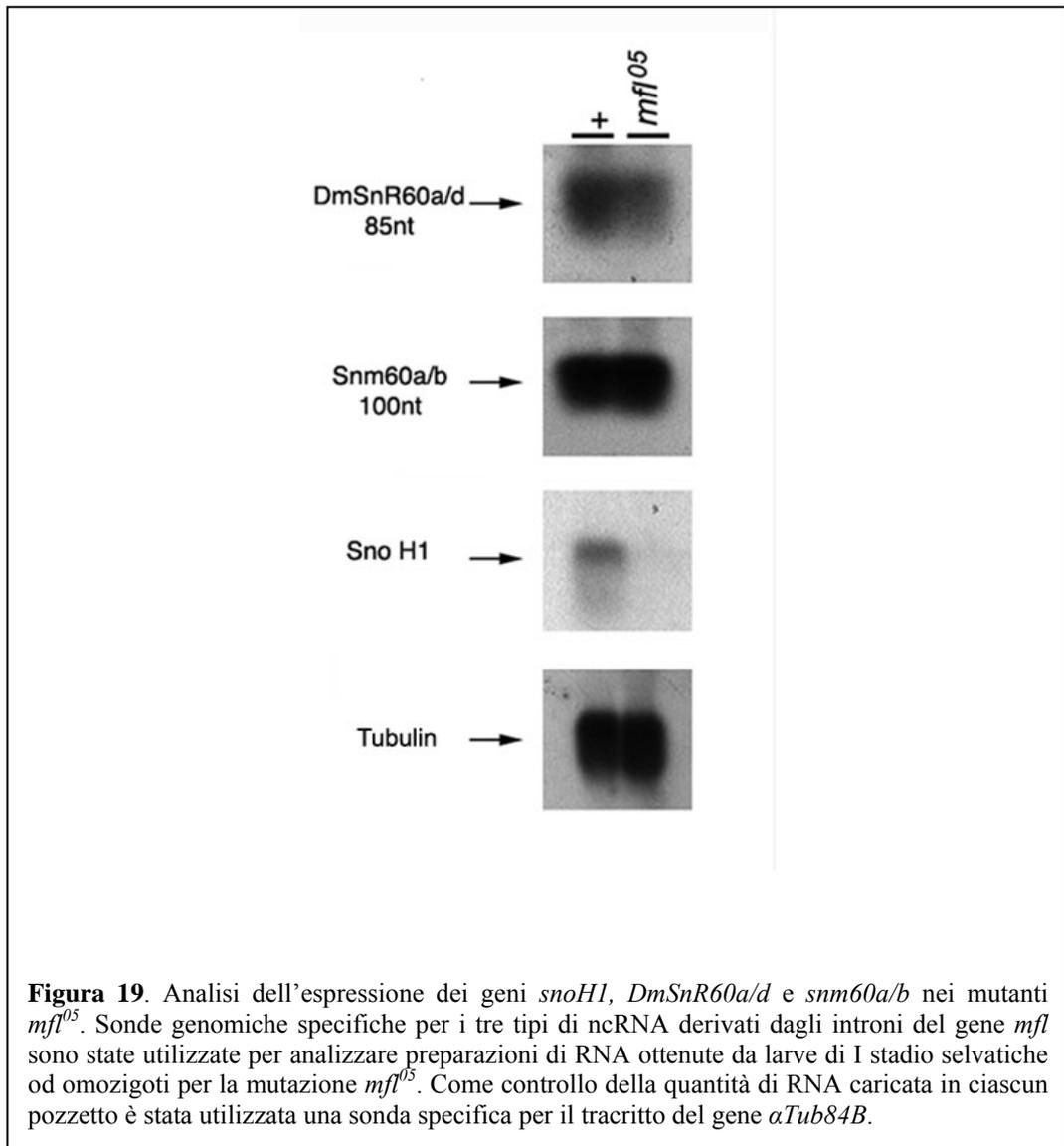


Figura 19. Analisi dell'espressione dei geni *snoH1*, *DmSnR60a/d* e *snm60a/b* nei mutanti *mfl*⁰⁵. Sonde genomiche specifiche per i tre tipi di ncRNA derivati dagli introni del gene *mfl* sono state utilizzate per analizzare preparazioni di RNA ottenute da larve di I stadio selvatiche od omozigoti per la mutazione *mfl*⁰⁵. Come controllo della quantità di RNA caricata in ciascun pozzetto è stata utilizzata una sonda specifica per il trascritto del gene *aTub84B*.

Questo risultato potrebbe essere dovuto ad una differente regolazione post-trascrizionale di queste molecole, anche se non è possibile escludere la possibilità che la trascrizione di questi geni possa, in via alternativa, partire da un promotore addizionale presente più a valle. Questa possibilità sembrerebbe essere avvalorata dalla presenza, nella banca dati Flybase (<http://flybase.bio.indiana.edu/>), di una sequenza di 2.2 kb classificata come un cDNA completo identificato nell'ambito di un' analisi di RNA estratto da ovari di femmine adulte di *Drosophila* (fig. 20).

BcDNA : GM08026

Total bases Genbank
2158 AY060937

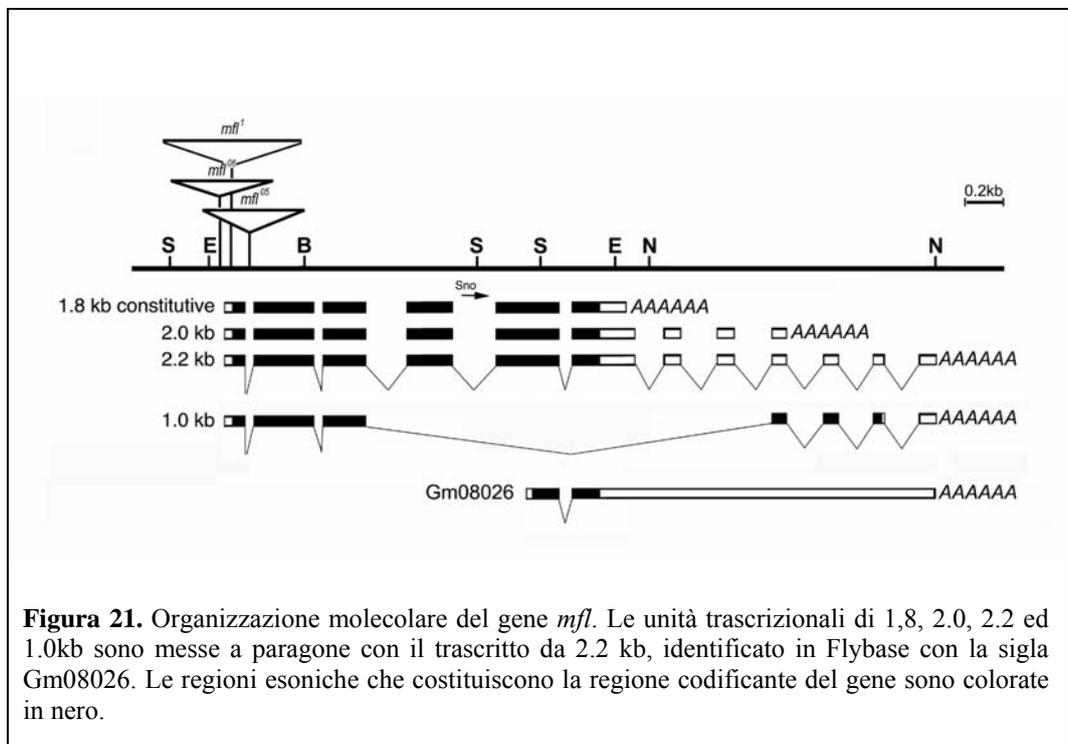
>GM08026.complete

```
CCCAATGAAAACACACCTAAGGAGTGGCTGACCGGTTATGTCTGACTACAATGCCAAGAAG
CCAGCAGCTCAAGAAGTCTCCCCACAAATGGCTCCAGTGAACCCAGCAAACGCAAGTTA
AGCACCTCCAGCGTTGAGGAACTGCTGCGGGCGGCGGTATCCGAGGAGACTCCTTCCAAG
GACAAAAAGAAGAAGAAAAAGAAGCACAAGGGCGACGAGGAAGCCCCAGAAGCCGCTGAG
GAGGAGGCAGAGCCAGTAGAAAAGGAGAAGAAAAAGAAGAAGAAGGACAAGGATAGG
GACAGAGACGAAGCTCAGGAATAGGAATCAGAAGTTTCTATAGGATTAGGTTTTAAAGAG
AATTCTACGCTTTTTTGTAGAGGCACCTACCGATTGCTAGCTGTTAGGTTTAAGGAGTGT
TTTTATGTGTAGAAGTGCATACAAATTACAAATGCATTTAAGTGACATTTCGCAATGCAAT
AAATTGAACATTTTATTGTGGAACCTATCGCTAGTTTCTTAGGTCTTAGCTTTTCCTTA
GGTTTTTCAGATAGAAAGGCAGGAATACAGACAGCAAAGGTGAGTTAACATATTACTAC
GTTCCCTTCAAGAGTGGCTTTTCGTTTTCGCTTATCTGATGAGCAATAACTTTTCGCCCTAGC
TGAAAAATGATACTTCTCATGTTGAATGCGAGATTGGTCTTACTGAAAAATAACGGAAAAAT
AGAAAAGCTCATTTGAGTTTCTAAAAACGTTGACCTGGCCTCTTATTAACCTAATTATA
TTTTCTATTGTAGCTTGGAAATCTGCTCAGAGGACAAAGCGGGAGTGACGGAAAGATACGGT
TAGTATCTATTTGATGACCATCTTTACGTTATATTAGTTCTGATGAGCATTAACTTTTCGC
CCCTATCTGAAATTGAAACTTCTCTGTTGAATGCGAGATTGGTCTTACTGAAAAACTAA
AAAACTATAAGCACATGTGGGCTATCTATATCCAATATATTTTTAAAAACCAATATATATT
TACAGATCGGCGGAACGAAAACTAAAAACGTGCTTCAATAAACATAGGAATTAAGGTAA
GAATGCCTTAATTTATAGTTGAATTTCTTATTATCATGATGAGCATTAACTTTTCGCCCT
ATCTGAAATTGAAACTTCTCATGTTGAATGCGAGATTGGTCTTACTGAAAAGAATTTTCGA
TTTTAAAAACACTAAAACAATTTTCTAAGCACTTTCTAACCAATTTTTTTTCCATTCTCTTT
ATAGATTTGTTGATAAGGTCATGCAATATATGGACTATAACTTTAATAAAAAGAGGTAAAT
ATTTAATAACTAAAAGTTATTAATAATGTTTTTAAGTGATGAGCTTTAACTTTTCGCCCTA
TCTGAAATTGAAACTTCTCCTGTTGAATGCGAGATTGGTCTTACTGAAATGACTCACAAAT
TAAAAACAAGCCTCAATCTTTTCGATTGCCTTTCTAATAAAATGCTACTTTTTCTTTACA
GATAAGTTCAAAAATTTTTTTGAAAATCATTGAATGCCACAGGAATCTTAAAAGGTAAAT
GTATTCAAAATGAATTATTTCTTTGCGAGAATAACTAACTTCTTTGAAAATTAAGTATG
TAAAAAAAACCTGAATTATGTGGCATGATTTTCACATTAACCGCAACAAATCCTACTCC
AATAGGTCTACTGACAATTTCTCACAAAAATAATGCATTTAGGCCAATACAATAACAACG
ATGCCTAACAATTGACTCTCCTTTGCCTTTTCAGTTTCGCATGACACTTTATTGCGTTGAG
TAAACAAGGTAAAAATAAATTTAAAAAACTCGTTTGGCCAGTTTCTAGTGAAAATGTTG
AAATAAAAATACTGATTTATGTGGCATGATTTTCACATTAACCGCAACAAATCCTACTCC
AATAGGTCTACTGAATATATTTCAAGAACTTAAGGTGTGTTTTAACCCAGATGTGTTT
CTAAGTGAATTTTTTCTTATATTTGCGAGATGTGCTTTTGGCGACTTTCCACTGACTGTT
CAATTGAGTTCATTTTATGCCCTTTTTTAGGTACAACACTGCCCAAGACATCATATAAT
ACTTTATAAGCAAATAAAAGAAACATACCCAAAAAATAAAAAAAAAAAAAAAAAAAAAAA
```

Figura 20. Sequenza completa del cDNA GM08026 derivato dalla regione 3' del gene *mfl*

La sequenza di questo cDNA (Gene Bank accession number AY060937; annotato in Flybase con la sigla Gm08026) indica che esso è rappresentativo di un trascritto non codificante (l'ORF più lunga è infatti di soli 39 aa) la cui estremità 5' cade all'interno del quinto esone del gene *mfl* e la cui estremità 3' coincide con quella da me precedentemente descritta per gli mRNA di 2.2 e 1.0 kb (fig. 21), confermando ulteriormente l'uso in vivo di questo nuovo sito di poliadenilazione più a valle. Come mostrato in figura 21, l'organizzazione

molecolare di Gm08026 indica che questo clone è caratterizzato dal ritenere sequenze introniche. Ciò potrebbe suggerire che in realtà esso potrebbe derivare da un intermedio parzialmente processato piuttosto che da un trascritto maturo. Anche se la coesistenza di forme maturate e non maturate è un evento riscontrato frequentemente per i trascritti che derivano da geni che producono ncRNA intronici, non è tuttavia al momento possibile escludere che questo lungo ncRNA poliadenilato rappresenti invece un veicolo alternativo per la produzione delle molecole DmSnR60 e snm60.



Analizzando la putativa sequenza promotrice di Gm08026 è stato inoltre osservato che, come quella del gene *mfl*, essa mostra delle caratteristiche tipiche della classe dei geni 5'-TOP. Infatti, sebbene la regione 5' del trascritto non presenti un lungo tratto di pirimidine, l'inizio della trascrizione è caratterizzato da 3 residui di citosina adiacenti che precedono una A, una sequenza considerata peculiare per i geni 5'-TOP di *Drosophila* (Tycowski and

Steitz, 2001). Un' altro aspetto comune ai geni 5'-TOP è rappresentato dall' assenza di una TATA box canonica 30-40 bp a monte del sito d'inizio della trascrizione. E' interessante notare che in uomo e topo lo snoRNA U80, l'ortologo funzionale degli snoRNA DmSnR60a/d, deriva dalla maturazione di un lungo ncRNA poliadenilato che appartiene alla classe dei trascritti 5' TOP (Smith and Steitz., 1998).

Questo lungo ncRNA, chiamato *gas5* (growth arrest-specific transcript 5; fig. 22), è stato inizialmente identificato in uno screening per potenziali geni onco-soppressori. Il livello del trascritto *gas5* è infatti molto basso in cellule in attiva crescita, mentre incrementa di molto in seguito a blocco della crescita (Schneider et al., 1998).

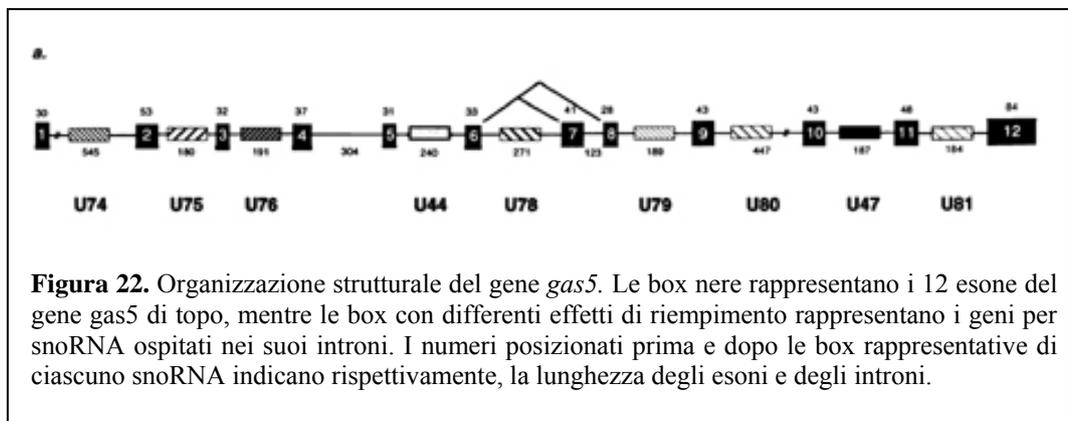


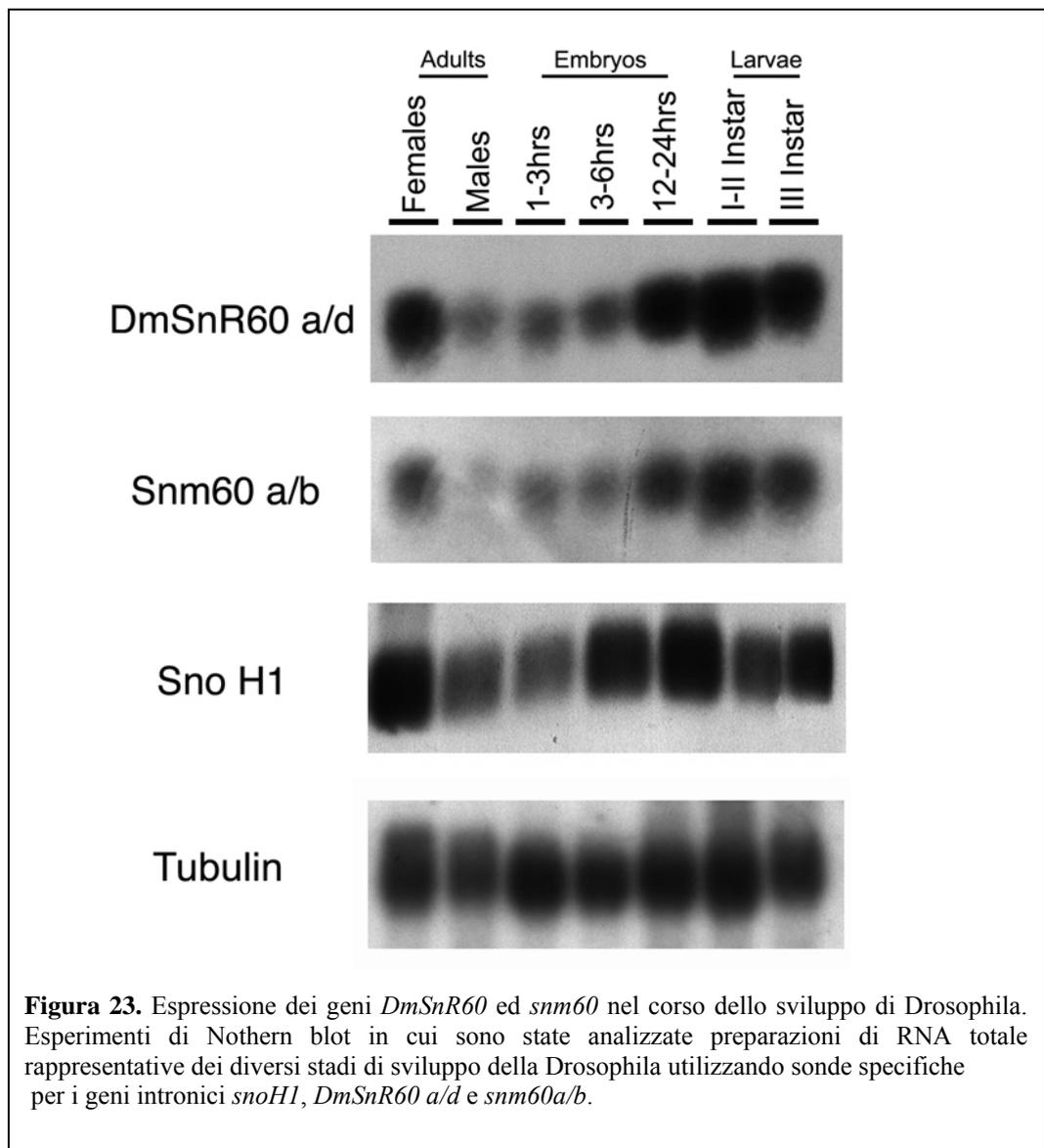
Figura 22. Organizzazione strutturale del gene *gas5*. Le box nere rappresentano i 12 esoni del gene *gas5* di topo, mentre le box con differenti effetti di riempimento rappresentano i geni per snoRNA ospitati nei suoi introni. I numeri posizionati prima e dopo le box rappresentative di ciascuno snoRNA indicano rispettivamente, la lunghezza degli esoni e degli introni.

L' osservazione che *mfl*, Gm08026 e *gas5* appartengono tutti alla classe dei geni 5'TOP è in accordo con l'ipotesi che la presenza di questa classe di sequenze all'estremità 5' possa garantire il legame di specifici fattori necessari per la corretta produzione degli snoRNA intronici a partire dal trascritto primario (de Turrís et al., 2004). E' inoltre plausibile supporre che, analogamente a quanto osservato per *gas5*, Gm08026 potrebbe costituire un

substrato per il meccanismo del non-sense-mediated RNA decay, noto degradare gli mRNA contenenti codoni “nonsense” (Rehwinkel et al., 2005). Quest’ultima ipotesi farebbe prevedere che la stabilità di questo trascritto dovrebbe essere inversamente proporzionale al tasso di sintesi proteica. Conseguentemente, in condizioni di arresto della crescita, le molecole di DmSnR60 ed snm60 dovrebbero derivare essenzialmente dalla maturazione di questo ncRNA e non dallo splicing dei trascritti primari del gene *mfl*. Se si considera che nelle larve omozigoti per la mutazione *mfl*⁰⁵ la biogenesi dei ribosomi è fortemente alterata (Giordano et al., 1999), la maggiore stabilità del trascritto Gm08026 dovuta al meccanismo del non-sense-mediated RNA decay potrebbe fornire una via alternativa alla sintesi delle molecole DmSnR60 e snm60 in condizioni di bassi livelli di sintesi proteica e di arresto della crescita. Tuttavia, è necessario invocare l’esistenza di un differente meccanismo di regolazione post-trascrizionale delle molecole snm60 per spiegarne lo specifico incremento osservato nei mutanti *mfl*⁰⁵.

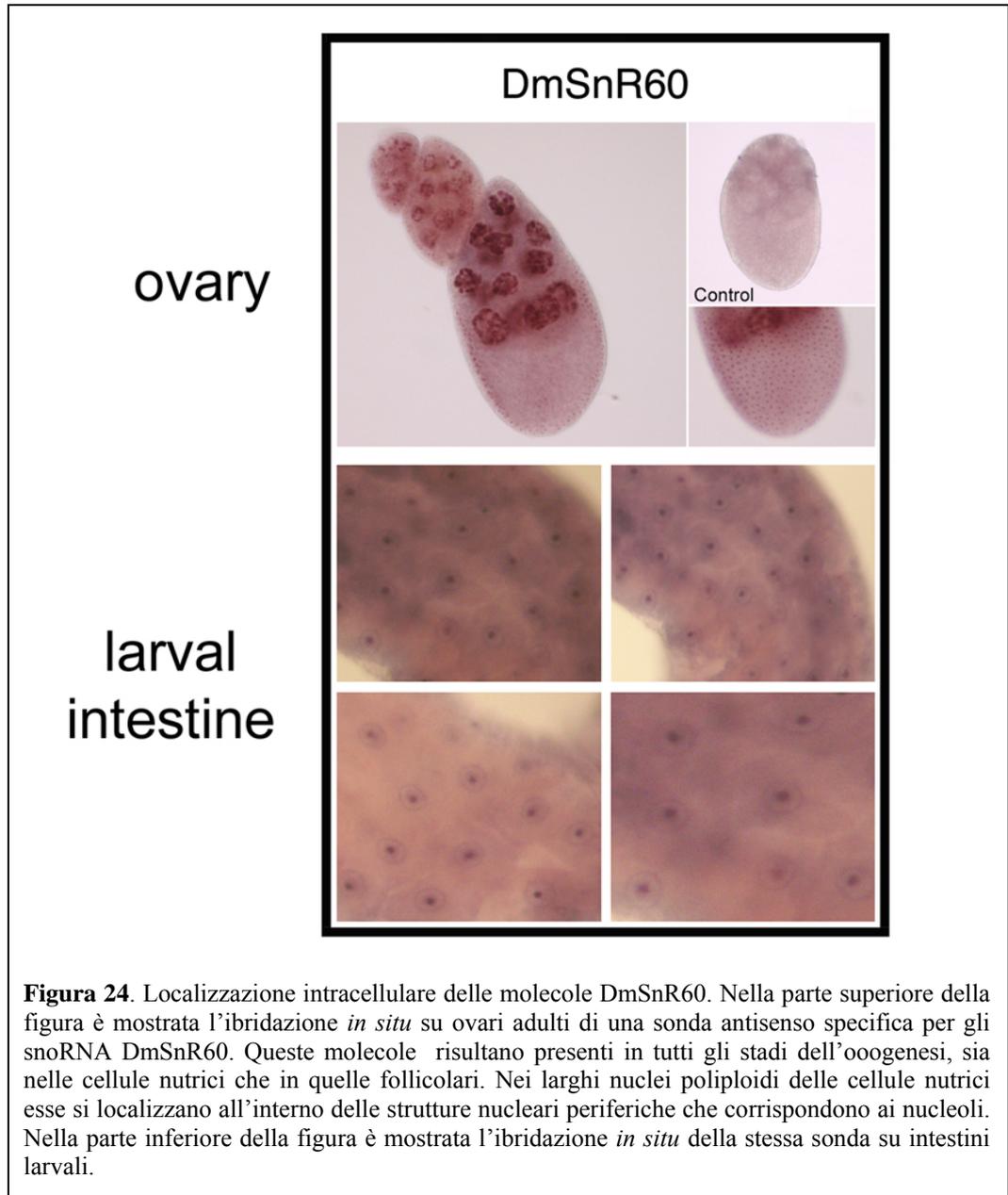
Espressione delle molecole di ncRNA derivate dagli introni del gene *mfl*

Per ottenere informazioni più complete riguardo la modalità di espressione dei piccoli RNA non codificanti che derivano dagli introni del gene *mfl* sono stati eseguiti esperimenti di Northern blot in cui preparazioni di RNA totale ottenute in vari stadi di sviluppo sono state analizzate con sonde specifiche per i geni *snoH1*, *DmSnR60* e *snm60*, utilizzando come controllo della quantità di RNA caricata in ciascun pozzetto una sonda corrispondente al gene *αTub84B*.



Come si può osservare dalla figura 23, nonostante le differenze strutturali e funzionali che caratterizzano questi piccoli ncRNA, il loro profilo di espressione nel corso dello sviluppo risulta molto simile ed è essenzialmente uniforme a quello del gene ospite (Giordano et al., 1999).

Queste piccole molecole di ncRNA si accumulano costitutivamente lungo tutte le fasi del ciclo vitale, ma la loro espressione risulta modulata dal punto di vista quantitativo, raggiungendo livelli sensibilmente più alti in larve di primo stadio ed in femmine adulte. E' inoltre interessante notare che esse mostrano un profilo di espressione marcatamente sesso-specifico negli adulti, una caratteristica davvero insolita per le molecole degli snoRNA, la cui funzione è essenzialmente legata alla modificazione dell'RNA ribosomale. L'insolita differenza fra i livelli d'espressione dei due sessi potrebbe però essere una conseguenza diretta degli elevati livelli di espressione descritta per il gene ospite negli ovari (Giordano et al., 1999), che nelle femmine adulte possono rappresentare più della metà del peso corporeo totale. Anche in considerazione dell'insolito profilo d'espressione osservato per questi piccoli ncRNA intronici ho voluto quindi definire la loro distribuzione intracellulare mediante esperimenti di ibridazione *in situ*, già precedentemente effettuati per caratterizzare l'espressione del gene *snoH1* (Giordano et al., 1999). Poiché era stato evidenziato un picco di espressione in larve di primo stadio (L1) ed in femmine adulte, per quest' analisi ho selezionato un'organo larvale, l'intestino, e come organo adulto l'ovario. Entrambi gli organi sono stati dissezionati manualmente ed analizzati con sonde di DNA a singolo filamento opportunamente marcate con digossigenina (vedi Materiali e Metodi).



L' ibridazione di una sonda “antisenso” specifica per gli snoRNA DmSnR60 su preparazioni di ovari ha evidenziato che queste molecole sono espresse in tutti i tipi di cellule. Come mostrato in figura 24, queste molecole sono infatti presenti in tutti gli stadi dell'oogenesi e si localizzano sia nelle cellule nutrici che in quelle follicolari. Esse mostrano una prevalente localizzazione nucleolare che è ben evidente nei grossi nuclei poliploidi delle cellule nutrici

dell'ovariolo. In queste cellule i nucleoli assumono infatti una morfologia caratteristica, formata da un fascio di fibre interconnesse che circondano la periferia del nucleo (Spradling, 1993), ed è ben evidente che gli snoRNA DmSnR60 si localizzano in queste strutture (fig. 24).

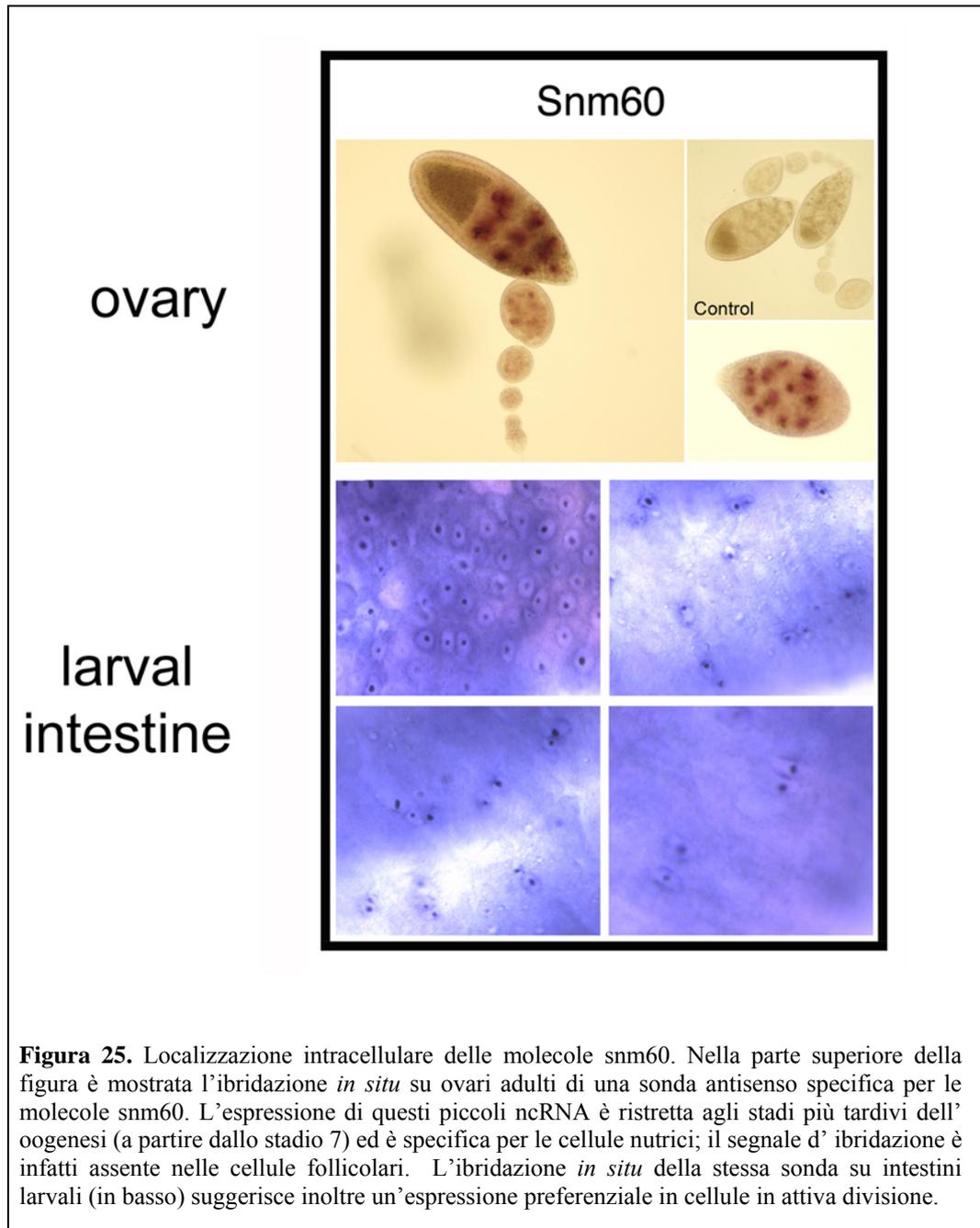


Figura 25. Localizzazione intracellulare delle molecole snm60. Nella parte superiore della figura è mostrata l'ibridazione *in situ* su ovari adulti di una sonda antisenso specifica per le molecole snm60. L'espressione di questi piccoli ncRNA è ristretta agli stadi più tardivi dell'oogenesi (a partire dallo stadio 7) ed è specifica per le cellule nutrici; il segnale d'ibridazione è infatti assente nelle cellule follicolari. L'ibridazione *in situ* della stessa sonda su intestini larvali (in basso) suggerisce inoltre un'espressione preferenziale in cellule in attiva divisione.

Un profilo di espressione decisamente diverso è emerso dall'ibridazione *in situ* con una sonda "antisenso" specifica per i trascritti dei geni *snm60*, che ha invece evidenziato come l'espressione di questi geni sia ristretta alle fasi tardive dell' oogenesi (successive allo stadio 7). Inoltre all'interno dell'ovariolo questi piccoli ncRNA si accumulano specificamente nelle cellule nutrici, ma non nelle cellule follicolari e, pur essendo confinate nel nucleo, non mostrano una localizzazione prevalentemente nucleolare (fig. 25). Queste molecole mostrano un'espressione interessante anche negli intestini larvali, dove l'ibridazione risulta spesso presente solo in cellule in attiva divisione (fig. 25). La maggioranza delle cellule in cui è possibile individuare il segnale d'ibridazione appare infatti essere nelle fasi terminali della mitosi o è formata da cellule figlie, ancora vicine l'un l'altra, che derivano dallo stesso evento mitotico. Anche se sono necessari ulteriori esperimenti, queste osservazioni potrebbero suggerire che l'espressione dei geni *snm60* possa variare lungo il ciclo cellulare. In ogni caso, gli esperimenti di ibridazione *in situ* indicano chiaramente che, mentre il profilo di espressione dei geni *DmSnR60* appare essere di tipo costitutivo, quello dei geni *snm60* è invece regolato durante l'oogenesi e ristretto alle cellule nutrici. Il differente profilo d'espressione mostrato da questi due tipi di molecole fa quindi supporre che la loro espressione sia regolata in maniera differenziale, a livello trascrizionale o post-trascrizionale. In ogni caso, la fine regolazione dell'espressione dei geni *snm60* renderà ancora più interessante la loro futura caratterizzazione funzionale.

Conservazione evolutiva del cluster di geni intronici localizzato nella regione 3' del gene *mfl*.

Per ottenere ulteriori informazioni sui geni *snm60* e *DmSnR60* è stata analizzata la loro conservazione filogenetica in altre tre diverse specie del genere *Drosophila*, *D. yakuba*, *D. virilis* e *D. ananassae*. Preparazioni di RNA totale sono state quindi analizzate in esperimenti di Northern blot con sonde specifiche per questi geni. Come mostrato in figura 26, l'espressione dei geni *DmSnR60* è stata osservata in tutte e tre le specie, mentre quella dei geni *snm60* è stata osservata solo in *D. yakuba*, una specie strettamente correlata alla *D. melanogaster*.

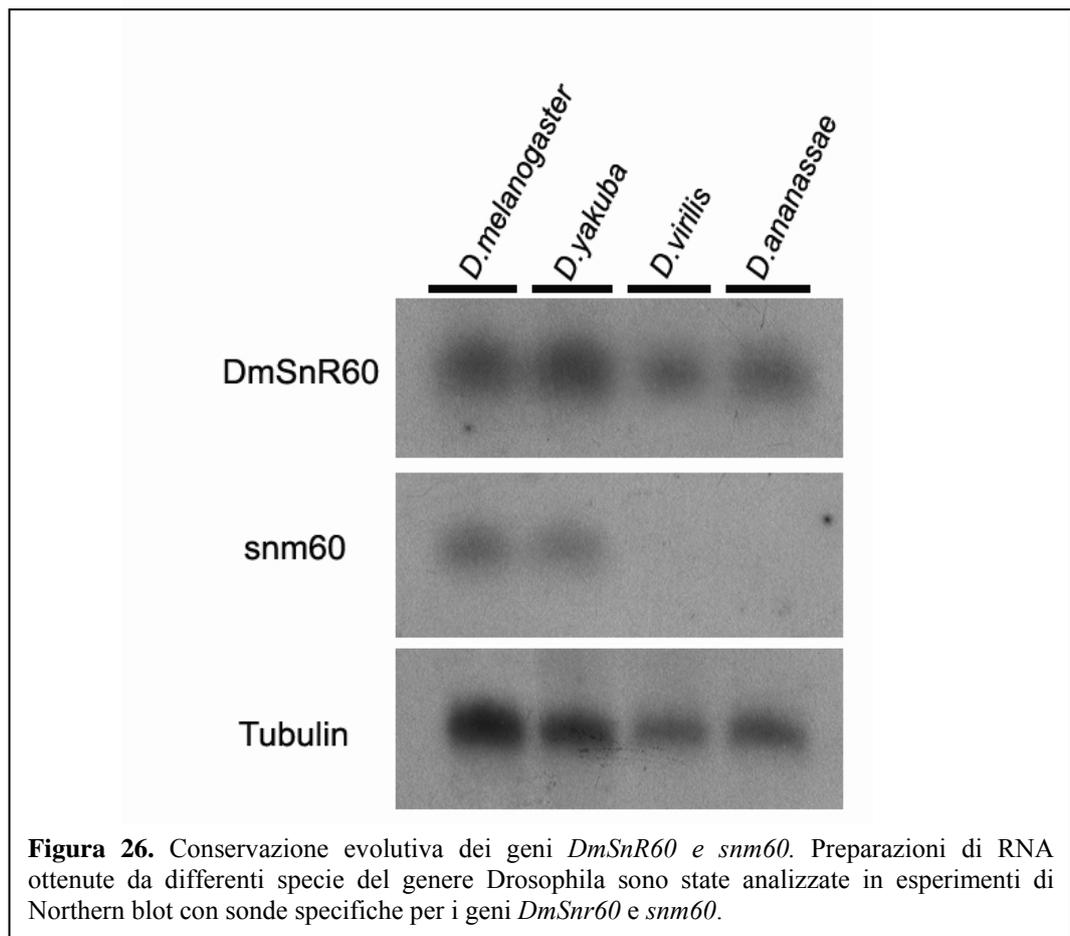


Figura 26. Conservazione evolutiva dei geni *DmSnR60* e *snm60*. Preparazioni di RNA ottenute da differenti specie del genere *Drosophila* sono state analizzate in esperimenti di Northern blot con sonde specifiche per i geni *DmSnr60* e *snm60*.

I dati ottenuti sono in perfetto accordo con i risultati di un'analisi bioinformatica che ha permesso di estendere le informazioni sulla conservazione di queste sequenze geniche anche ad altre specie del genere *Drosophila*. Un'analisi complessiva della conservazione della corrispondente regione genomica è mostrata in figura 27, dove è riportato il risultato di allineamenti multipli ottenuti sottomettendo la sequenza d'interesse all' UCSC genome Browser (<http://genome.ucsc.edu>).

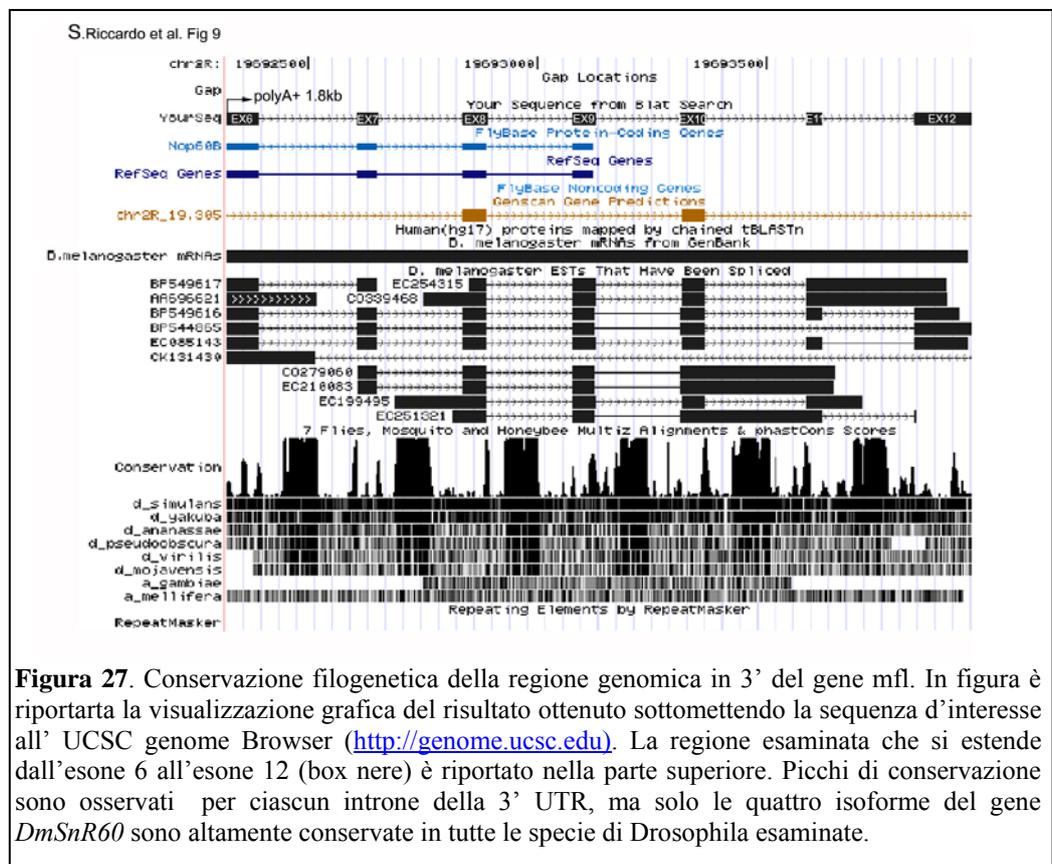


Figura 27. Conservazione filogenetica della regione genomica in 3' del gene *mfl*. In figura è riportata la visualizzazione grafica del risultato ottenuto sottomettendo la sequenza d'interesse all' UCSC genome Browser (<http://genome.ucsc.edu>). La regione esaminata che si estende dall'esone 6 all'esone 12 (box nere) è riportata nella parte superiore. Picchi di conservazione sono osservati per ciascun introne della 3' UTR, ma solo le quattro isoforme del gene *DmSnR60* sono altamente conservate in tutte le specie di *Drosophila* esaminate.

Dal diagramma è possibile individuare picchi di conservazione in tutti gli introni delle sequenze 3'UTR del gene *mfl*, ma il maggior grado di conservazione è mostrato dalle isoforme del gene *DmSnR60*. Il minor grado di conservazione filogenetica dei geni *snm60* lascerebbe supporre una più recente origine evolutiva.

Putative funzioni dei geni *snm60*

Nel tentativo di attribuire una potenziale funzione ai geni *snm60 a/b* ho voluto verificare la possibilità che i rispettivi trascritti presentassero tratti di complementarità con molecole di mRNA presenti nella banca dati Flybase. Quest'analisi è stata svolta utilizzando il programma BLASTN ed in parte con il programma MACAW (Multiple Alignment Construction and Analysis Workbench; Schuler et al., 1991) che permette di fare allineamenti con sequenze degenerate. In questi allineamenti è stata inclusa la possibilità dell'appaiamento G:U, molto comune nella formazione di duplex RNA:RNA ma generalmente non considerato nei convenzionali programmi di allineamento. Il risultato ottenuto con quest'approccio bioinformatico è stato decisamente inatteso, poiché è emerso che il trascritto del gene *snm60b* presenta al 5' terminale un tratto di 57 nt potenzialmente capace di appaiarsi in maniera imperfetta al pre-mRNA dello stesso gene ospite, il gene *mfl* (fig. 28).

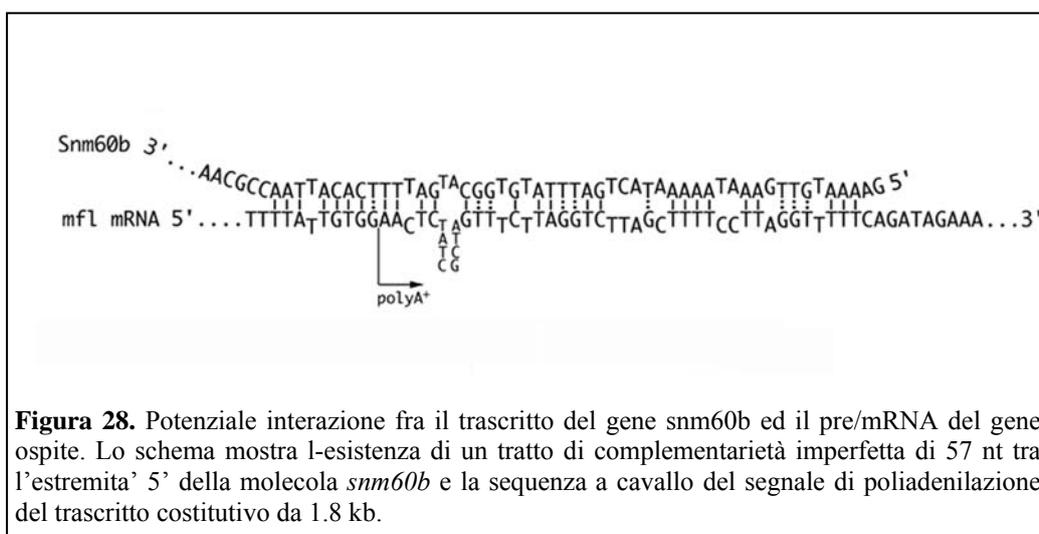


Figura 28. Potenziale interazione fra il trascritto del gene *snm60b* ed il pre/mRNA del gene ospite. Lo schema mostra l'esistenza di un tratto di complementarità imperfetta di 57 nt tra l'estremità 5' della molecola *snm60b* e la sequenza a cavallo del segnale di poliadenilazione del trascritto costitutivo da 1.8 kb.

E' interessante notare che questo tratto di complementarità comprende il sito di poliadenilazione del trascritto costitutivo da 1.8 kb, lasciando supporre che il verificarsi di questo appaiamento potrebbe svolgere un ruolo regolativo sull'uso dei tre siti di terminazione alternativi identificati per il gene *mfl*.

E' da notare che un segmento di 14 nt di questo potenziale elemento antisenso che si appaia in corrispondenza del sito di poliadenilazione è interno al segmento di 59 bp perfettamente identico nei geni *snm60a* e *snm60b*, così che i trascritti di entrambi i geni sono potenzialmente complementari a questo sito di terminazione, sebbene con elementi antisenso di differente lunghezza. Poiché è noto che molti dei piccoli RNA non codificanti con funzione regolatoria selezionano le molecole bersaglio mediante appaiamento di basi, che può essere perfetto o imperfetto, l'esistenza di una complementarità tra *snm60b* e l'mRNA da 1.8 kb potrebbe avere una valenza funzionale, suggerendo l'ipotesi che le molecole *snm60a/b* potrebbero svolgere un ruolo regolativo sull'espressione del proprio gene ospite, influenzandone l'uso alternativo dei diversi siti di poliadenilazione.

Analisi funzionale dei geni *DmSnR60*

I geni *DmSnR60* appartengono alla classe dei geni che codificano per piccoli RNA della classe C/D, ed è ben noto che gli snoRNA di questa classe sono responsabili della metilazione dei gruppi 2'-OH dei ribosi (Kiss-Laszlo et al., 1998). Gli snoRNA di tipo C/D contengono due corti motivi di sequenza conservati: la box C (5'RUGAUGA3') e la box D (5'CUGA3'), localizzati a pochi nucleotidi dalle estremità 5' e 3', rispettivamente (fig. 29). I due motivi si trovano generalmente ravvicinati in una struttura secondaria terminale che coinvolge 4-5 nt ad entrambe le estremità e che è critica per la stessa biogenesi e la localizzazione nucleolare di queste molecole. Il motivo C/D inoltre è essenziale per la maturazione, la stabilità, la localizzazione nucleolare e la funzione degli snoRNA di questa classe (Cavaillé et al., 1996; Caffarelli et al., 1996; Samarsky et al., 1998; Villa et al., 2000). Gli snoRNA di tipo C/D contengono talvolta, nella loro porzione centrale, una copia meno conservata della box C, chiamata per questo motivo box C' (Kiss-Laszlo et al., 1998) e una copia aggiuntiva della box D, chiamata similmente box D' (Tycowski et al., 1996). Immediatamente a monte della box D e/o D' si trovano elementi guida antisenso, lunghi 10-21 nt, complementari in maniera perfetta alla molecola di RNA bersaglio (fig. 29). Tramite questo appaiamento con la molecola target si viene a formare una doppia elica nella quale è presente il nucleotide da modificare, che è sempre posizionato esattamente 5 nt a monte delle box D o D' (fig. 29) (Kiss-Laszlo *et al.*, 1996). Ciascuna copia del gene *DmSnR60* (denominate *a*, *b*, *c*, *d*) possiede i canonici box C (5'-UGAUGA-3'),

snoRNA di classe C/D

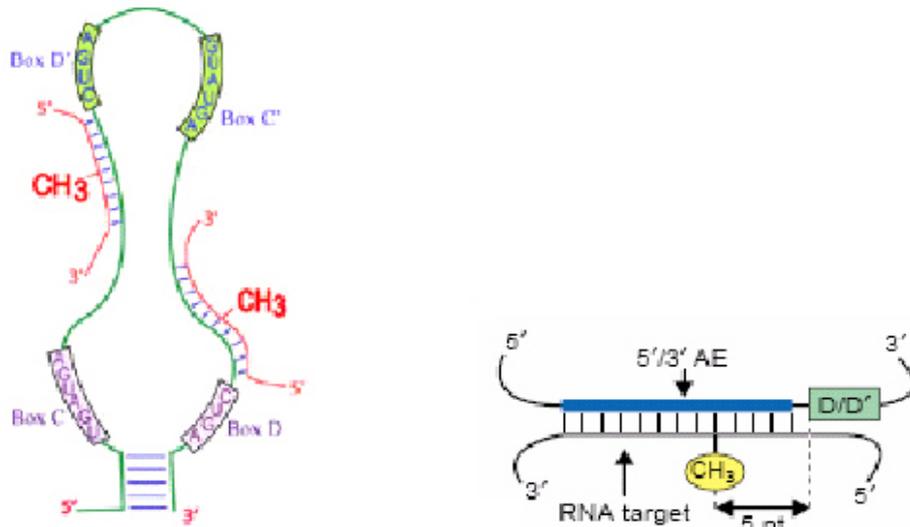
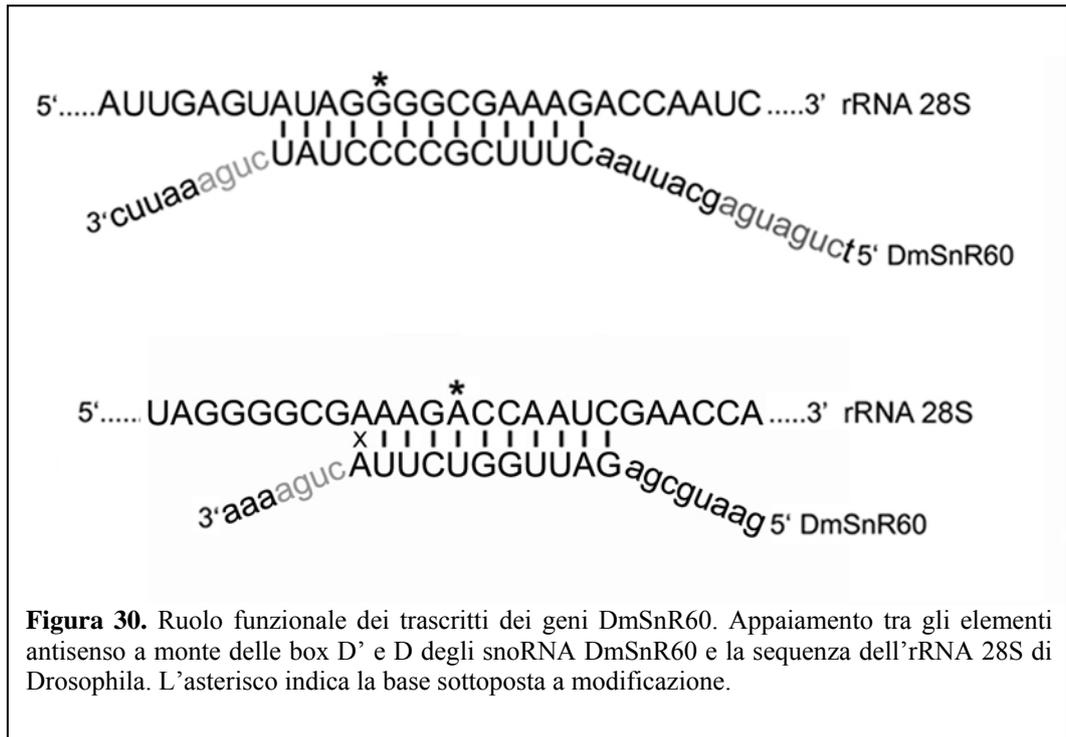


Figura 29. Struttura degli snoRNA di classe C/D. Rappresentazione schematica di una canonica molecola di snoRNA di classe C/D. Il gruppo metilico rappresenta il nucleotide metilato sullo RNA target all'interno della doppia elica formata tra la sequenza guida dello snoRNA e l'RNA target. Ogni sequenza guida antisenso (AE) media la 2'-O-metilazione del nucleotide posizionato 5 nt a monte delle box D/D'.

D e D' (5'-CUGA-3'), suggerendo che questi geni possano codificare per snoRNA bifunzionali (fig. 16). Le quattro isoforme *DmSnR60* sono disposte in tandem, con la stessa polarità, e le loro sequenze sono altamente conservate, con polimorfismi occasionali che si verificano essenzialmente all'esterno delle regioni funzionali (fig. 16). È da notare che i geni *DmSnR60* sono tutti localizzati circa 70 nucleotidi a monte del sito 3' di splicing dei loro introni ospiti. Nei vertebrati, questo tipo di localizzazione è stata osservata essere ideale per l'espressione degli snoRNA C/D di derivazione intronica. Si è ipotizzato che questa distanza possa essere ottimale per determinare una

sinergia con lo splicing, favorendo l'assemblaggio delle snoRNP di tipo C/D (Hirose and Steitz, 2001). La localizzazione delle isoforme del gene *DmSnR60* indica che la regola osservata da Hirose e collaboratori potrebbe essere valida anche negli invertebrati. Poiché gli snoRNA della classe C/D, come già detto, selezionano per la metilazione la base posizionata 5 nucleotidi a monte del box D'/D (Kiss, 2002), queste sequenze complementari sono state predette dirigere la 2'-O metilazione rispettivamente al residuo G1082 e A1092 (Accardo et al., 2004; Huang et al., 2005) dell'rRNA 28S di *Drosophila* (fig. 30). È interessante notare che la metilazione di questi due siti è conservata dal lievito ai vertebrati. Infatti, come riportato nel "database" che racchiude ed aggiorna tutte le informazioni disponibili sugli snoRNA (<http://www-snorna.biotoul.fr>), la metilazione al residuo G1083 è equivalente a quella dei residui modificati dallo snR60 in lievito e dallo snoRNA U80 in mammifero; analogamente, la 2'-O-metilazione al residuo Am1092 è stata osservata in lievito, dove è guidata dallo snR84, ed in uomo, dove è guidata da uno snoRNA non ancora identificato. In figura 30 è schematizzato l'appaiamento degli elementi antisense a monte del box D' e D dei geni *DmSnR60*; come si può osservare, per entrambi è presente un lungo tratto di perfetta complementarità (rispettivamente di 13 e 10 nt) con l'rRNA 28S di *Drosophila* (Accardo et al., 2004; Huang et al., 2005). Ad oggi, tranne poche eccezioni, i siti di metilazione dell'rRNA ribosomale di *Drosophila* non sono stati ancora determinati sperimentalmente. Di conseguenza, la definizione del ruolo funzionale degli snoRNA *DmSnR60* ha richiesto l'allestimento di un saggio che permettesse di verificare l'effettiva modificazione dei residui predetti.



Per questo scopo è stata utilizzata la tecnica di Maden (2001), secondo cui la specifica sequenza di rRNA viene retro-trascritta in condizioni di alta e bassa concentrazioni di nucleotidi. In questo saggio, uno specifico segnale di stop può essere individuato al sito di metilazione in condizioni di bassa concentrazione di nucleotidi, ma non di alta concentrazione di nucleotidi. Come mostrato in figura 31, quest'analisi ha confermato la presenza di un nucleotide metilato in entrambe le posizioni predette. Nessun altro snoRNA della classe C/D potenzialmente capace di dirigere la metilazione del nucleotide G1082 o del nucleotide A1092 è stato finora descritto in *Drosophila*. Di conseguenza, l'evidenza sperimentale della reale modificazione di questi residui conferma il ruolo di guida ipotizzato per le molecole DmSnR60. E' da notare che l'esperimento descritto in figura 31 ha portato ad individuare anche la modificazione del residuo Gm1108.

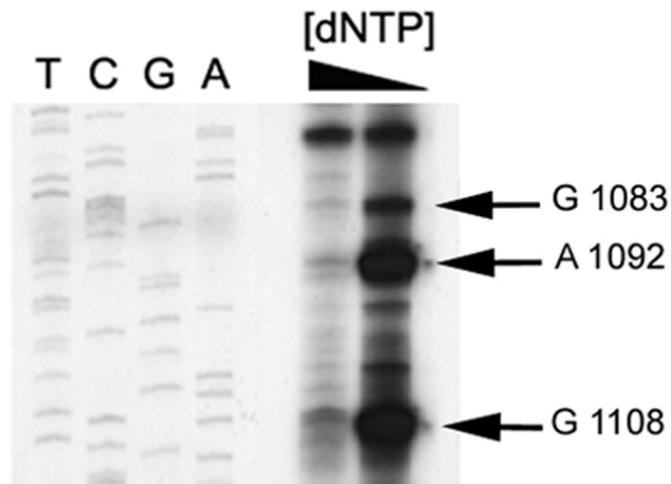


Figura 31. Saggio di metilazione. La reazione di retro trascrizione, innescata da un oligonucleotide complementare all'rRNA 28S di *Drosophila* in condizione di alta e bassa concentrazioni di nucleotidi, ha confermato le metilazione dei nucleotidi G1083 ed A1092. L' esperimento ha anche messo in evidenza un'addizionale sito di mutilazione, al nucleotide G1108.

Poiché la modificazione di questo residuo non è stata descritta in nessun altro organismo si può supporre che possa essere specifica dell'rRNA 28S di *Drosophila*. Un unico esperimento ha quindi permesso di dimostrare l'effettiva mutilazione dei nucleotidi alle posizioni predette (Accardo et al., 2004; Huang et al., 2005), confermando la natura bifunzionale ipotizzata per gli snoRNA della classe C/D ospitati negli introni del gene *mfl*. E' interessante notare che l'allineamento dell'elemento antisenso a monte della box D, identico in tutte e quattro le isoforme del gene DmSnR60, alla sequenza del pre-rRNA aveva anche identificato un terza putativa regione di appaiamento, lunga 13 nt, sull'rRNA 28S. Tale appaiamento, che dovrebbe portare alla 2'O-metilazione della "A" in posizione 850 (fig. 32), appariva in realtà poco probabile dato che prevedeva un numero di appaiamenti G:U superiore a tre (Yang et al., 2006).

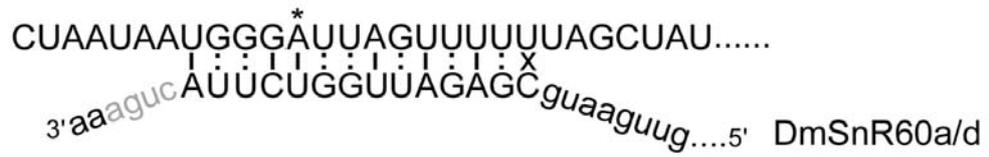


Figura 32. Ipotetico appaiamento tra l'elemento antisenso a monte della box D dei geni DmSnR60 e l'rRNA 28S di Drosophila.

In effetti, nessun segnale di stop è stato individuato in corrispondenza del nucleotide predetto in seguito a retro-trascrizione della regione specifica nelle condizioni precedentemente descritte (fig. 33). Tuttavia, un segnale di stop è stato osservato 2 nucleotidi prima del residuo A850; la metilazione di questo residuo non è stata descritta in altri organismi, così che anche questa modificazione potrebbe essere specifica dell'rRNA 28S di Drosophila.

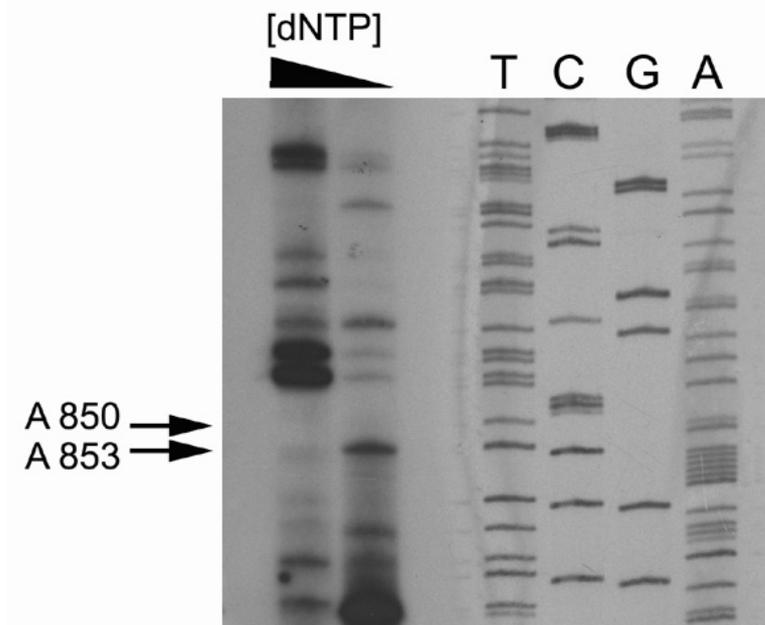
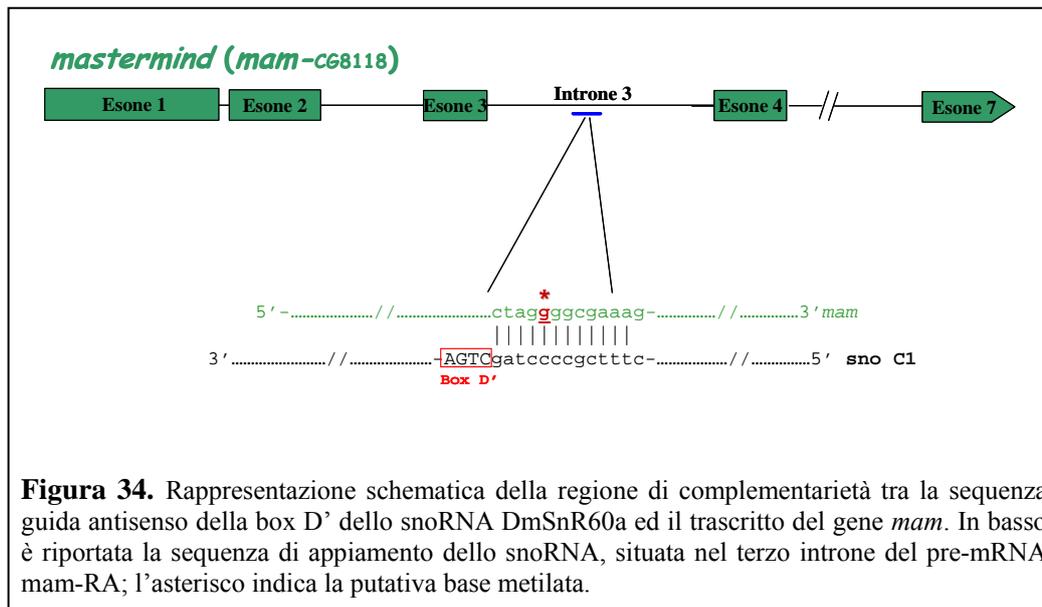


Figura 33. Saggio di metilazione. La sequenza dell'rRNA 28S riconosciuta dall'elemento antisenso a monte della box D è stata analizzata per la presenza di metilazione del residuo A850.

Identificazione computazionale di ulteriori putativi target di modificazione degli snoRNA DmSnR60

E' interessante notare che un singolo polimorfismo nucleotidico nell'elemento antisenso a monte della box D' dell'isoforma *DmSnR60a* interrompe la perfetta complementarità con l'rRNA 28S. Poiché la variabilità della sequenza antisenso dell'RNA guida può variare il repertorio delle molecole target, non era possibile escludere che questa isoforma potesse dirigere la modificazione di una molecola di RNA bersaglio non ancora identificata. Di conseguenza, l'elemento antisenso a monte della box D' dell'isoforma *DmSnR60a* rientra a pieno titolo nella categoria dei cosiddetti "orfani", per i quali restano cioè ancora da definire le molecole bersaglio. A questo proposito è importante ricordare che recentemente è stato riportato in letteratura un esempio di snoRNA "orfano" di tipo C/D, lo snoRNA di mammifero MBII-52, che è in grado di modificare come molecola bersaglio una molecola di RNA messaggero, cioè l'mRNA del recettore della serotonina 5-HT_{2C} (Vitali et al., 2005). E' quindi plausibile ipotizzare che specifiche molecole di RNA messaggero possano rappresentare i bersagli degli snoRNA "orfani". Partendo da questi presupposti è stata condotta, tramite BLAST, una ricerca sull'intero trascrittoma di *Drosophila* allineando la sequenza antisenso a monte della box D' dell'isoforma *DmSnR60a* con tutti i trascritti (EST, cDNA) la cui sequenza è stata depositata in Flybase (<http://flybase.bio.indiana.edu/>). Questa ricerca ha rivelato che l'elemento antisenso a monte della box D' dell'isoforma "a" può potenzialmente guidare la modificazione del trascritto primario del gene *mam* (*mastermind*, *CG8118*), con la putativa base bersaglio localizzata nel terzo introne del pre-mRNA (fig. 34).



Poiché gli snoRNA, sia della classe C/D che di quella H/ACA, possono guidare anche la modificazione degli small nuclear RNA coinvolti nel macchinario dello splicing, ho verificato anche questa possibilità allineando l'elemento antisenso a monte del box D' con le sequenze depositate nella banca dati "noncode" (<http://noncode.bioinfo.org.cn>), che raggruppa le sequenze delle molecole di snRNA identificate finora nei diversi organismi. Questa ricerca ha portato a stabilire che l'elemento antisenso a monte della box D' dell'isoforma *DmSnR60a* potrebbe appaiarsi, per 9 nt, a tutte le isoforme note dello snRNA U2 di *Drosophila* (fig. 35). Lo stesso appaiamento, lungo però 8 nt, è prevedibile anche per il corrispondente elemento antisenso delle altre isoforme del gene *DmSnR60*. Nel genoma di *D. melanogaster* sono state identificate 6 isoforme dello snRNA U2; questo snRNA rappresenta, insieme alla snRNP U6, il core catalitico dello spliceosoma di tipo U2 che è responsabile dello splicing della maggior parte

degli introni. E' interessante notare che le modifiche guidate dagli snoRNA sulle molecole di snRNA possono influenzarne il legame con le specifiche proteine che compongono le snRNP; infatti, lo snRNA U2 pare infatti capace di legare gli specifici fattori di splicing SF3a ed SF3b solo dopo la sua modificazione (Nesic and Kramer, 2001).

snRNA:U2:14B-RA/ snRNA:U2:34ABa-RA/ snRNA:U2:34ABb-RA
/snRNA:U34ABc-RA
snRNA:U2:38ABa-RA/ snRNA:U2:38ABb-RA

```

1 atcgcttctc ggccttatgg ctaagatcaa agtgtagtat ctgttcttat cagcttaaca
61 tctgatagtt cctccattgg aggacaacaa atgttaaact gatttttga atcagacgga
121 gtgctaggggcttgctccac ctctgtcacg ggttggcccg gatttgcagt accgccgga
      |||||
3'-AGTCgatccccgctttc-5' DmSnR60a
181 cttcggccca ac

```

Figura 35. Putativo sito di metilazione presente in tutte le isoforme note dello snRNA U2 di *Drosophila*. In giallo è evidenziata la sequenza dello snRNA che si appaia in maniera perfetta alla sequenza guida; la base in rosso rappresenta la putativa base mutilata; anche la box D' dello snoRNA è evidenziata in rosso.

Per completare l'analisi bioinformatica ho poi voluto controllare la possibilità che anche l'elemento antisenso posizionato a monte della box D, comune a tutte le isoforme del gene *DmSnR60*, potesse riconoscere come target addizionali altre molecole di RNA (mRNA o snRNA). L'allineamento mediante il programma BLAST ha portato così a stabilire che quest'elemento può in effetti potenzialmente riconoscere come putativo target, oltre l'rRNA 28S, anche il trascritto del gene *CG7112*, con la putativa base bersaglio localizzata nel quarto esone (fig. 36). Lo stesso elemento antisenso è anche in grado di appaiarsi a 6 delle 7 isoforme note dello snRNA U5 ed all'isoforma RA dello snRNA U7 (fig. 37).

Nel genoma di *D. melanogaster* sono note sette varianti dello snRNA U5; sorprendentemente, l'espressione di ciascuna di queste varianti presenta un profilo d'espressione regolato durante lo sviluppo, con la variante U5-38Aba specificamente espressa nella testa. Questi differenti livelli d'espressione in diverse condizioni di sviluppo lascerebbero supporre un possibile ruolo delle diverse isoforme negli eventi di splicing stadio-specifici. E' possibile che queste varianti possano contribuire agli eventi di splicing mediati dal complesso minore dello spliceosoma, che è noto catalizzare lo splicing di un sotto-gruppo di introni che contengono le sequenze AT/AC alle giunzioni di taglio. Altra possibilità è che le differenti isoforme di U5 siano coinvolte in eventi di splicing alternativo; infatti, la scarsa conservazione nucleotidica riscontrata all'estremità 3' di queste isoforme, potrebbe mediare, durante l'assemblaggio dello spliceosoma sul pre-mRNA, l'interazione di ciascuna variante con fattori di splicing alternativi, indirizzando il taglio verso siti di splicing non canonici (Chen et al., 2005).

Per quel che riguarda lo snRNA U7, bisogna innanzi tutto dire che questo snRNA è coinvolto nella formazione del 3' terminale dell'mRNA degli istoni. Di conseguenza la sua funzione, pur essendo correlata a quella delle snRNP che intervengono nello splicing, è da considerarsi biochimicamente distinta. Nei metazoi, infatti, la formazione del 3' terminale non poliadenilato degli istoni è l'unica reazione di maturazione che è nota prevedere l'intervento di una specifica famiglia di proteine (Muller and Schumperli, 1997) e che è strettamente legata al ciclo cellulare. In *Drosophila*, i geni per gli istoni sono caratterizzati dalla presenza di siti criptici di poliadenilazione, che in

condizioni normali non vengono riconosciuti. In moscerini mutanti per i geni dello snRNA U7, questi siti vengono invece riconosciuti, portando alla formazione di trascritti poliadenilati degli istoni che non possono essere correttamente maturati. I mutanti da perdita di funzione degli snRNA U7, pur presentando sterilità maschile, risultano tuttavia vitali, poiché le molecole di snRNA U7 che vengono trasmesse all'embrione per via materna sono sufficienti per dirigere la normale maturazione dei pre-mRNA istonici fino ad oltre il terzo stadio larvale (Godfrey et al., 2006). La tabella 1 riassume quindi tutti i dati ottenuti dalle ricerche bioinformatiche sopra descritte, riportando anche le caratteristiche dei geni i cui trascritti rappresentano putativi target per gli snoRNA *DmSnR60*, la localizzazione precisa della putativa sequenza bersaglio, nonché la regione di appaiamento tra la sequenza guida dello snoRNA e quella del putativo target.

Tabella 1:

SnoRNA	Putativo/i snRNA target e base put. meth.	Appaiamento Seq.Antisenso/Seq.target	Putativi target aggiuntivi	Appaiamento Seq.Antisenso/Seq.target	
DmSnR60	<p>B O X D'</p> <p>snRNA:U2:14B-RA, snRNA:U2:34ABa-RA snRNA:U2:34ABb-RA snRNA:U2:34ABc-RA snRNA:U2:38ABa-RA snRNA:U2:38ABb-RA G</p>	<p>D' box 3'-ACUCgaucccccgcuuuc-5' AS sno 5'-gctaggggc - 3'snRNA U2</p>	<p>A G S E N E M B L Y A N G E N T A T E S D A N N G E N T A T E S</p>	<p><i>mastermind</i>(CG8118): Notch signaling pathway, ectoderm development, asymmetric cell division, nervous system development INTRONE 3 G</p> <p><i>CR40459</i>: Cytosolic ribosomal rna gene 28SrRNA:CR40459 G</p> <p><i>CG7112</i>: Chromosome segregation, mitosis, transport, Rab GTPase activator activity ESONE 4 G</p>	<p>D' box 3'-ACUCgaucccccgcuuuc-5' AS sno 5'-ctaggggcgaaag-3'CG8118</p> <p>D' box 3'-ACUCgaucccccgcuuuc-5' AS sno 5'- uaggggcgaaag-3'CR40459</p> <p>D box 3'-ACUCauuccuggguaga-5' AS sno 5'- aagaccaatct - 3'CG7112</p>
	<p>B O X D</p> <p>snRNA:U5:14B-RA, snRNA:U5:23D-RA snRNA:U5:34A-RA snRNA:U5:35D-RA snRNA:U5:38ABa-RA snRNA:U5:38ABb-RA snRNA:U5:63BC-RA A</p>	<p>D box 3'-ACUCauuccuggguaga-5' AS sno 5'-uggggcca - 3'snRNA U5</p>		<p>D box 3'-ACUCauuccuggguaga-5' AS sno 5'- aagaccaatct - 3'CG7112</p>	

Sovra-espressione *in vivo* dei geni *DmSnR60* e *snm60*

Come primo approccio per analizzare in modo più diretto il ruolo biologico svolto dai geni *DmSnR60* e *snm60* ho voluto verificare se la sovra-espressione di questi geni *in vivo* potesse produrre effetti fenotipici evidenti.

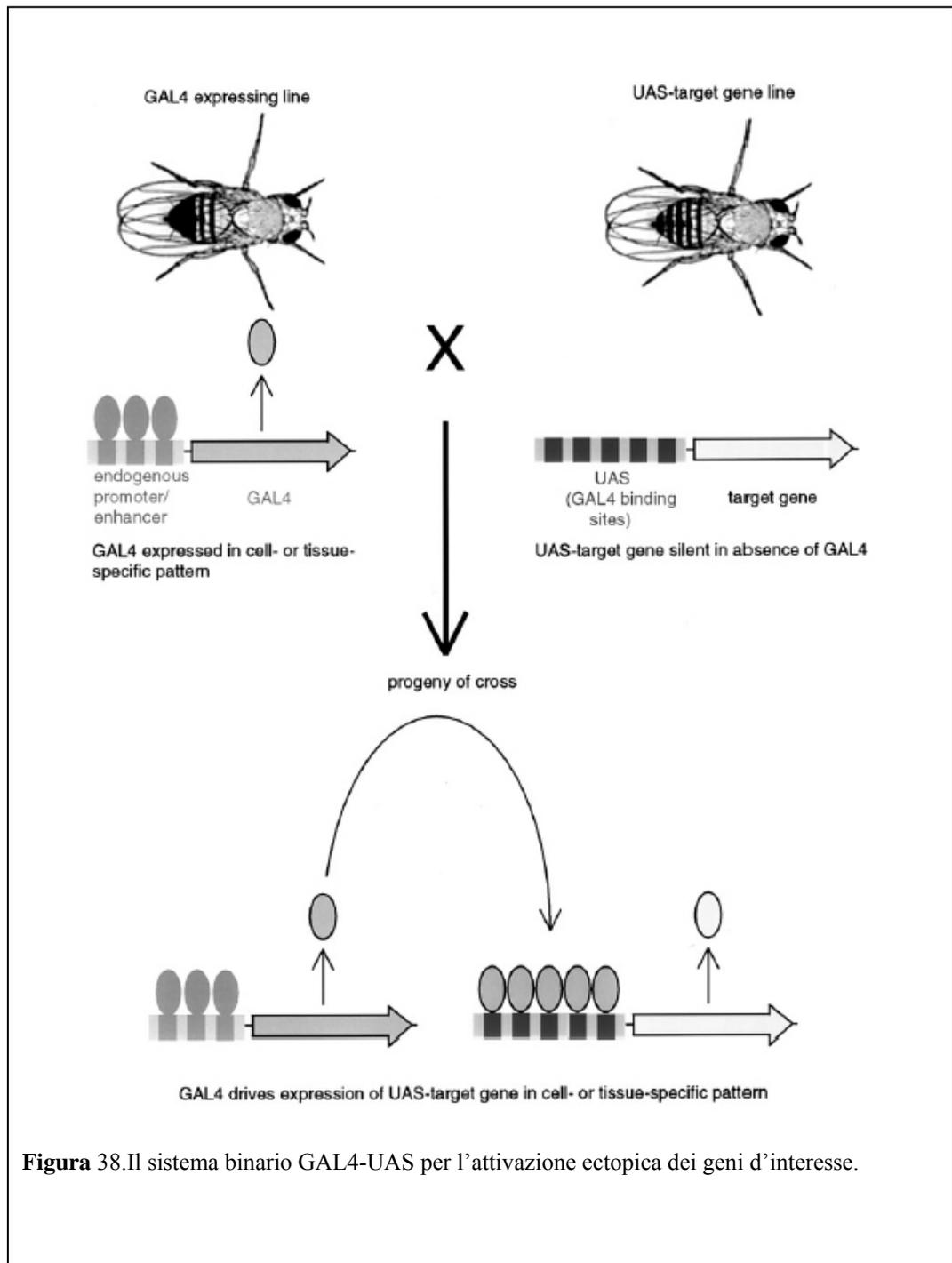


Figura 38. Il sistema binario GAL4-UAS per l'attivazione ectopica dei geni d'interesse.

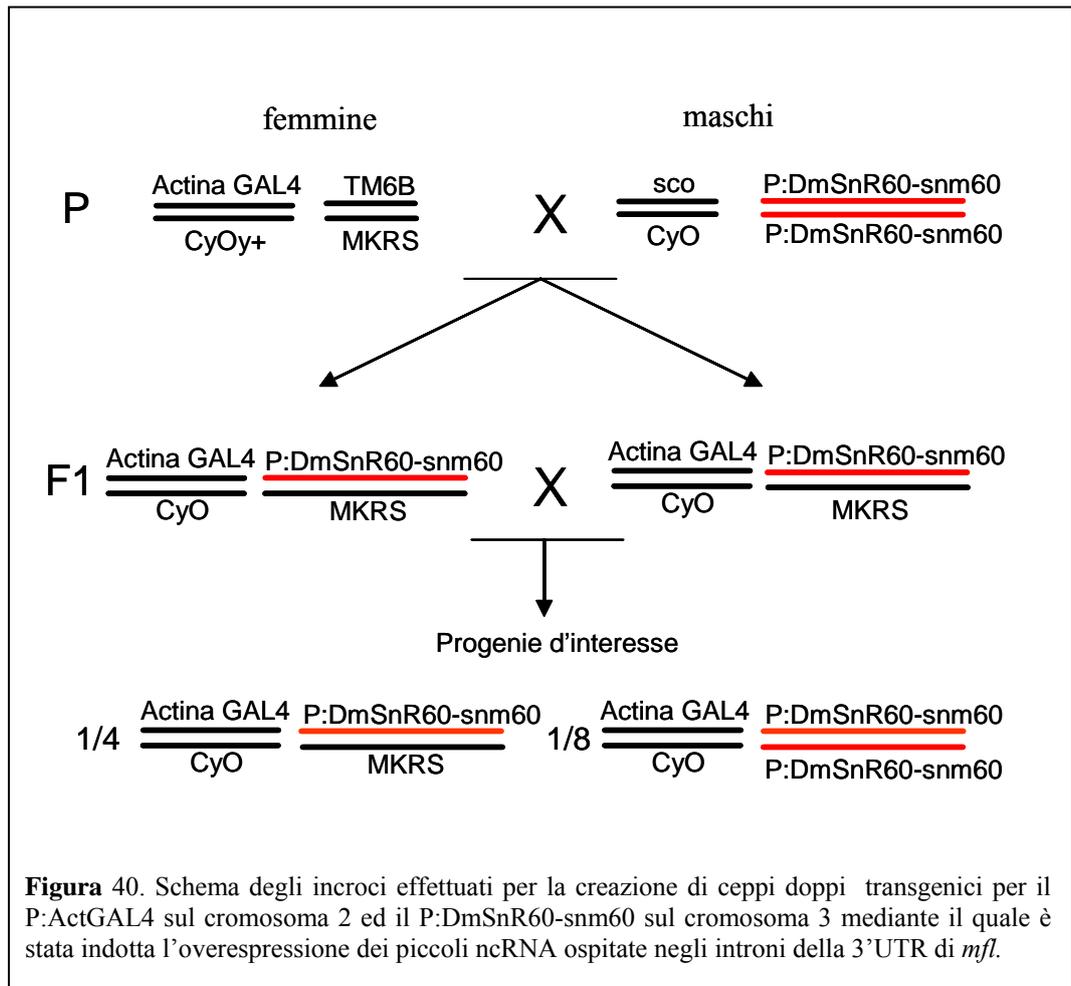
A questo scopo ho condotto esperimenti di transgenesi utilizzando il sistema binario GAL4-UAS, che risulta particolarmente adatto per dirigere in maniera ectopica l'espressione dei geni d'interesse.

Questo sistema utilizza il gene che codifica per la proteina GAL4 per attivare il gene bersaglio, che viene a sua volta clonato a valle della sequenza UAS, cioè la sequenza riconosciuta dalla proteina attivatrice GAL4. I due transgeni, quello "attivatore" e quello "bersaglio" sono portati da linee distinte e solo l'avvio di opportuni incroci innesca l'attivazione ectopica del gene d'interesse nella progenie. La possibilità di utilizzare differenti linee attivatrici GAL4 permette poi di attivare l'espressione del gene con modalità diverse, che possono essere sia di tipo costitutivo che regolato durante lo sviluppo o nei diversi tipi di cellule o tessuti (fig. 38).

La progenie dell'incrocio in cui è stata attivata la sovra-espressione del gene d'interesse è individuata mediante opportuni marcatori fenotipici. Nei miei esperimenti la trasformazione della linea germinale è stata indotta in embrioni mutanti per il gene *w*, cioè con fenotipo "occhio bianco", il transgene portava quindi l'allele *w*⁺ e le linee transgeniche sono state quindi selezionate in base al colore selvatico dell'occhio. In questi esperimenti di trasformazione stabile della linea germinale di *Drosophila* è stato utilizzato il vettore pUAST, e la struttura del transgene è mostrata in figura 39.

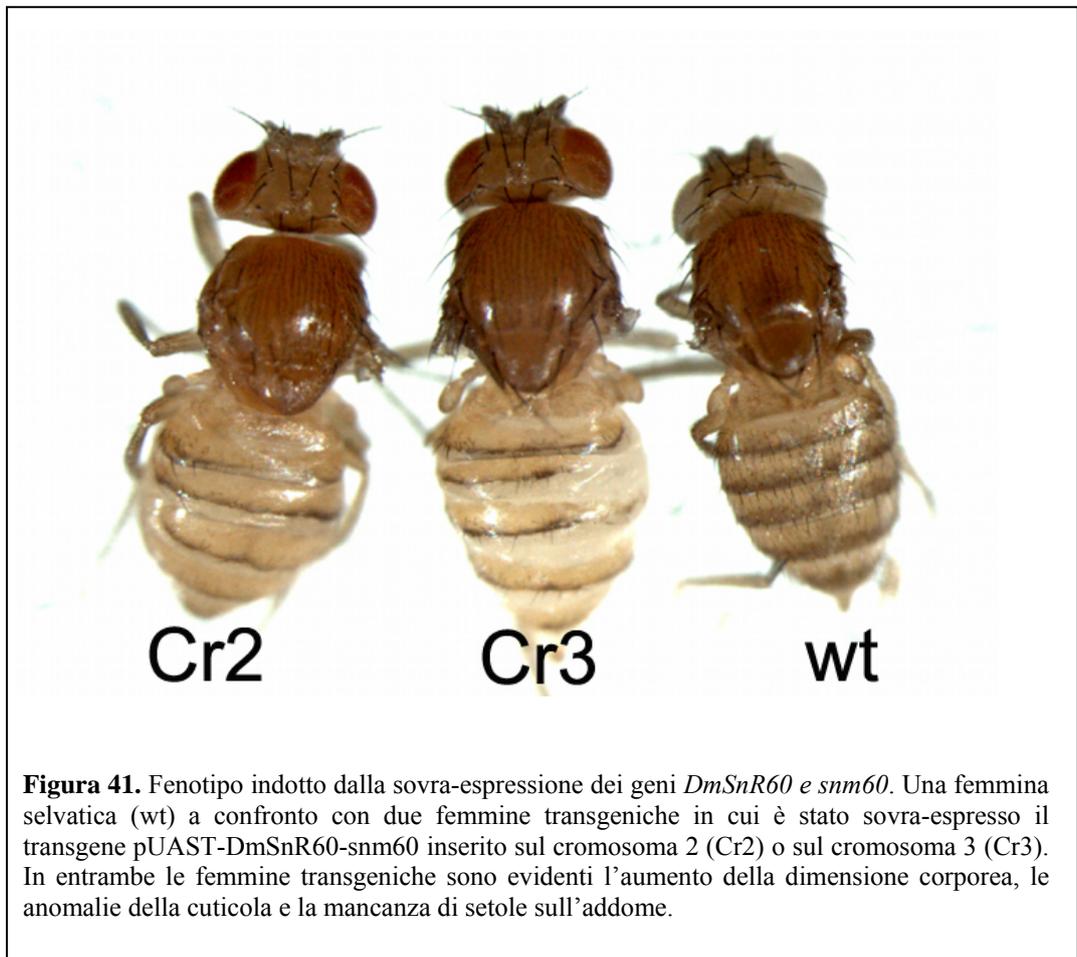
Un frammento genomico di circa 2.0 kb, contenente i geni *DmSnR60* e *snm60*, è stato amplificato per PCR e successivamente clonato nel vettore plasmidico pMOS blue.

I moscerini derivati dagli embrioni iniettati con il transgene pUAST:DmSnR60-snm60 sono stati incrociati singolarmente con moscerini del ceppo *w* (occhio bianco); dalla progenie di questo incrocio sono state selezionati gli individui transgenici (che mostravano colore dell'occhio *w+*). In ciascuna linea transgenica il cromosoma interessato dall'inserzione è stato identificato e messo in eterozigosi con l'opportuno cromosoma bilanciatore mediante appositi incroci. Infatti, affinché un cromosoma in cui si è inserito un transgene possa essere mantenuto stabilmente nelle generazioni successive esso deve essere posto in eterozigosi con un cromosoma omologo "bilanciante", che cioè è portatore di inversioni tali da non consentire fenomeni di ricombinazione. Ogni cromosoma "bilanciante" porta anche un marcatore fenotipico dominante che consente di riconoscerlo facilmente dall'omologo recante l'inserzione del transgene e da un letale recessivo, che impedisce la possibilità che il bilanciante venga portato in omozigosi. Dall'iniezione del transgene pUAST:DmSnR60-snm60 sono state ottenute quattro linee, delle quali solo due sono risultate derivare da differenti eventi di inserzione. Una di queste linee, recante un'unica inserzione del transgene sul cromosoma 3, è stata selezionata per i successivi esperimenti di sovra-espressione. Individui omozigoti per l'inserzione del transgene pUAST:DmSnR60:snm60 sono stati incrociati con individui che portavano sul cromosoma 2 il gene *Gal4* sotto il controllo del promotore costitutivo dell'actina, e gli individui della progenie che presentavano entrambi i transgeni sono stati poi selezionati sulla base della marcatori fenotipici portati dai rispettivi cromosomi bilanciatori. In figura 40 è riportato lo schema degli incroci effettuati.

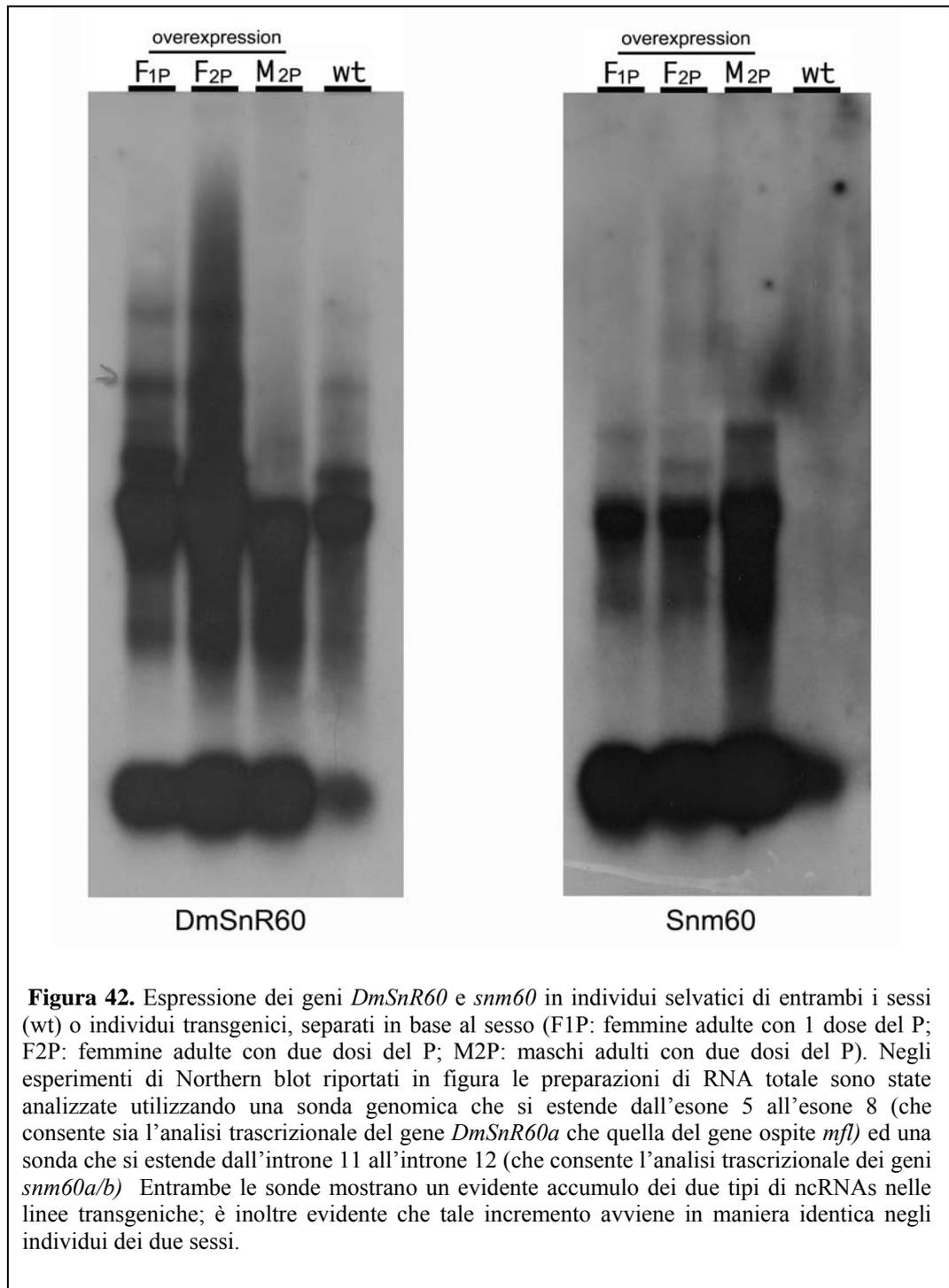


L'osservazione allo stereomicroscopio degli individui portatori di entrambi i transgeni ha evidenziato che la sovra-espressione costitutiva dei geni *DmSnR60* e *snm60* provoca un sensibile incremento della dimensione corporea, visibili alterazioni della cuticola addominale e forte riduzione del numero, spessore e lunghezza delle setole addominali (vedi fig. 41). E' spesso evidente anche la presenza di vere e proprie fessure a livello della cuticola, una possibile testimonianza di una crescita cellulare incontrollata. L'insieme di queste caratteristiche fenotipiche si presenta con una frequenza di circa il 10% nelle femmine adulte, mentre la progenie maschile non manifesta alcun apprezzabile fenotipo mutante. E' interessante notare che alcuni tratti, quali le anomalie a carico della cuticola addominale e la forte riduzione del numero, spessore e lunghezza delle setole, sono caratteristici anche del fenotipo del mutante

ipomorfo *mfl^l*. Al contrario, l'aumento del peso e della dimensione corporea costituiscono elementi fenotipici esattamente opposti, dato che i mutanti *mfl^l* sono invece tipicamente caratterizzati da una sensibile riduzione delle dimensioni corporee. Per verificare che il fenotipo osservato fosse direttamente causato dalla sovra-espressione dei piccoli ncRNA intronaci, e non dalla presenza di eventuali mutazioni provocate dall'inserzione del transgene, individui omozigoti della seconda linea transgenica, in cui l'inserzione del costrutto pUAST:DmSnR60-snm60 era avvenuta sul cromosoma 2, sono stati incrociati con individui recanti sul cromosoma 3 l'attivatore Gal4 sotto il controllo del promotore costitutivo della tubulina.



L'osservazione della progenie ottenuta da quest'incrocio, oltre a confermare la presenza del fenotipo già descritto (vedi fig. 41), ha ulteriormente confermato che esso si manifesta esclusivamente nelle femmine, mentre la progenie maschile è invariabilmente caratterizzata dal fenotipo selvatico.

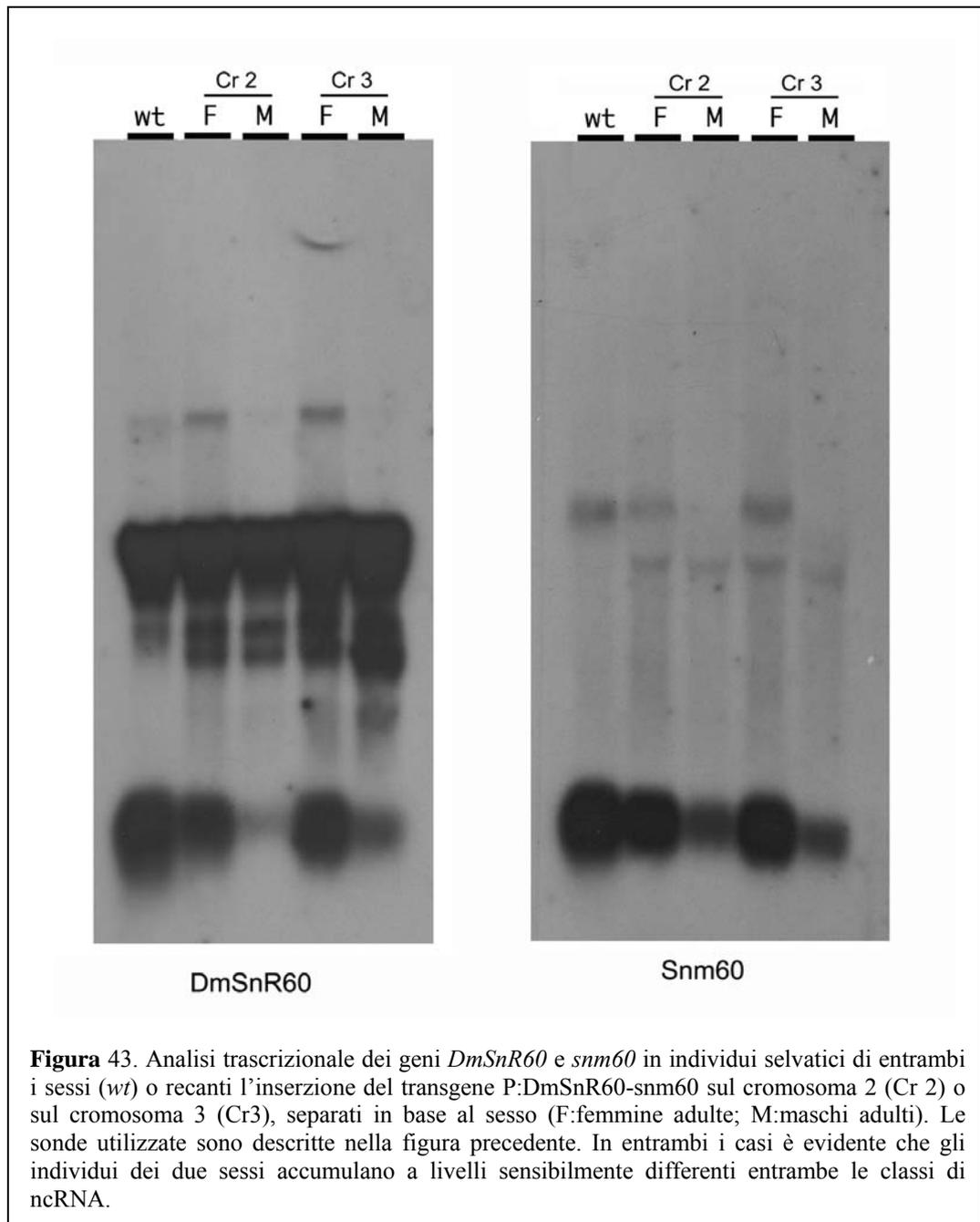


Per controllare l'effettivo aumento dei livelli d'espressione dei geni *DmSnR60* e *snm60*, l'RNA totale estratto da maschi e femmine recanti entrambi i transgeni "attivatore" e "bersaglio" è stato analizzato con sonde specifiche.

Per l'analisi dell'espressione dei geni *DmSnR60* è stata utilizzata una sonda genomica più lunga rispetto alle precedenti, che si estende dall'esone 5 all'esone 8, nel tentativo di saggiare anche il livello generale di accumulo degli mRNA da 1.8, 2.0 e 2.2 kb. L'espressione dei geni *snm60* è stata invece saggiata con una sonda interna all'introne 12.

Come atteso, il livello quantitativo delle molecole *DmSnR60* e *snm60* risulta fortemente aumentato negli individui transgenici (fig. 42). E' inoltre evidente che quest'incremento è identico nei due sessi mentre, come descritto precedentemente, nel ceppo selvatico essi si differenziano sensibilmente per l'espressione di entrambi i tipi di ncRNA. Nelle linee transgeniche questa differenza quantitativa risulta invece del tutto mascherata dall'espressione ectopica del transgene attivato. Per controllare ulteriormente che la sovra-espressione dei geni *DmSnR60* e *snm60* fosse direttamente controllata dal transgene attivatore Gal4 ho analizzato mediante Northern blot anche preparazioni di RNA totale ottenuto da moscerini adulti, separati in base al sesso, recanti il solo transgene "bersaglio". Come si può chiaramente evincere dalla figura 43, in entrambe le linee transgeniche indipendenti (che portano l'inserzione di pUAST:*DmSnR60*-*snm60* sul cromosoma 2 o sul 3) non è osservabile alcuna alterazione dei livelli di espressione dei geni *DmSnR60* e *snm60*, ed è ben visibile il differente profilo che caratterizza i due sessi. Di

conseguenza, il fenotipo osservato può essere con certezza attribuito all'attivazione ectopica dei geni *DmSnR60* e *snm60*.



Un dato inaspettato è che, pur essendo stato analizzato RNA estratto da femmine eterozigoti o omozigoti per l'inserzione del costrutto

pUAST:DmSnR60-snm60, i livelli di accumulo delle molecole DmSnR60 e snm60 sono risultati del tutto paragonabili fra gli individui che portano una o due copie del transgene (fig. 42).

Analogamente, nessuna differenza di espressività fenotipica è stata notata fra le femmine portatrici di una o due copie del transgene, un dato che rimane essenzialmente incompreso.

Devo anche aggiungere che gli esperimenti di Northern blot riportati in figura 42 non mi hanno permesso di trarre alcuna conclusione certa sui possibili effetti causati dalla sovra-espressione delle molecole DmSnR60 e snm60 sulla produzione dei diversi tipi di mRNA che derivano dal gene *mfl*. Infatti, come già descritto, le sonde contenente in geni *DmSnR60* cross-ibridano fortemente con l'rRNA 28S, la cui mobilità elettroforetica è essenzialmente molto simile a quella dei trascritti del gene *mfl*; di conseguenza, il forte segnale di ibridazione in questa zona del filtro non mi ha permesso un'analisi sufficientemente informativa. Per quanto riguarda invece la sonda specifica per i geni *snm60*, essa ha messo in evidenza l'accumulo di un trascritto di circa 2.0 kb che non è invece osservato negli individui *wt*. Ulteriori esperimenti sono però necessari per chiarire se questo segnale d'ibridazione corrisponde all'aumentato accumulo di un trascritto endogeno o dalla presenza di un precursore che deriva specificamente dal transgene utilizzato. Intendo quindi svolgere ulteriori esperimenti di Real-Time RT-PCR con l'ausilio di oligo specifici per ottenere indicazioni più precise.

Discussione

Discussione

La produzione di trascritti multipli con differenti proprietà codificanti è uno degli aspetti che contribuisce maggiormente alla complessità del genoma degli organismi eucariotici. L'identificazione di tutti i differenti trascritti che si originano da un gene può quindi essere essenziale per raccogliere importanti indicazioni sulle sue funzioni biologiche. Per quel che riguarda il gene *minifly*, la dettagliata analisi trascrizionale svolta nel corso della mia tesi di dottorato ha evidenziato come la sua organizzazione molecolare fosse ben più complessa di quanto inizialmente sospettato. Il repertorio di trascritti sia codificanti che non codificanti che deriva da questo gene è risultato infatti essere molto articolato e vario. Va sottolineato fra l'altro che fino ad oggi *mfl* è l'unico membro della famiglia genica delle pseudouridina sintasi per cui siano state descritte forme alternative di mRNA. In primo luogo, la proteina MFL canonica può essere prodotta da tre differenti mRNA che si distinguono solo per l'eterogeneità delle loro sequenze 3' UTR e per il loro livello di espressione, che per due di essi è sensibilmente più marcato in femmine. Le differenti regioni 3' UTR che caratterizzano i tre mRNA da 1.8, 2.0 e 2.2 kb potrebbero svolgere un ruolo importante nel regolare il loro profilo di espressione, dato che è ben noto come queste regioni siano spesso bersaglio di specifici fattori regolativi, quali proteine e/o microRNA. Per quel che riguarda l'eventuale interazione con le molecole di microRNA, una preliminare analisi bioinformatica ha confermato la possibilità che le sequenze 3'UTR che caratterizzano i diversi trascritti del gene potrebbero effettivamente rappresentare un bersaglio per la regolazione post-trascrizionale operata da questa classe di molecole; al contrario, la ricerca di eventuali sequenze capaci legare specifici fattori proteici non ha fornito alcuna apprezzabile indicazione.

Un' ulteriore ipotesi sul ruolo regolativo svolto dalle sequenze 3' UTR è stato recentemente suggerito da Ledda et al. (2005), che hanno indicato come per la classe degli mRNA 5' TOP la stringenza della regolazione traduzionale sembra essere inversamente proporzionale alla lunghezza della 3'UTR. In accordo con quest'ipotesi, è possibile considerare che la lunghezza stessa delle diverse sequenze 3'UTR potrebbe rappresentare un elemento regolativo importante. Quindi, la diversa estensione delle sequenze 3'UTR degli mRNA da 1.8, 2.0 e 2.2 kb prodotte dal gene *mfl*, che appartiene alla classe dei geni 5'-TOP, potrebbe di *per se* influenzarne la stabilità e/o l'efficienza della traduzione in differenti condizioni di crescita.

L'analisi trascrizionale da me effettuata ha anche portato a stabilire che un nuovo trascritto da 1.0 kb, che si origina per splicing alternativo e che termina in corrispondenza del sito di poliadenilazione più distale, codifica una rara subforma proteica cui è stato dato il nome di MFL α . Questa subforma rappresenta il primo esempio di variante proteica codificata da un membro della famiglia delle pseudouridina sintasi eucariotiche.

MFL α condivide con la più abbondante proteina canonica un segmento di 211 aa della regione ammino-terminale, in cui sono contenuti un segnale di localizzazione nucleare ed i due domini catalitici altamente conservati, TrubI e TrubII. La caratteristica distintiva di questa nuova subforma è la presenza di una corta sequenza C-terminale specifica, lunga 43 aa, che rimpiazza il ben più ampio dominio carbossi-terminale della proteina MFL, in cui sono presenti il dominio RNA-binding denominato PUA, nonché il dominio tyr ed il segnale bipartito di localizzazione nucleare. Poiché è stato recentemente dimostrato che nella proteina aCbf5 di Archaea il dominio PUA è essenziale per il legame con gli snoRNA che agiscono da guida nel processo di pseudouridilazione, (Aravind et al., 1999;

Raschid et al., 2006), la mancanza di questo dominio potrebbe avere conseguenze importanti sulla funzionalità della subforma MFL α , lasciando supporre che essa possa partecipare con scarsa efficienza alla formazione di complessi snoRNP. Sarà quindi importante stabilire se la nuova subforma sarà in grado di legare efficientemente le altre componenti proteiche dei complessi snoRNP, perché in questo caso MFL α potrebbe svolgere un ruolo di regolazione negativa sull'attività del gene *mfl*.

Un altro elemento distintivo di questa nuova subforma è rappresentato dall'assenza del segnale di localizzazione nucleare carbossi-terminale. Data la presenza di un ulteriore segnale di localizzazione nucleare all'estremità ammino-terminale, questa caratteristica non dovrebbe abolire la capacità della proteina di localizzarsi nel nucleo, anche se potrebbe tuttavia influenzarne l'efficienza di trasporto oppure la sua ritenzione, o ancora la sua localizzazione subnucleare. A tal proposito è interessante notare che nell'uomo la delezione di un cluster ricco in lisine presente all'estremità C-terminale della discherina non annulla la capacità di questa proteina di localizzarsi nel nucleo, ma ne influenza il tasso di trasporto (Heiss et al., 1999), mentre in lievito la delezione del dominio basico C-terminale della proteina Cbf5 porta ad un allungamento dell'intervallo G2/M del ciclo cellulare (Jiang et al., 1993). Il confronto delle sequenze amminoacidiche di MFL e MFL α ha anche indicato che la porzione C-terminale assente in MFL α è ricca di putativi siti di fosforilazione, sottolineandone un ulteriore punto di distinzione. Poiché è ben noto che numerosi processi cellulari, tra i quali la sintesi proteica e la divisione cellulare, sono controllati mediante un meccanismo di fosforilazione dipendente dalla progressione del ciclo cellulare, è possibile ipotizzare che questo tipo di modifica post-traduzionale possa regolare in maniera differenziale le due subforme MFL ed MFL α in risposta ai diversi stimoli cellulari.

E' anche interessante notare che, pur essendo le sequenze aminoacidiche delle pseudouridina sintasi eucariotiche altamente conservate, la maggiore variabilità risiede nelle regioni C-terminali. Quest'osservazione suggerisce la possibilità che queste regioni potrebbero svolgere un ruolo nello stabilire specifiche interazioni funzionali con fattori proteici accessori. In analogia, la regione C-terminale specifica della subforma MFL α potrebbe fornire a questa proteina la capacità di interagire con differente efficienza e/o specificità con partner proteici differenti, eventualmente partecipando alla formazione di specifici complessi snoRNP o, alternativamente, formando complessi con funzioni totalmente differenti.

Una caratteristica poco chiara della subforma MFL α riguarda il suo profilo di espressione, che nei moscerini adulti appare essere molto più abbondante in femmine. Sebbene questa modalità di espressione sia caratteristica di molte proteine che sono necessarie al corretto svolgimento sia dell'oogenesi che dei primi stadi dello sviluppo, è comunque prematuro escludere che in differenti stadi la proteina MFL α possa accumularsi in maniera identica nei due sessi, o che nei maschi il suo profilo d'espressione possa essere di tipo transiente o tessuto specifico.

In realtà, un livello di espressione più elevato in femmine appare essere una caratteristica generale di tutti i trascritti del gene *mfl*, sia codificanti che non codificanti, ad eccezione dell'mRNA di 1.8 kb, la cui abbondanza è identica tra maschi e femmine (Giordano et al., 1999). Quest'osservazione è anche in accordo con le indicazioni riportate in un recente lavoro che include *mfl/Nop60b* nella lista dei geni di *Drosophila* la cui espressione risulta significativamente incrementata in cellule staminali purificate dalla linea germinale femminile (Kai et al., 2005). Bisogna inoltre notare che durante il ciclo vitale della *Drosophila* i trascritti del gene *mfl* presentano un picco quantitativo di espressione negli stadi e nei tessuti

maggiormente caratterizzati da crescita e divisione cellulare, quali gli embrioni, le larve e gli ovari di femmine adulte. Quest'osservazione spingerebbe ad ipotizzare che l'espressione del gene possa essere strettamente legata al tasso di crescita e di proliferazione cellulare, un'ipotesi in accordo con il ruolo biologico essenziale giocato dalla proteina MFL nella biogenesi dei ribosomi (Giordano et al., 1999). Tale ipotesi è ulteriormente sostenuta dai risultati di recenti analisi di microarray che indicano *mfl/Nop60B* come un target preferenziale dell'induzione trascrizionale diretta dal gene *d-myc* (Grewal et al., 2005). Sulla base di queste considerazioni e della nozione che in femmine adulte l'ovario può rappresentare fino a metà dell'intero peso corporeo, è possibile ipotizzare che i differenti livelli di espressione della proteina MFL α osservati nei due sessi potrebbero essere solo una conseguenza del differente tasso di sintesi proteica, noto essere marcatamente più elevato nel corso dell'oogenesi.

Un aspetto particolarmente interessante emerso dall'analisi trascrizionale del gene *mfl* riguarda il variegato repertorio di trascritti codificanti e non codificanti che derivano da questo gene. Quattro isoforme di un gene che codifica per uno snoRNA della famiglia C/D (DmSnR60 a/d) sono infatti localizzate all'interno degli introni presenti nelle regioni 3'UTR, rispettivamente negli introni 6, 7, 8 e 9. Inoltre, negli introni 10 ed 11, caratteristici dei due mRNA (di 2.2 e 1.0 kb) che terminano in corrispondenza del sito di poliadenilazione più distale, sono ospitate due copie di un piccolo gene privo di motivi di sequenza e di struttura noti. In accordo con una nomenclatura recentemente proposta (Huttenhofer et al., 2002), le molecole prodotte da questi geni sono state quindi definite "small non messenger RNA" derivate dalla regione politenica 60 (snm60). La mancanza di caratteristiche strutturali note o di motivi di sequenza conservati rende difficile l'analisi funzionale di queste molecole, mentre è stato ovviamente più facile attribuire una specifica

funzione agli snoRNA DmSnR60. Queste molecole presentano, infatti, due lunghi elementi antisenso che mostrano complementarità perfetta con l'rRNA 28S di *Drosophila*, ed è stato possibile confermare l'effettiva metilazione di entrambi i residui bersaglio. L'analisi dell'espressione dei geni *DmSnR60a/d* e *snm60a/b* mediante esperimenti di ibridazione *in situ* su intestini larvali e su ovari adulti ha poi fornito ulteriori utili informazioni. Infatti, mentre gli snoRNA DmSnR60 hanno mostrato un profilo di espressione costitutivo ed ubiquitario, nonché una localizzazione prettamente nucleolare, in perfetto accordo con la loro funzione, le molecole *snm60* non sono risultate espresse in maniera ubiquitaria ed hanno mostrato una fine regolazione sia spaziale che temporale. All'interno dell'ovaio, l'espressione dei geni *snm60a/b* è risultata evidente solo nelle fasi finali dell'oogenesi e solo nelle cellule nutrici, che appartengono alla linea germinale, e non in quelle follicolari, che appartengono alla linea somatica. Infine, le molecole *snm60* si accumulano nei nuclei, ma non si concentrano nei nucleoli. Nelle cellule proliferanti dell'intestino, i geni *snm60a/b* appaiono essere preferenzialmente espressi in gruppi di cellule che sembrano essere in attiva divisione. La fine regolazione osservata per l'espressione dei geni *snm60a/b* suggerisce quindi che la loro trascrizione non sia dovuta semplicemente a "transcriptional noise", bensì finalizzata ad un preciso ruolo biologico. Il potenziale appaiamento notato tra la molecola *snm60b* ed il pre-mRNA del gene *mfl* apre la possibilità di una correlazione funzionale tra l'espressione di questo piccolo ncRNA e quella del gene ospite, anche se non si può escludere che l'eventuale ruolo biologico di questa molecola possa essere totalmente scorrelato da quello di *mfl*. Al contrario, per i geni intronici *snoH1* e *DmSnR60* la correlazione funzionale con il gene ospite è ben definita, dal momento che questi geni sono stati dimostrati direttamente

responsabili della modificazione di residui altamente conservati sull'rRNA 18S e 28S, e quindi intervengono anch'essi nel processo di biogenesi dei ribosomi.

I primi approcci di mutagenesi *in vivo* per verificare il ruolo biologico delle molecole *snm60* sono attualmente in corso, ed è interessante notare che la sovra-espressione ubiquitaria e costitutiva di queste molecole mediante il sistema binario GAL4/UAS in un background selvatico è in grado di indurre la comparsa di un fenotipo caratterizzato da “overgrowth”. Un ulteriore aspetto emerso dall'analisi trascrizionale è poi costituito dalla presenza di un lungo trascritto poliadenilato di 4.4 kb che è evidenziato prevalentemente in preparazioni di RNA ottenute da embrioni e femmine adulte. Sebbene in esperimenti di Northern blot questo trascritto sia identificato sia da sonde esoniche che introniche, né in precedenti screening di librerie di cDNA né nel corso degli esperimenti di 3' RACE ed RT-PCR descritti in questa tesi è stato possibile ottenere informazioni sulla sua struttura. Una possibile ipotesi, basata sull'osservazione che esso include sequenze introniche e che la sua lunghezza è identica a quella del segmento genomico compreso tra il sito di inizio della trascrizione del gene *mfl* ed il più distale dei suoi siti di poliadenilazione, è che essa rappresenti un precursore non maturato. Poiché la presenza di trascritti non o parzialmente maturati è una caratteristica comune dei geni che ospitano nei propri introni geni per ncRNA, è possibile ipotizzare che l'accumulo di questo trascritto possa essere correlato alla biogenesi dei piccoli ncRNA intronici mediante il meccanismo “splicing indipendente”.

In conclusione, la complessa organizzazione molecolare evidenziata per il gene *mfl* suggerisce che le funzioni biologiche svolte da questo gene siano legate all'appropriata e dinamica produzione di tutti i suoi prodotti, siano essi proteine o molecole di ncRNA.

E' da notare che l'organizzazione mostrata da *mfl* mostra forti analogie con quella di molti geni umani localizzati sul cromosoma 21 e 22, per i quali è stato dimostrato che trascritti non codificanti derivano frequentemente dalle rispettive regioni 3' (Cowley et al., 2005). Quest' aspetto, oltre ad essere per se stesso interessante, potrebbe fornire indicazioni utili per lo studio degli ortologi del gene *mfl* nei mammiferi. Infatti, mutazioni nell'ortologo umano, il gene *DKCI*, causano una rara sindrome nota come discheratosi congenita X-linked, che colpisce cute mucose e midollo osseo e le cui basi molecolari sono ancora da chiarire. Le connessioni strutturali e funzionali emerse nel corso degli anni tra il gene *mfl* di *Drosophila* ed il gene umano *DKCI* sono davvero sorprendenti. Questi due geni ortologi, oltre a codificare proteine altamente correlate, mostrano infatti anche una simile e complessa organizzazione genomica. Per esempio, Huttenofer et al. (2001) hanno identificato un gene che codifica per uno snoRNA della classe H/ACA nell'ottavo introne del gene *DKCI*, dimostrando che la proprietà di produrre ncRNA intronaci è un' ulteriore caratteristica che accomuna questo gene al suo ortologo di *Drosophila*. Le similitudini sono ulteriormente accresciute dall'osservazione che anche *DKCI* produce un lungo trascritto di circa 4.5 kb la cui struttura non è stata ancora definita, probabilmente a causa della sua scarsa abbondanza o di un profilo di espressione molto ristretto (Heiss et al., 2000). Queste somiglianze, accoppiate alla disponibilità di potenti e sofisticate tecniche di analisi genetica, rendono la *Drosophila* un ottimo sistema modello per lo studio della discheratosi congenita X-linked. In effetti, l'analisi genetica e funzionale dei mutanti *mfl* di *Drosophila* ha concretamente anticipato molte delle osservazioni successivamente riportate per i topi in cui è stato deletato, tutto o in parte, il gene *DKCI*. Alcuni anni fa, la caratterizzazione molecolare e funzionale di mutanti *mfl* da perdita di funzione indicò che alleli nulli causavano letalità larvale, mentre alleli

ipomorfi erano vitali e causavano una serie di difetti pleiotropici, quali la riduzione della dimensione corporea, ritardo di sviluppo, difetti della cuticola addominale, riduzione nel numero, lunghezza e spessore delle setole ed infine ridotta fertilità femminile (Giordano et al., 1999). Successivamente, anche in topi knock-out per il gene *DKC1* è stata evidenziata letalità (He et al., 2002), mentre alleli ipomorfi risultavano vitali e ricapitolavano le caratteristiche cliniche della DC, quali la suscettibilità a sviluppare tumori (Ruggero et al., 2003). Inoltre la letalità è stata osservata solo quando l'allele deleto era di origine materna (He et al., 2002), rivelando per la prima volta l'esistenza di un effetto materno nella trasmissione di questa malattia. E' interessante notare che la ridotta fertilità femminile e la presenza di effetto materno erano tratti fenotipici già descritti per i mutanti *mfl*. Infatti, negli ovari delle femmine mutanti *mfl*, i nuclei delle cellule nutrici apparivano condensati e di forma irregolari e gli embrioni in degenerazione presentavano alterazioni morfologiche caratteristiche del processo apoptotico. Questi dati avevano anche suggerito un possibile coinvolgimento dell'apoptosi nella progressione della malattia (Giordano et al., 1999). A sostegno di quest'ipotesi sono anche i dati ottenuti successivamente, inducendo RNAi *in vivo*. Mediante l'attivazione di transgeni capaci di silenziare specificamente il gene *mfl* in tessuti specifici è stato infatti dimostrato che l'inattivazione del gene era sufficiente per indurre la morte cellulare localizzata (Giordano et al., 2001). Analogamente ai mutanti *mfl*, la maggior parte degli embrioni degenerati di topo che presenta l'allele materno *DKC1* deleto ha presentato nuclei apoptotici nel trofoblasto (He et al., 2002). In *Drosophila*, l'esistenza di effetto materno è spiegata a livello molecolare dall'osservazione che il gene produce, oltre ad un mRNA costitutivo, almeno tre trascritti addizionali che si accumulano preferenzialmente in ovari e sono forniti per via materna all'embrione in sviluppo. Sebbene nei mammiferi non sia stato ancora

caratterizzato alcun trascritto alternativo del gene *DKCI*, è interessante notare che nei topi knock-out l'effetto letale può essere causato anche da una piccola delezione che copre solo la regione 3' del gene (He et al., 2002), e che delezioni della regione 3' sono state riscontrate in alcuni pazienti affetti dalla malattia. Queste osservazioni sottolineano la possibilità che anche per il gene umano *DKCI* la regione 3'-terminale possa svolgere un ruolo biologico rilevante. La similarità fra il gene *mfl* ed il gene *DKCI* si estende anche a livello regolativo. Infatti, dati recenti hanno indicato che l'espressione di entrambi i geni è regolata dall'oncogene *myc*. La famiglia dei geni *myc* è strettamente coinvolta nella regolazione della crescita e del differenziamento cellulare, nonché nell'apoptosi e nello sviluppo di tumori. Le proteine Myc agiscono come fattori di trascrizione modulando l'espressione di geni che promuovono la crescita cellulare, in special modo controllando il biometabolismo e stimolando la sintesi delle proteine. Recentemente è stato dimostrato che *myc* controlla diversi aspetti della biogenesi dei ribosomi, sia in cellule di *Drosophila* che in cellule di mammifero, regolando la trascrizione mediata da tutte e tre le RNA polimerasi (vedi Oskarsson e Trump, 2005). In particolare, nei mammiferi *c-myc* influenza la sintesi dell'rRNA, la maturazione del pre-rRNA e presenta una localizzazione nucleolare dopo l'induzione da siero di cellule quiescenti. Nel nucleolo, la proteina c-Myc si associa all'rDNA per stimolarne direttamente la trascrizione. In cellule di *Drosophila*, il controllo della sintesi dello rRNA sembra avvenire in modo più indiretto. Utilizzando esperimenti basati sull'uso di microarrays, Grewal et al. (2005) hanno mostrato infatti che un largo numero di geni di *Drosophila* coinvolti nella biogenesi dei ribosomi, tra i quali il gene *mfl/Nop60b*, sono regolati da *d-myc*. D'altra parte, in *Drosophila* i mutanti ipomorfi di *d-myc* (chiamati anche *diminutive: dm*; Scribeber-Agus et al., 1997) presentano sterilità femminile, ridotta taglia corporea, ritardo di crescita e

riduzione nel numero delle setole. L' insieme di queste caratteristiche fenotipiche è essenzialmente identico a quello causato dalle mutazioni ipomorfe del gene *mfl*, sottolineando ulteriormente la correlazione funzionale esistente tra questi due geni.

Materiali e Metodi

Materiali e Metodi

Ceppi di *Drosophila melanogaster*

- Il ceppo white (occhi bianchi) di *Drosophila melanogaster* è stato usato come ceppo di riferimento.
- Ceppo 3664: genotipo $yw;ry506 Sb^1 P\{ry \Delta 2-3\}99B/TM6$. Questi elementi sono capaci di produrre elevate quantità di trasposasi, ma non possono trasporre essendo privi delle sequenze ripetute ed invertite di 31bp; contengono sequenze del ceppo selvatico rosy e sequenze del gene che codifica la trasposasi. Il gene per la trasposasi presente nell'elemento è mutagenizzato, essendo privo dell'esone 3. In questo ceppo il fenotipo del colore dell'occhio $ry+$ è "mascherato" dal background white. Sb=Stable: peluzzi sullo scutello più corti. Cy=Curly:ali arricciate.

Ciclo vitale di *Drosophila melanogaster* a 25°C

STADIO	ORE	GIORNI
Embrione	0-24	0-1
Larva 1° stadio	24-48	1-2
Larva 2° stadio	48-72	2-3
Larva 3° stadio	72-120	3-5
Prepupa	120-132	5-5.5
Pupa	132-216	5.5-9
Adulto	216	9-10

Terreno per la coltura in massa di *D.melanogaster*

Il terreno su cui vivono i moscerini nel nostro laboratorio viene preparato utilizzando un terreno per la coltura di insetti (Instant *Drosophila medium*) fornito dalla Sigma-aldrich. Per la preparazione di 1 lt di terreno di coltura vengono aggiunti ad un litro di acqua, 300 gr. di Instant *Drosophila medium*, 10 gr. di agar e il tutto viene portato ad ebollizione; si lascia bollire per 5' e si aggiungono successivamente 2.5 gr. di metil p-idrossibenzoato, sciolto in una appropriata quantità di alcool etilico 95%, come antimuffa. Il terreno di coltura così preparato viene poi versato in appositi barattoli cilindrici utilizzati per il mantenimento dei ceppi di *Drosophila*. Si lascia raffreddare coprendo i barattoli con della garza per evitare che possano verificarsi contaminazioni del terreno e si aspetta fino a che il terreno sia completamente asciutto; infine si tappano i barattoli con del cotone da batteriologia.

Mantenimento dei ceppi

I ceppi vengono fatti crescere in camere termo-regolate ad una temperatura di 22°C in barattoli contenenti il terreno suddetto e vengono trasferiti su terreno fresco ogni 14/15 giorni. Il trasferimento viene operato previa analisi al microscopio del fenotipo dei moscerini per accertarsi che nessun individuo con caratteristiche differenti da quelle corrispondenti al ceppo in esame sia presente nel barattolo. Si selezionano poi una trentina di individui di ciascun sesso e si trasferiscono in un nuovo barattolo. Il numero di individui da trasferire non è fisso ma dipende dal genotipo dei moscerini e un numero maggiore è richiesto se essi sono caratterizzati da bassa fertilità.

Analisi delle sequenze in banche dati

Le banche dati rappresentano il materiale sperimentale della bioinformatica. Esse possono essere classificate in quattro grosse categorie: Banche Dati di biosequenze; Banche Dati genomiche; Banche Dati di espressione genica e Banche Dati di proteomica. La Banca Dati del National Center for Bioinformaic Information, comunemente chiamata NCBI, è il classico e più importante esempio di database per gli studi computazionali in biologia molecolare, in cui sono sviluppati programmi appositi per l'analisi delle sequenze genomiche in essa presente e ricca di informazioni biomediche per meglio conoscere i processi molecolari che sono alla base delle malattie umane. Nel mio lavoro di tesi ho utilizzato ampiamente il programma BLAST (**B**asic **L**ocal **A**lignment **S**earch **T**ool) che, grazie ad un algoritmo specifico, ottimizza l'operazione di allineamento di 2 o più sequenze. Questo programma mi ha permesso di selezionare i primer idonei all'amplificazione tramite PCR delle sequenze del gene *mfl*. La banca dati Flybase (www.flybase.org) ha fornito i dati di sequenza e molte altre informazioni specifiche sull'organizzazione genica di *D. melanogaster*. La banca dati "Flybase" è un database genetico e molecolare ed include dati su tutte le specie appartenenti alla famiglia delle *Drosophilidae*, tra cui *Drosophila melanogaster*. La banca dati Flybase è prodotta da un consorzio di ricercatori del "National Institutes of Health" (USA) e del "Medical Research Council" (London). Il consorzio include anche biologi Drosophilisti e bioinformatici dell'Università di Harvard, di Cambridge, dell'Indiana e della California. "Flybase" contiene informazioni sull'organizzazione, l'espressione e la funzione genica, le proprietà di trascritti e di proteine e la funzione dei geni noti con accurati ed aggiornati riferimenti alla letteratura scientifica. Essa comprende inoltre i dati derivanti dai Progetti Genomici di Drosophila. Per la ricerca di sequenze funzionali note all'interno delle sequenze 3'UTR dei trascritti del gene *mfl* è stato utilizzato il database UTRsource disponibile previa registrazione al sito www.ba.itb.cnr.it/UTR (Mignone F, Grillo G, Licciulli F, Iacono M, Liuni S, Kersey PJ, Duarte J, Saccone C, Pesole G. UTRdb and UTRsite: a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. 2005. Nucleic Acids Res. 33). La ricerca nelle sequenze 3'UTR di putativi siti di legame per microRNA è stata condotta con un programma di predizione disponibile in rete, chiamato *MicroInspector* (<http://www.imbb.forth.gr/microinspector>). Le analisi sono state condotte in accordo con le indicazioni fornite dagli ideatori del programma e disponibili nello stesso sito. Infine, La conservazione filogenetica della regione gnomica in 3' del gene *mfl* è stata ottenuta sottomettendo la sequenza d'interesse all'UCSC genome Browser (<http://genome.ucsc.edu>).

Protocollo per l'estrazione del DNA genomico di *D. melanogaster*

Si raccoglie un numero adeguato di mosche (circa 30-40 individui) in tubi Eppendorf da 1.5ml. Si pestano le mosche in 200 µl iniziali di buffer di omogenizzazione poi si porta il volume a 500µl finali. Si centrifuga per 1 min a 1000 rpm. Si preleva il surnatante, facendo attenzione a non prendere le carcasse e si centrifuga per 5 min a 8000 rpm, per far precipitare i nuclei. A fine centrifugazione si elimina il surnatante ed al pellet si aggiungono 500µl di buffer e la proteinasi K ad una concentrazione finale di 100 µg/ml. Si aggiungono 50 µl di SDS 10% ed si incuba a 37°C per 45-60 min. Terminata l'incubazione si addiziona un ugual volume di fenolo e si agita per inversione per 10 min. Si centrifuga per 5 min a 13200 rpm e si preleva la fase acquosa superiore. Si ripetono gli ultimi tre punti ancora una volta. Si preleva la fase superiore e si aggiunge un ugual volume di cloroformio/alcool iso-amilico e si centrifuga per 5 min a 13200 rpm. Si preleva la fase acquosa superiore e procede alla precipitazione degli acidi nucleici aggiungendo NaCl ad una concentrazione finale di 200 mM e due volumi di etanolo 95%. Il campione viene tenuto a -20°C per almeno 1 ora. Finita la precipitazione, si centrifuga per 15 min a 13200 rpm, si elimina l'etanolo lasciando indisturbato il "pellet" e si aggiungono 200µl di etanolo 70%. Si centrifuga per 5 min a 13200 rpm, si elimina l'etanolo 70% e si lascia asciugare il "pellet". Si risospende il "pellet" in un adeguato volume di TE e si conserva a 4°C.

Amplificazione di molecole di DNA mediante la reazione a catena della Taq polimerasi (PCR)

La tecnica di PCR (Polymerase Chain Reaction) permette di amplificare in vitro una regione specifica di DNA.

Essa necessita di due oligonucleotidi (corte sequenze di 15-25 nt) come inneschi complementari all'estremità 5' di ciascuno dei due filamenti del frammento da amplificare. Le estremità 3'OH degli oligonucleotidi servono da innesco per una DNA polimerasi che sintetizza i filamenti complementari. La polimerasi utilizzata è quella di *Thermus aquaticus* (Taq polimerasi) che è termostabile e, quindi, non degradabile nella tappa di denaturazione a 95°C.

A seconda della finalità dell'amplificazione possono essere utilizzati diversi tipi di polimerasi. La miscela di reazione è costituita da

- un DNA che funge da stampo, la cui quantità richiesta è sempre inferiore a 1µg;
- una Taq polimerasi e il suo buffer specifico;
- MgCl₂ che attiva la Taq polimerasi e stabilizza le interazioni tra oligonucleotidi e stampo;
- dNTPs che costituiscono i precursori per la sintesi del DNA;
- 2 oligonucleotidi che fiancheggiano il tratto di DNA che deve essere amplificato;

Tale miscela viene sottoposta ad un numero "n" di cicli di:

- 1) Riscaldamento a 95°C con conseguente separazione delle eliche (denaturazione).
- 2) Raffreddamento per permettere l'appaiamento delle singole eliche del DNA con gli oligonucleotidi alla temperatura specifica (annealing).
- 3) Raggiungimento di una temperatura per la sintesi del DNA ad opera della Taq polimerasi (extension).

La miscela di reazione utilizzata per la PCR è composta da: 5µl di Buffer PCR (Tris-HCl 100mM pH 8.3; 500mM KCl; 15mM MgCl₂); 5µl dNTP 2mM; 2µl Primer 50ng/µl; 1µl DNA 100ng/µl e acqua sterile per un volume finale di 50µl.

Il programma di cicli tipico invece è il seguente:

- 4 min a 94°C (denaturazione) ed aggiungere 1µl di Taq polimerasi circa a fine dei quattro minuti
- 40 sec a 94°C (denaturazione)
- 40 sec a 50°C (appaiamento) *
- 1 min a 72°C (polimerizzazione)
- 30 cicli da 2) a 4)
- 7 min a 72°C (polimerizzazione)
- La T (°C) di appaiamento dei "primers" deve essere di circa 6-10°C inferiore al valore di T_m dei primer utilizzati.
- I tempi di polimerizzazione dipendono dalla lunghezza del frammento da amplificare ed in media sono di 1min per 1kb

Per ogni reazione sono stati preparati corrispondenti controlli negativi (eliminando il DNA stampo) e positivi (usando un DNA stampo o degli oligonucleotidi già usati in precedenza).

Solitamente 1/10 dei prodotti di reazione vengono sottoposti a elettroforesi su gel di agarosio all'1% con EtBr insieme ad un marcatore di lunghezza per stimare la dimensione del frammento amplificato e valutare la purezza, oltre che per valutare la specificità delle condizioni di reazione o eventuali contaminazioni dei reagenti.

Purificazione dei frammenti amplificati per PCR mediante colonne NUCLEOSPIN Extract II (Machery-Naigel)

FASI: corsa elettroforetica su gel di agarosio all'1%; escissione del frammento dal gel; Gel lisi; Lavaggi della membrana; Eluizione del frammento.

Si prepara un gel di agarosio al 1%. Caricare direttamente l'amplificato con BBF 10X ed elettroforesi. Si preriscalda il multi-block a 50°C. A fine corsa, si taglia con bisturi la parte di gel contenente il frammento di interesse. Si aggiungono 300µl di NT1 del kit per 100 mg di gel tagliato e si mette a 50°C per 5-10 min. Una volta sciolto il gel, si trasferisce il tutto nell'apposito tubicino con membrana nel fondo a funzione di filtro. Si centrifuga 1 min a 8000 rpm e si elimina il liquido precipitato. Si aggiungono 600µl NT3 e si centrifugano 1 min a 11000rpm. Si elimina il liquido precipitato e si aggiungono altri 200 µl NT3. Si centrifuga 2 min a 11000rpm. Si trasferisce il tubicino in Eppendorf. Si aggiungono 20-25 µl di NE e si lascia 1 min a T ambiente. Si centrifuga 1 min a 11000rpm. Il frammento di DNA è stato eluito dalla membrana nell'Eppendorf.

Si conserva a -20°C.

Protocollo per la preparazione di ceppi competenti di *E.coli* mediante soluzioni di CaCl₂

Un ceppo di *E.Coli* batterico non naturalmente competente è un ceppo la cui parete e relativa membrana plasmatica è stata resa permeabile al DNA plasmidico, mediante un trattamento con CaCl₂. Per preparare un ceppo *E.Coli* che presenta tali caratteristiche, si inocula in 5ml di LB il ceppo di *E.coli* che deve essere reso competente e lo si cresce a 37°C per 16 ore in agitazione. L'inoculo viene rinfrescato effettuando una diluizione di 1:100 in LB, e cresciuto a 37°C fino ad un valore di assorbanza a 620 nm compreso tra 0.3 e 0.6 O.D. Le cellule vengono centrifugate per 5 min a 5000 g, risospese in metà del volume di CaCl₂ (50 mM) iniziale e poste in ghiaccio per 30 min, trascorsi i quali, vengono nuovamente centrifugate per 5 min a 5000 g. Il precipitato ottenuto viene risospeso in 1/10 del volume di CaCl₂ 50 mM e posto in ghiaccio da un'ora a tutta la notte. Infine, si aggiunge alla soluzione il glicerolo (30%) e si trasferisce in provette Eppendorf da 1.5ml (aliquote da 100, 200 e 400 µl). Le aliquote di cellule vengono conservate a -80°C.

Protocollo per il clonaggio in vettore plasmidico pMOSBlue Blunt Ended Cloning Kit (Amersham biosciences)

Tutti i passaggi sono in accordo con quanto suggerito nel protocollo di accompagnamento del Kit.

Per prima cosa bisogna calcolare la quantità di inserto richiesto con l'equazione: $Z \times 50 / 2887 \times 2.5/1 = X$ ng di inserto, in cui Z è la lunghezza in bp dell'inserto, 2887 è la lunghezza in bp del vettore, 50 sono i ng di vettore utilizzati e 2.5/1 il rapporto molare inserto/vettore utilizzato (può andare da 1/1 a 2.5/1). Conoscendo la concentrazione dell'inserto si ricava il volume in µl da aggiungere.

Il primo passaggio prevede la fosforilazione del frammento

H ₂ O	Y µl
inserto	X µl
PK buffer 10X	1 µl
DTT	0.5 µl
PK enzima mix	1 µl

Si centrifuga brevemente a 13200 rpm per raggruppare i volumi. La reazione viene incubata a 22°C per 40 min. Si blocca la reazione incubando a 75°C per 10 min. Si tiene in ghiaccio per 2 min e si spinnia a 13200 rpm per raggruppare i volumi. Si esegue la reazione di ligasi aggiungendo: 1 µl di vettore pMosBlue; 1 µl di DNA ligasi T4 per un volume totale di 12 µl. Si incuba a 22°C per 2h.

Si procede alla trasformazione di cellule DH5α

Informazioni sul vettore plasmidico pMosBlue:

- Linearizzazione al sito di taglio EcoR V
- Lunghezza = 2887 bp
- Gene per la resistenza all'antibiotico ampicillina
- Promotore dell'Rna Polimerasi T7
- Polilinker (a monte e a valle del sito EcoRV)
- Efficienza di trasformazione circa del 75%.

La quantità di inserto utilizzata può essere in eccesso se la probabilità di avere multimeri di inserto è bassa (inserto molto lungo).

Trasformazione batterica

Con questo procedimento è possibile trasformare cellule batteriche con DNA plasmidico.

- Si aggiungono da 5-20ng di plasmide o tutta la reazione di ligasi ad un'aliquota di 100µl di cellule competenti (solitamente DH5α), tenute in ghiaccio
- Si tiene in ghiaccio per almeno 15 min
- Si sottopongono le cellule ad "heat shock" a 37°C per 10 min e poi subito in ghiaccio.

Si piastrano tutte le cellule con una bacchetta di vetro in condizioni di sterilità su terreno solido selettivo (LB agar + ampicillina). Si incubano le piastre capovolte a 37°C overnight. L'antibiotico fa sì che crescano solo le cellule che hanno assunto il plasmide poiché esso contiene il gene che conferisce resistenza.

Estrazione di DNA plasmidico su piccola scala mediante lisi alcalina (“mini-prep”)

Tale metodo permette di estrarre 2-20µg di DNA plasmidico da cellule trasformate per successivo sequenziamento o digestione diagnostica.

- Da una piastra di E.coli recante il plasmide di interesse, viene effettuato in maniera sterile un inoculo di una singola colonia che viene posta in un tubo batteriologico da 10-15ml contenente 3ml di brodo di coltura LB con ampicillina (100µg/ml finali).
- Il tubo è incubato per 12-16 ore a 37°C in agitazione, affinché la coltura batterica raggiunga la fase di crescita stazionaria.
- Si versa in tubo Eppendorf 1.5 ml di coltura. [Riporre il resto degli inoculi a 4°C].
- Si centrifuga per 1 min a 13200 rpm per ottenere un “pellet” di cellule batteriche.
- Si risospende il pellet in 200 µl di Sol.A (al vortex). Si aggiungono 200 µl di Sol.B, si agita bene per inversione. Si aggiungono 200 µl di Sol.C, si agita per inversione.
- Si centrifuga 20 min a 13200 rpm a 4°C. Si trasferisce il surnatante (600 µl) in Eppendorf pulite e si aggiunge 1 ml di EtOH 95%. Si agita per inversione.
- Si lascia a -20°C per almeno 30 min.

Dopo la precipitazione, si centrifuga per 15 min a 13200 rpm a 4°C. Si elimina il surnatante ed si sciaccia con 200 µl di EtOH 70%. Si centrifuga per 5 min a 13200 rpm. Si elimina il surnatante ed si lascia asciugare bene il pellet.

Si risospende in 30 µl di TE e 2 µl di RNasi A (10mg/ml). Si incuba 20 min a 37°C.

Conservare le miniprep. a -20°C (fino al sequenziamento del clone positivo).

Estrazione di DNA plasmidico su media scala mediante metodo LiCl (“midi-prep”)

Con questa tecnica è possibile estrarre da 40-70µg di DNA plasmidico altamente purificato a partire da 50ml di sospensione batterica satura.

- Da una piastra di E.coli recante il plasmide di interesse o direttamente dalla “mini-prep” positiva viene effettuato in maniera sterile un inoculo in 50ml di brodo di coltura LB con ampicillina (100µg/ml finali).

Il tubo è incubato per 12-16 ore a 37°C in agitazione, affinché la coltura batterica raggiunga la fase di crescita stazionaria.

I 50 ml di terreno LB, con le cellule batteriche vengono versate in apposite falcon da centrifuga. Si centrifuga per 6 min a 5000 rpm. Si elimina il surnatante e si risospende il pellet con 1 ml di GTE lasciando per 5 min a T ambiente. Si aggiungono 2 ml di Sol.B si agita e si lascia 10 min in ghiaccio. Si aggiunge 1.5 ml di Sol III, si agita e si lascia 5 min in ghiaccio. Si centrifuga per 15 min a 10000 rpm e si recupera il surnatante, utilizzando una garza per evitare di versare anche una piccola quota di pellet costituito esclusivamente di proteine denaturate. Si aggiungono 0.6 volumi di isopropanolo, si agita e si lascia in ghiaccio per almeno 20 min. Si centrifuga per 20 min a 10000 rpm. Si versa il surnatante e si lascia asciugare. Si risospende in 1 ml di TE 1X e si aggiunge ugual volume di LiCl 5M. Tenere 20 min in ghiaccio. Si centrifuga per 20 min a 10000 rpm e si trasferisce il surnatante in una nuova falcon. Si precipita di nuovo con 2.5 volumi di etanolo 95% e si lascia a -20°C per almeno 30 min. Si centrifuga per 15 min a 10000 rpm. Si elimina il surnatante e si risospende in 500µl di TE e 20µl di RNasi A (10mg/µl). Si incuba per 20 min a 37°C. Si aggiunge ugual volume di Polietilenglicol (13%PEG8000-1.6M NaCl) e si tiene in ghiaccio per almeno 1 ora. Si centrifuga per 30 min a 13200 rpm a 4°C e si elimina il surnatante. Si risospende il pellet in 200µl di TE 1X e si lascia a T ambiente per qualche minuto. Si aggiunge ugual volume di Fenolo, si agita e si centrifuga per 15 min a 13200 rpm. Si preleva la fase acquosa e si aggiungono 200µl di Fenolo/Cloroformio e si ripete la centrifugazione. Si preleva la fase acquosa e si aggiunge 200 µl di Cloroformio/Alcool-isoammilico, si centrifuga. Si preleva la fase acquosa e si aggiungono 20µl di Sodio Acetato 3M e si precipita con 2.5 volumi di Etanolo 95%. Si lascia a -20°C overnight. Si centrifuga per 15 min a 13200 rpm, si aspira il surnatante e si aggiungono 200µl di Etanolo 70%. Si centrifuga per 5 min a 13200 rpm. Si

elimina il surnatante e si asciuga bene il pellet di DNA. Si risospende in una adeguata quantità di TE 1X. Si conserva a 4°C.

Elettroforesi su gel di Agarosio

Si è fatto ricorso alla corsa elettroforetica su gel di agarosio all'1-2,5%(peso/volume) in tutti i casi in cui è stato necessario verificare la completezza di una digestione, valutare la purezza di un DNA estratto, di stimare la concentrazione o la lunghezza in paia di basi del DNA nelle varie preparazioni

Stima della concentrazione di DNA in soluzione

La stima della concentrazione dei campioni di DNA è stata effettuata mediante due metodi:

- *Elettroforesi su gel di agarosio:*

si corre il campione in un gel di agarosio all'1% in presenza di Bromuro di etidio e di adeguati marker di quantità a concentrazione nota: comparando la fluorescenza agli UV della banda di interesse con quella dei marker di quantità si può stimare approssimativamente la concentrazione del campione iniziale.

- *Spettrofotometria UV:*

Si tara lo spettrofotometro effettuando una misurazione con 1ml di acqua mQ.

Si diluisce 1µl del campione in 1ml di acqua mQ in una cuvetta di quarzo e si misura l'assorbanza alla lunghezza d'onda di 260nm: lo strumento restituisce direttamente il valore di concentrazione del campione e un'indicazione della purezza in base al rapporto OD260/OD280 (se maggiore o uguale a 1,8 il campione si ritiene privo di contaminanti quali sali e proteine).

Digestione con enzimi di restrizione

L'uso di enzimi di restrizione è stato necessario per effettuare digestioni diagnostiche, rimuovere sequenze non desiderate, verificare la presenza e l'orientamento di un inserto all'interno di un vettore, linearizzare plasmidi per un clonaggio o per usarli come stampo lineare in esperimenti di trascrizione in vitro o per la preparazione di sonde marcate a singolo filamento. Per quest'ultima applicazione è stata fatta particolare attenzione ad usare enzimi che lasciassero un'estremità 3' "blunt".

Le reazioni di digestione sono state assemblate di norma con 3unità/µg di DNA da digerire di enzima di restrizione e con il suo tampone specifico, facendo attenzione che il volume di enzima non eccedesse il 10% del volume finale di reazione: in questo modo il glicerolo presente negli stock di enzimi non interferisce con l'attività enzimatica. L'incubazione è stata protratta da 2h a ON a seconda dell'enzima e della quantità di DNA da digerire, a una temperatura scelta in base alle specifiche del singolo enzima.

La completezza della reazione è stata valutata sottoponendo un'aliquota della miscela di reazione a elettroforesi su gel di agarosio all'1% colorato con Bromuro d'etidio.

Reazione di Ligasi

La reazione di ligasi consente di saldare le estremità di un plasmide a quelle di un frammento lineare di DNA grazie all'attività dell'enzima DNA ligasi, rendendo così possibile il clonaggio dei frammenti di DNA all'interno di vettori plasmidici.

Ogni reazione è stata allestita come segue:

- T4 DNA ligasi (usb) 1 unità
- T4 DNA Ligasi Buffer 10X 1,5µl
- Plasmide linearizzato 150ng
- Quantità variabili di inserto
- H2O mQ a 15µl

Incubare 1h a ambiente

Per ciascun clonaggio è stato valutato di volta in volta il rapporto vettore/inserto in grado di spostare efficientemente la reazione verso la saldatura tra DNA lineare e DNA plasmidico; una miscela di reazione priva dell'inserto è stata inoltre utilizzata come controllo per valutare la capacità del vettore di richiudersi su sè stesso.

Defosforilazione con CIP

(Phosphatase alcalina (AP) from Calf Intestine (CIP))

Quando non è stato possibile effettuare clonaggi direzionali (vettore digerito con due differenti enzimi di restrizione per favorire la reazione di saldatura con l'inserito oltre che il suo inserimento in un verso noto), per impedire la risaldata spontanea del vettore si è reso necessario defosforilare le estremità mediante CIP (Roche) con il seguente protocollo;

Dopo aver digerito il plasmide in 10µl con gli enzimi d'interesse e averne controllato l'integrità su gel di agarosio colorato con Bromuro di Etidio, si porta il volume della reazione a 20µl aggiungendo: 2µl di CIP Buffer 10X e H₂O a volume.

Si aggiunge 1µl di CIP (10 unità)

Si incuba 30 min a 37°C.

Il DNA è poi purificato ed utilizzato nelle reazioni di ligasi.

Protocollo per la reazione di sequenza partendo da DNA plasmidico ricombinante

Si utilizza come fonte di DNA stampo il plasmide pMOSBlue, utilizzato per il clonaggio dei frammenti di snoRNA e di alcune sequenze del gene *mfl*. Ai lati del sito di clonaggio (EcoRV), questo plasmide ingegnerizzato contiene le sequenze per i due primer, T7 e U19, che consentono la lettura di entrambi i filamenti delle sequenze di interesse. Si prepara la soluzione di denaturazione, composta da 4% idrossido di sodio 10N e acqua distillata. Si aggiungono 20µl di Sol.denaturante a 20 µl di DNA (5 µg di DNA) e si incuba a temperatura ambiente per 10 min. Si precipita in seguito con 7 µl di sodio acetato 3M e 120µl di etanolo 95% e si incuba a -80°C per almeno 1 ora. Si centrifuga a 13200 rpm per 15 min a 4°C e si lascia asciugare bene il pellet. In seguito, si risospende, in 2µl di primer, 2µl di buffer di annealing e 10µl di acqua distillata. Questa mix1 viene incubata prima a 65°C per 5 min poi a 37°C per 10 min ed infine a temperatura ambiente per almeno 5 min. Si diluisce la RNA polimerasi T7 modificata per le sequenze (sequenasi versione 2.0) nell'apposito buffer. Alla mix1 si aggiungono 3µl di labelling mix, 0,5µl di dATP α -³⁵S e 2µl di T7 polimerasi. Si incuba questa mix2 a temperatura ambiente per 5 min e poi si aggiungono 4,5µl della mix2 in quattro tubi Eppendorf contenenti ciascuna le mix di ddNTP/dNTP 1/100 T C G A e si incuba a 37°C per 15-20 min. Infine si aggiungono 5µl di soluzione STOP, composta di formammide e colorante per la corsa, e si conserva il tutto a -20°C per un breve tempo.

Tutte le reazioni richieste sono contenute nel kit della USB.

Elettroforesi su gel denaturante di poliacrilammide per la determinazione dei prodotti finali delle reazioni di sequenza

Si mettono in agitazione 50ml di soluzione contenente 6% acrilammide/bis-acrilammide (40% w/v), 8M urea, TBE 1X e acqua. Si versano contemporaneamente 250µl di persolfato di ammonio 10% (APS) e 25µl di TEMED. Si lascia polimerizzare per circa 1 min, quindi si versa con attenzione la soluzione tra le lastre precedentemente pulite con alcool ed acetone per eliminare granelli di polvere e evitare la formazione di bolle. In seguito all'estremità del gel si pone un pettine per creare il fronte di corsa e si applicano delle pinze per diminuire lo spessore del gel ed assicurare le due lastre tra di loro. Si lascia il tutto overnight. La mattina seguente si procede al caricamento delle sequenze preparate sul gel di poliacrilammide. Per prima cosa si bagna con un po' di tampone di corsa, TBE, il pettine per sciogliere l'acrilammide superflua, poi si toglie delicatamente il pettine e lo si posiziona al contrario per creare i pozzetti di caricamento. Dopo questa operazione si pone il gel tra le lastre nel macchinario apposito per la corsa e si versa il tampone di corsa ai due elettrodi. Si fa una precorsa per riscaldare il gel per almeno 30 min, dopodichè si caricano le sequenze campione, precedentemente denaturate a 90°C per 5 min e subito messe in ghiaccio. La durata della corsa è indicativa della lunghezza della sequenza campione da leggere. Generalmente nel nostro laboratorio, per la lettura delle sequenze degli snoRNA clonati in frammenti lunghi 400-500 bp, si adoperano due combinazioni di corsa: una corsa di circa 6-7 ore a 1800V, che consente la lettura delle basi poste 300-500 nucleotidi a valle dell'inizio di trascrizione, e una corsa di circa 3-4 ore, nelle stesse condizioni, che invece permette di leggere quelle basi poste 10-300 nucleotidi a valle.

Trasferimento del gel su carta, esposizione delle lastre per autoradiografia e sviluppo

Terminata la corsa dei campioni, il gel viene trasferito su un foglio di carta Whatmann 3MM tramite semplice compressione della carta sul gel, il quale aderisce ad essa. In seguito, il tutto

si lascia asciugare per circa 30 min a 80°C nel gel-dryer. Finita questa operazione, si controlla prima la bontà delle incorporazioni del radioattivo nelle reazioni di sequenza, tramite la misura delle conte per minuto (cpm) con il geiger, dopodiché il gel è esposto ad una lastra autoradiografica e lasciato nell'apposita cassetta da esposizione per tutto il tempo necessario (1-3 giorni). Lo sviluppo delle lastre autoradiografiche consiste nell'immergere la lastra esposta, dapprima in una soluzione di sviluppo, che consente la visualizzazione delle bande tipiche di sequenza, quindi nella soluzione di fissaggio, dove le bande sono definitivamente fissate alla lastra.

Nota

L'emulsione delle lastre autoradiografiche è una sospensione di cristalli di argento alide in gelatina. L'esposizione alle radiazioni attiva i cristalli, convertendo un numero di ioni argento in argento metallico che precipitano dal lattice cristallino e producendo una immagine latente che può essere convertita in immagine reale dopo lo sviluppo. La risultante immagine latente però è instabile e tende a perdere gli elettroni catturati ed a ritornare nella matrice gelatinosa. A temperatura ambiente questa reazione di ritorno ha un tempo inferiore al secondo, mentre a -70°C la reazione è molto più lenta.

L'isotopo ³⁵S ha una scarsa energia di emissione, circa 0,17 MeV, e può penetrare solo a 0,25mm di profondità, quindi è essenziale che la lastra e la fonte di radiazioni siano in stretto contatto l'un l'altro.

Protocollo per l'estrazione di RNA totale da adulti ed embrioni di *D. melanogaster*

Si raccoglie un numero adatto di individui in tubi Eppendorf da 1.5 ml sterili, in cui vengono pestate, tramite un apposito pestello, in 300µl di buffer di omogeneizzazione. Si aggiungono altri 200µl di buffer (vol.totale 500µl) e 100µg/µl di proteinasi K e si lascia incubare a 37°C per 1 ora. A fine incubazione, si aggiungono 500µl di fenolo/0.1%idrossichinolina e si centrifuga per 5 min a 13200 rpm a 4°C. Si preleva la fase superiore acquosa e si trasferisce in una nuova Eppendorf sterile. Si ripete l'aggiunta di fenolo e si centrifuga per 5 min a 13200 rpm a 4°C. Si preleva la fase superiore acquosa trasferendola in una nuova Eppendorf sterile. Si aggiungono 500 µl di cloroformio/alcool iso-amilico (24:1) e si centrifuga per 5 min a 13200 rpm a 4°C. Si preleva la fase superiore acquosa e la si trasferisce in un tubo Eppendorf sterile tenuta in ghiaccio (da questo punto lavorare sempre in ghiaccio). Si aggiunge sodio acetato a concentrazione finale di 200mM. Si precipita con due volumi di etanolo assoluto (95%) e si pone il tutto a -80°C over night. Si centrifuga per 15 min a 13200 rpm, si estrae l'etanolo 95% e si aggiunge 200µl di etanolo 70% in H₂O DEPC. Si centrifuga per 5 min a 13200 rpm e si estrae l'etanolo 70%, lasciando asciugare il campione. Si risospende in 30µl di H₂O DEPC e si conserva il campione di RNA a -80°C.

Protocollo per l'estrazione di RNA totale da larve di *D. melanogaster* (di 1°,2° e 3° stadio)

Si raccoglie un numero adatto di larve in tubi Eppendorf da 1.5ml sterile. Si aggiungono 300µl di buffer di omogeneizzazione per RNA di larve e si pesta velocemente il tutto. Si aggiungono altri 200µl di buffer (vol.totale 500µl) e 250µg/ml di proteinasi K e si incuba a 50°C per 1 ora. A fine incubazione si aggiungono 250µl di fenolo/idrossichinolina e si mettono in agitazione per 5 min. Poi altri 250µl di fenolo/idros. e si agita per 5 min.

Si centrifuga per 5 min a 13200 a 4°C. Si preleva la fase acquosa superiore e si trasferisce in un tubo Eppendorf sterile tenuta in ghiaccio (da questo punto lavorare sempre in ghiaccio). Si aggiungono 500µl di cloroformio/alcool iso-amilico ed agitare per 5 min.

Si centrifuga per 5 min a 13200 rpm a 4°C. Si preleva la fase acquosa superiore e si trasferisce in un nuovo tubo Eppendorf sterile tenuto in ghiaccio. Si precipita con due volumi di etanolo 95% e si mette il tutto a -80°C over night. Si centrifuga per 15 min a 13200 rpm e si estrae l'etanolo 95% e si aggiunge 200µl di etanolo 70% in H₂O DEPC.

Si centrifuga per 5 min a 13200 rpm e si estrae l'etanolo 70% e si lascia asciugare il campione. Si risospende in 30µl di H₂O DEPC e si conserva il campione di RNA a -80°C.

Estrazione dell'RNA totale con il TriReagent

Tutti i passaggi sono in accordo con quanto suggerito nel protocollo di accompagnamento del Kit.

Si mette il campione in tubo Eppendorf sterile da 1.5 ml e si pesa. Si aggiunge il Tri Reagent (1ml di Tri Reagent per 50-100 mg di tessuto) e si pesta il campione per omogeneizzarlo. Si centrifuga a 14000 rpm per 10 minuti a 4⁰C.

Si trasferisce il surnatante in un nuovo tubo Eppendorf sterile. Si lascia il campione per 5 minuti a temperatura ambiente. Si aggiunge cloroformio (0.2ml per ml di Tri Reagent usato). Si agita per inversione e si tiene 5 minuti a temperatura ambiente.

Si centrifuga a 14.000 rpm per 15 minuti a 4⁰C, si separano così le tre fasi contenenti, dalla fase acquosa più leggera a quella organica più pesante, RNA, DNA e proteine, rispettivamente. Si trasferisce la fase acquosa, contenente l'RNA, in una nuova provetta eppendorf sterile e si aggiunge isopropanolo (0.5ml per ml di Tri Reagent usato). Si agita per inversione e si lascia a temperatura ambiente per 5 minuti. Si centrifuga per 10 minuti a 14.000 rpm a 4⁰C (per far precipitare l'RNA).

Si rimuove il supernatante e si aggiunge al pellet (RNA) etanolo 70% (1ml di etanolo per ml di Tri Reagent usato). Si centrifuga per 5 minuti a 4⁰C a 14.000rpm.

Si rimuove il supernatante e si lascia asciugare il pellet (RNA). Si risospende l'RNA in H₂O DEPC sterile.

Protocollo per la purificazione degli mRNA (polyA⁺) mediante "Quick PREP MICROMRNA PURIFICATION KIT" Amersham Pharmacia Biotech Inc

Tutti i passaggi sono in accordo con quanto suggerito nel protocollo di accompagnamento del Kit, ad eccezione dei primi passaggi a cui è stata apportata qualche modifica.

L'RNA estratto con uno dei metodi sopra citati è stato risospeso in H₂O-DEPC e la concentrazione è stata aggiustata in modo da avere non più di 100 µg di RNA in 100 µl.

A questo volume si aggiungono 400µl (per campione) di Buffer Extraction tenuto preventivamente a 37⁰C e 800µl (per campione) di Buffer Elution.

Contemporaneamente In un tubo Eppendorf da 1.5 ml è stata preparata per ogni campione 1 ml di oligo(dT)cellulosa .

Si centrifuga per un minuto sia la cellulosa oligo(dT) che il campione. Si aspira il buffer dal tubo contenente la cellulosa oligo (dT) e su questa si trasferisce l'RNA. Si mette il campione sulla ruota per almeno 15'-20' per mescolare bene la cellulosa con l'RNA. Si centrifuga per 10 secondi alla massima velocità.

Si lava per cinque volte con 1 ml di Buffer High Salt (mix, centrifuga per 10 sec ed aspira il supernatante ogni volta). Si lava per due o tre volte con 1ml di Buffer Low Salt (mix, centrifuga per 10 sec ed aspira il supernatante ogni volta).

Si risospende, pipettando, con 0.3 ml di Buffer Low Salt il pellet di cellulosa oligo (dT) con l'RNA legato. Si trasferisce la risospensione in una colonnina Micro Spin e si lava per tre volte la colonnina con 0.5 ml di Buffer Low Salt. Si eluisce per due volte l'mRNA polyA⁺ con 0.2 ml di Buffer Elution preriscaldato a 65⁰C dopo aver messo la colonnina Micro Spin su un tubo Eppendorf sterile da 1.5 ml. Si uniscono le due eluizioni dell'RNA polyA⁺.

Precipitazione dell'RNA polyA⁺ eluito: si aggiungono al campione 10 µl di soluzione glicogeno fornito dal Kit. Si aggiungono 40 µl di potassio acetato (1/10 del volume totale). Si aggiunge 1ml di etanolo 95%. Si lascia il campione a -80⁰C per almeno 30 minuti per far precipitare l'mRNA. Si centrifuga per 15 minuti alla temperatura di 4⁰C e si elimina il supernatante. Si fa asciugare il pellet e si risospende in H₂O DEPC (circa 20 µl). Si controlla la purificazione mediante elettroforesi su gel di agarosio all'1% e poi si quantizza allo spettrofotometro.

Protocollo per la RetroTrascrizione (RT) degli mRNA (sintesi del cDNA).

Il kit Super Script First-Strand Synthesis System per RT-PCR (Invitrogen) è ottimizzato per sintetizzare il primo filamento di cDNA partendo da piccole quantità di RNApolyA⁺ o da RNA totale. Questo Kit è stato utilizzato anche nei primi passaggi di 5' e 3'RACE.

Si prepara la seguente

miscela RNA/primer in tubi Eppendorf sterili:

componenti	campione	Controllo No RT	RNA kit (+)	No mRNA
mRNA	x µl	x µl	–	–
RNA kit (50ng/0.001ml)	–	–	1 µl	–
10mMdNTPmix	1 µl	1 µl	1 µl	1 µl
oligo(dT) ₁₂₋₁₅	1 µl	1 µl	1 µl	1 µl
oligo(dT) ₁₂₋₁₅ Adaptor oligo specifico (05µg/µl)				
H ₂ O DEPC	x µl	x µl	7 µl	8 µl
Volume totale	10 µl	10 µl	10 µl	10 µl

Si incubano i campioni (miscela RNA/primer) a 65°C per 5 minuti e si mettono in ghiaccio per almeno 1 minuto. Si prepara la seguente mix:

componenti	N ^o reazioni x
10RT Buffer	2 µl
25mM MgCl ₂	4 µl
0.1M DTT	1 µl

Si aggiunge RNasi OUT Recombinant RNasi I Inibitor: 1 µl x n^o di reazioni. Si aggiungono 9 µl di mix reazione ad ogni mix RNA/primer e si lascia incubare a 42°C per 2 minuti. Si aggiunge a tutti i campioni, tranne ai controlli no RT, 1 µl di RT (SSII : Super Script II) e si incubano a 42°C per 50 minuti. Si mettono a 70°C per 15 minuti per interrompere la reazione di retrotrascrizione. Dopo breve centrifugazione si aggiunge 1 µl di RNasi H ad ogni campione e si lascia incubare 20 minuti a 37°C. La reazione di RT viene amplificata mediante PCR (vedi protocollo PCR), secondo il seguente schema tipo:

- 5 min a 95°C (denaturazione), si aggiunge 1µl di Taq polimerasi alla fine dei cinque minuti
- 55°C per 5 minuti (appaiamento)
- 72°C per 40 minuti (polimerizzazione)
- 95°C per 2 minuti (denaturazione)
- 40 sec a 95°C (denaturazione)
- 1 min a 55°C (appaiamento) *
- 3 min a 72°C (polimerizzazione)
- 40 cicli da 5) a 7)
- 7 min a 72°C (polimerizzazione)

3' RACE: La retro trascrizione è stata condotta utilizzando un oligodT caratterizzato all'estremità 5' dalla presenza di una sequenza Adaptor (5'-GACTCGAGTCGACATCGA(T)17-3'). I successivi passaggi di amplificazione sono stati condotti con un set di primer derivati dalla sequenza del gene mfl sempre in associazione con il primer Adaptor(5'-GACTCGAGTCGACATCGA-3'). I cicli e le temperature della PCR sono stati adattati in funzione delle esigenze.

5'RACE:

Questo metodo è stato utilizzato per amplificare il 5' del trascritto da 1.0kb che deriva dal gene mfl con un evento di splicing alternativo.

La reazione di RT è stata condotta utilizzando un'oligo posizionato tra l'esone 3 e l'esone 9(5'-CCTAATCAACAAATCCATATTTTCGGG-3') e la reazione è stata eseguita secondo le indicazioni riportate sopra.

Il cDNA così ottenuto viene concentrato a 10 μ l

Aggiunta di una coda di polyA al 3' del cDNA mediante Terminal Transferase:

ai 10 μ l di cDNA aggiungere

4 μ l di terminal transferase buffer 5x

4 μ l dATP; 1mM

2 μ l terminal transferase (15-25 unità)

Incubare la reazione a 37°C per 15 min. Disattivare l'enzima con un trattamento al calore (80°C per 3min). Diluire il cDNA ad 1ml ed utilizzarlo nei successivi passaggi di amplificazione. Il primo ciclo di PCR è condotto con un oligo posizionato nell'esone 3 di mfl (5'-GGGCACCACGCAGCTTCTC-3') e lo stesso oligo (T)17-Adaptor utilizzato nella reazione di 3'RACE. Il secondo ciclo di PCR è stato condotto con un oligo più interno all'esone 3 (5'-GGGACTTCACCAGACGGG-3') ed il primer Adaptor descritto precedentemente.

Tutti i prodotti ottenuti nelle varie reazioni di amplificazioni sono stati recuperati da gel e clonati nel vettore pMOSBlue.

Analisi dell'RNA mediante elettroforesi su gel d'agarosio denaturante e trasferimento su filtro di nylon (Northern blot)

Si prepara un gel di agarosio all'1,2% ad esempio per un gel del volume di 150ml si sciolgono 1,8gr. di agarosio in 110 ml di acqua distillata a cui si aggiungono 15 ml di RB 10X (Running Buffer 10X). Si lascia bollire e quando la temperatura del gel scende a circa 65°C, si aggiungono, sotto cappa, come denaturante 24,9 ml di formaldeide e si versa nel lettino di corsa opportuno. Si lascia solidificare il gel. Si caricano almeno 8 μ g di RNA (in un volume che non deve superare i 6 μ l) in 19 μ l di soluzione di caricamento, composta da formammide 50%, formaldeide 2.2M, RB 1X e 0,5% bromuro d'etidio in acqua DEPC. Si caricano anche 4 μ l di marker (composto da frammenti di RNA a peso molecolare noto) più 2 μ l di H₂O DEPC in 19 μ l di soluzione di caricamento. Si denaturano i campioni per 15 min a 65°C, e si aggiungono 2,5 μ l di una soluzione di Ficoll 25%, 0,05% Bromofenolo, 0,05% Xilene Cianolo (BBF). Si caricano i campioni sul gel, precedentemente preparato, e si effettua la corsa elettroforetica per circa 4-5 ore a 70 V nel tampone di corsa RB 1X. A fine corsa, il gel viene adagiato su un supporto di vetro che poggia su una vaschetta contenente 300-400 ml di SSC 10X. La soluzione è in contatto con il gel mediante un ponte formato da due fogli imbevuti di carta Whatman 3MM. Un filtro di nylon delle stesse dimensioni del gel viene adagiato sul gel evitando di formare bolle d'aria. Successivamente, si pongono due fogli di carta Whatmann 3MM, imbevuti della soluzione di SSC 10X, e su di essi una pila di circa 10 cm di comune carta da banco e con alla sommità un peso adatto. Il tutto lo si lascia overnight per il trasferimento. Al mattino seguente si rimuove delicatamente il filtro di nylon, lo si fa asciugare a temperatura ambiente e infine si pone 1 ora in stufa a 80°C sotto vuoto per fissare l'RNA al filtro.

Protocollo di preibridazione ed ibridazione dell' RNA trasferito sul filtro di nylon

Il filtro viene incubato in una busta contenente 20ml di Soluzione di preibridazione/ibridazione 1X composta da soluzione fosfato 0,5M, SDS 5% e EDTA 10mM precedentemente riscaldata a 55°C. Il filtro viene incubato per 15 min a 55°C. Successivamente si aggiunge la sonda di DNA marcata con α -³²PdCTP e precedentemente denaturata per 20 min a 100°C alla soluzione di preibridazione/ibridazione. Il filtro viene incubato in questa soluzione a 55°C over night.

La rimozione del radioattivo, ibridato specificamente sul filtro, viene eseguita mediante lavaggi successivi con la soluzione di ibridazione 0,1X preriscaldata a 55°C. Dopo 3 lavaggi successivi di circa 15-20 min, il filtro viene asciugato a temperatura ambiente ed è pronto per essere esposto ad una lastra autoradiografica a -80°C per un tempo variabile di esposizione.

Protocollo per la reazione di marcatura radioattiva della sonda DNA mediante Nick translation Kit (Roche)

I frammenti di DNA clonati in pMosBlue sono stati digeriti, recuperati da gel e quantizzati sono stati marcati mediante reazione di nick translation. Per ogni marcatura sono stati utilizzati 500 ng di frammento purificato. Ai 500 ng di frammento si aggiungono 1 μ l ciascuno di nucleotidi 0,5mM (dATP, dGTP e dTTP), 2 μ l di Buffer 10X, 5 μ l di α -³²PdCTP (3000Ci/mmol) e 2 μ l di mix di enzimi, composto dalla DNA polimerasi I e dalla DNAsi I. Il tutto si incuba a 37°C per circa 1 ora. Per allontanare i nucleotidi non incorporati in eccesso si

prepara una colonnina di Sephadex G50. Si introduce la miscela dei prodotti di reazione marcati radioattivamente nella colonnina e si centrifuga a 3.2 rpm per 1 min. L'eluato si denatura a 100°C per 20 min. La sonda preparata è così pronta per l'ibridazione.

Primer Extension

Questa tecnica è utilizzata per mappare l'estremità 5' dell'RNA. L'RNA da testare è ibridato con un eccesso di DNA primer a singolo filamento marcato al 5'. Questo primer fornisce l'innescò alla trascrittasi inversa che lo estende producendo il cDNA complementare all'RNA stampo.

La sequenza dei primer utilizzati è:

oligo Pex DmSnR60a (5'- TTGCGTTATTTTCAGTAAGACCAATCTCG-3')

oligo Pex DmSnR60c (5'-GAAATTCTTTTCAGTAAGACCAATC-3')

oligo Pex snm60a (5'-GGCATCGTTGTTATTGTATTGGCCTAAATGC-3')

oligo Pex snm60b (5'-CACTTAGAACCACATCTGGTTAAAACACAC-3')

Si pone 100ng di DNA primer con 50µg di RNA totale. Si aggiunge 0.1 volumi di Sodio Acetato (pH 5.2) e 2.5 volumi di etanolo e si pone 30 min a -20°C.

Si centrifuga a 12.000g per 10min a 4°C.

Il pellet è sciacquato con etanolo 70% e lasciato asciugare.

Si risospende il pellet in 30µl di buffer di ibridazione.

“Buffer di ibridazione”

- 40mM di pipes (piperazine-N,N'-bis(2-etanesulfonico acido) pH6.4
- 1mM EDTA pH8.0
- 0.4M NaCl
- 80% formamide

Dopo aver risospeso bene si incuba la soluzione a 85°C per 10 minuti per far denaturare gli acidi nucleici.

Si trasferisce a 30°C per l'annealing e si lascia 12 ore.

Si addiziona 170µl di acqua e 400µl di etanolo, si lascia 1 ora a -20°C.

Si centrifuga a 12.000g per 15min a 4°C, e il pellet viene sciacquato con 500µl di etanolo 70%.

Si lascia asciugare.

L'ibrido RNA/primer viene risospeso in 20µl di buffer reverse transcriptase.

“Buffer reverse transcriptase”

- 50mM di TRIS-HCl pH 7.6
- 60mM di KCl
- 10mM di MgCl₂
- 1mM di ciascun dNTP
- 1mM di ditiotreitolo
- 1 unità/µl inibitori di RNasi
- 50µg/ml di actinomicina D

Si addiziona 50 unità di trascrittasi inversa pipettando bene.

Si addiziona 1µl di EDTA 0.5M pH8, ed 1µl di Dnase-Rnase free 5µg/µl.

Si incuba per 30 minuti a 37°C. si addiziona 150µl di TE pH 7.6 contenente 0.1M di NaCl e 200µl di fenolo:cloroformio. Si pone al vortex per 30 secondi, e si centrifuga a 12.000g per 5 minuti a temperatura ambiente.

Si trasferisce la fase superiore acquosa in tubo pulito, a cui si aggiunge 500µl di etanolo.

Si precipita 1 ora a -20°C.

Si centrifuga 15 minuti a 4°C, si sciacqua il pellet con etanolo 70% e si lascia asciugare.

Il pellet è risospeso in 4µl di TE pH 7.4.

Si addizionano 6µl di buffer formammide e si pipetta bene

“Buffer formammide”

- 80% formammide
- 10mM EDTA pH 8.0
- 1mg/ml di xilene cianolo
- 1mg/ml di blue di bromofenolo

Si incuba 5 minuti a 95°C per denaturare e si pone immediatamente in ghiaccio. Il DNA marcato è analizzato per elettroforesi su gel di poliacrilammide 6%-urea 7M. Si carica anche la sequenza di DNA preparata con gli stessi oligonucleotidi.

Saggio di Metilazione

Questo saggio è stato condotto in accordo con quanto descritto da Maden et al., 2001 senza apportare alcuna modifica al protocollo.

Quantitative Real Time RT-PCR

La **PCR real-time**, denominata anche **PCR quantitativa** o **PCR quantitativa in tempo reale** (rtq-PCR), è un metodo che permette di amplificare (PCR) e quantificare simultaneamente molecole di RNA e/o di DNA di interesse.

Per questa analisi; 500ng di RNA polyA+, è stato retrotrascritto con 250ng di esameri random, 100U di SSII reverse transcriptase, 10mM DDT e 1X finale First-Strand Buffer, in accordo con le istruzioni fornite nel kit dell'Invitrogen. La reazione è stata condotta in un volume di 20µl e tenuta a 42°C per 1 ora.

L'analisi quantitativa è stata condotta utilizzando iQTM 5 multicolor Real-Time PCR Detection System (Bio-Rad). La reazione di PCR è stata assemblata in 15µl finali con: 1µl di cDNA, 5pmol di ciascun primer e 7.5µl di iQTM SYBR Green Supermix 2X (Bio-Rad).

Cicli e temperature:

1X, 95°C per 3min

40X, 95°C per 10s; 60°C per 30s.

Tutti i primer utilizzati in questo tipo di analisi sono stati progettati utilizzando il programma Primer Express 2.0 software (applied Biosystem, Foster City, CA) e sottoposta ad allieamento con il BLASTN per verificare che si appaessero esclusivamente al gene d'interesse.

In tutti gli esperimenti i dati sono stati normalizzati per l'espressione del gene housekeeping di *Drosophila* α Tub84B. tre differenti preparazioni di RNA sono state testate per ciascun campione e ciascuna reazione è stata avviata in triplicato. I dati riportati nei risultati sono rappresentativi di tre esperimenti indipendenti.

Produzione di un anticorpo policlonale anti-MFL/MFL α

La produzione dell'anticorpo policlonale è stata commissionata alla ditta Eurogentec, EGT grop.

Il protocollo da loro eseguito ha previsto la sintesi di due peptidi presenti nella porzione ammino terminale presente in entrambe le subforme MFL/MFL α .

peptide 1 (NH₂-CADVEVRKEKKKKKIK-CONH₂)

peptide 2 (NH₂-CHGSSPLNRDIKEYH-CONH₂)

Gli anticorpi policlonali contro i due peptidi sono stati ottenuti in coniglio.

Estrazione di proteine da cellule S2 di *Drosophila melanogaster*

Centrifugare le cellule 3' a 1400rpm

Lavare le cellule con 5ml di PBS1X, centrifugare come sopra.

Risospendi in 500µl di Buffer di lisi (RIPA buffer+inibitori proteasi+ PMSF).

Tieni in ghiaccio 30'

Centrifuga 15' a 13200 rpm.

Prelevare il surnatante, quantizzarlo allo spettrofotometro, conservarlo a -80°C

RIPA Buffer:

50mM Tris-HCl pH 7.5

150mM NaCl

5mM EDTA

1% NP40

1mM PMSF

1X cocktail inibitori proteasi

Elettroforesi su gel di poliacrilammide (SDS-PAGE)

Questa tecnica permette di separare le proteine sulla base del loro peso molecolare.

L'SDS è un detergente carico negativamente che si lega alle regioni idrofobiche delle proteine, sovrappaccandone la carica intrinseca e provocandone lo svolgimento in catene polipeptidiche

estese. Inoltre, l'aggiunta di β -mercaptoetanolo permette la riduzione dei ponti disolfuro, contribuendo a separare le proteine da eventuali associazioni dovute a tale legame: in questo modo la proteina denaturata, quando sottoposta a campo elettrico, migrerà al polo positivo in maniera esclusivamente dipendente dalle sue dimensioni.

- Il gel di poliaccrilammide viene colato tra due lastre di vetro sovrapposte e separate da uno spessore di 0,5 mm.
- Si cola per primo il running gel ad una concentrazione del 10% di acrilammide fino a circa 2/3 della superficie.
- Una volta polimerizzato si cola lo stacking gel ad una concentrazione del 5% di acrilammide.
- Terminata la polimerizzazione in ciascun pozzetto sono stati caricati 30 μ g di estratto proteico ottenuto da cellule S2 di *Drosophila* (prima del caricamento i campioni sono stati diluiti in sample buffer denaturante e riscaldati a 95°C per 5min).

Composizione Running gel: 10% Acrilammide-BisAcrilammide, 0,375M Tris-HCl pH 8.8, SDS 0,1% (per la polimerizzazione in 10ml 80 μ l di APS 10% ed 8 μ l di TEMED)

Composizione Stacking gel: 5% Acrilammide-BisAcrilammide, 0,125M Tris-HCl pH 6.8, SDS 0,1% (per la polimerizzazione in 10ml 80 μ l di APS 10% ed 8 μ l di TEMED)

Buffer di corsa: Tris Glicina 1%, 0,1% SDS

La corsa elettroforetica è stata condotta ad Amperaggio costante: 0,01A ed il tempo necessario per la separazione è stato stimato seguendo la separazione delle bande del marker.

Terminata la corsa uno dei due gel è stato colorato con Comassie per verificare l'integrità degli estratti mentre l'altro è stato sottoposto a trasferimento su filtro mediante trans-blot.

Colorazione con Comassie Brilliant Bleu: Il gel viene staccato delicatamente dal vetro di supporto ed immerso in una vaschetta contenente la "Staining Solution"

- **"Staining Solution"**
- Metanolo 30%
- Acido Acetico 10%
- Comassie Bleu R 0,1%

Tenere in agitazione per un tempo variabile dai 30min alle 2ore, procedere alla decolorazione.

"Destaining Solution"

- Metanolo 30%
- Acido Acetico 10%

Il gel decolorato può essere conservato in bustine sigillate nella stessa soluzione di Destaining. TRASFERIMENTO DEL GEL SU FILTRO.

"Buffer di Trasferimento"

- 48mM di Tris
- 39mM Glicina
- 20% Metanolo
- 0,0375% SDS

Tagliare un pezzo di filtro della dimensione del gel e 6 fogli di carta 3MM

Equilibrare il gel nel buffer di trasferimento bagnare la carta ed inumidire il fitro

Montare l'apparato di trasferimento

Avviare il trasferimento: 15 Volt costanti per 30 min.

TBS: 100mM Tris-HCl pH8, 1,5M NaCl

TTBS: 1X TBS; 0,1% Tween 20

Terminato il trasferimento colorare il filtro con il Ponceau,

Sciacquare con TTBS fino a rimozione completa del colorante

Bloccare il filtro in TTBS+10% Milk per 2h in agitazione (anche o.n.)

Incubare per almeno 2h con l'anticorpo primario diluito 1:500 in TTBS-5% milk

Sciacquare 3 volte per 15 min in TTBS

Incubare per almeno 1h con l'anticorpo secondario (anti-rabbit 1:5000 in TTBS-5% milk)

Sciacquare 3 volte per 15 min in TTBS.

Per la rivelazione è stato usato il Kit ECL-Advance Western Blotting Detection Reagents (Amersham Biosciences) in accordo con le istruzioni fornite dalla ditta produttrice.

Preparazione di sonde a singolo filamento marcate con metodo non radioattivo.

Questo protocollo utilizza la tecnica della P.C.R. (Polymerase Chain Reaction) per l'amplificazione di DNA a singolo filamento in presenza di nucleotidi marcati con Digossigenina a partire da stampi di DNA inseriti in vettori di clonaggio. Per questa procedura, i DNA plasmidici contenenti lo stampo da amplificare vengono linearizzati all'estremità opposta a quella contenente l'oligonucleotide usato come innesco. I due "primers" specifici del vettore possono essere utilizzati alternativamente per creare filamenti antisenso e senso (come controllo).

Si preparano inizialmente le seguenti soluzioni:

Miscela di reazione 10X:

- 0.5M KCl
- 0.1M Tris-HCl pH 8.3
- 15mM MgCl₂

Miscela di nucleotidi 5X in acqua

- 1mM ciascuno di dATP, dGTP e dCTP
- 0.65mM di dTTP
- 0.35mM di DIG-dUTP

Dopo aver linearizzato il plasmide, l'enzima di restrizione utilizzato viene disattivato a 70°C per 15min. Il DNA viene diluito con acqua ad un'a concentrazione di 200ng/μl ed utilizzato per preparare la seguente miscela di reazione:

- 2.5 μl 10X Miscela di reazione
- 5.0 μl Miscela di nucleotidi
- 5.0 μl Primer (30ng/ μl)
- 2.0 μl DNA stampo linearizzato
- 8.5 μl di Acqua

Si aggiungono 40 μl di olio minerale, si denatura 5min a 100°C e si aggiungono 2 μl di Taq DNA Polymerase (0.6 unità/ μl). Si incuba per 30 cicli in una "thermal cycler" per PCR secondo le seguenti condizioni: 95°C per 45sec; 55°C per 30sec; 72°C per 1min.

Terminata la reazione si aggiungono 75 μl di acqua e si rimuove la miscela dall'olio minerale.

Si precipita aggiungendo 10 μg di glicogeno come carrier, NaCl a concentrazione finale di 0.1M e 3 volumi di Etanolo 96%. Il precipitato viene risospeso in 300 μl di soluzione di ibridazione (vedi protocollo ibridazione *in situ*).

Ibridazione in situ su tessuti ed ovari

Dissezionare tessuti e ovari in PB1X

- PBS 10X: NaCl 1.3M; Na₂HPO₄ 0.07M; NaH₂HPO₄ 0.03M

Fissarli in un soluzione composta da 200 μl di Paraformaldeide 4% e 600 μl di eptano a Tamb

Rimuovere l'eptano e sciacquare più volte in PBS 1X

Trattare i tessuti per 5min con proteinasi K 50mg/ml

Sciacquare più volte in PBS 1X

Rifissare i tessuti per 15min in 4% paraformaldeide e lavare più volte con PBS 1X

Hybridization solution

- 50% formammide
- 5X SSC
- 50 μg/ml di eparina
- 0.1% Tween

Incubare a Tamb per 10min i tessuti e gli ovari in una soluzione PBS: Sol di ibridazione (1:1)

Incubare per 10 min solo in soluzione di ibridazione

Bollire per 10min la sonda marcata con digossigenina e raffreddare in ghiaccio.

Ibridare tutta la notte a 45°C

Lavare 1h in sol. di ibridazione a 45°C

Lavare 30min in una sol 4:1 sol ibrid-PBS

Lavare 30min in una sol 3:2 sol ibrid-PBS

Lavare 30min in una sol 2:3 sol ibrid-PBS

Lavare 30min in una sol 1:4 sol ibrid-PBS

Sciacquare più volte in PBS

Incubare per almeno 1h con l'anticorpo secondario (Anti-Digoxigenina AP-Fab fragment Roche)

Lavare 3 volte in PBS
Lavare 2 volte in Buffer 3
Buffer 3

- 100mM NaCl
- 50mM MgCl₂
- 100mM Tris-HCl pH 9.5
- 0.1% Tween

Colorazione: ad 1ml di Buffer 3 aggiungere 18 µl di NBT/BCIP stock solution (Roche).
Fermare la reazione in TE 1X

Preparazione di ceppi transgenici di drosophila mediante trasformazione basata sull'iniezione di vettori derivati dall'elemento trasponibile P.

Preparazione dei costrutti adoperati per la trasformazione:

- Per la preparazione del costrutto chimerico utilizzato per l'esperimento di recupero della mutazione *mfl*⁹⁵ è stato inserito nel vettore pTUBU un frammento genomico di 8 kb comprendente la regione codificante del gene *mfl*, più 4 kb della regione genomica fiancheggiante il gene in 3'. Il clone ricombinante ottenuto, selezionato per la presenza dell'inserto è stato successivamente digerito con gli enzimi di restrizione *SacI/NcoI*, per rimuovere il frammento che contiene l'introne da cui si origina lo snoRNA H1. Contemporaneamente è stato digerito con gli stessi enzimi di restrizione un clone di cDNA del gene *mfl* per ottenere il frammento, di circa 500 cb, corrispondente alla sequenza del gene *mfl* che comprende la fine del quarto e l'inizio del quinto esone.
- Per la preparazione del costrutto utilizzato per l'esperimento di overespressione dei geni *DmSnR60a/d* ed *snm60a/b* è stato inserito nel vettore pUAS-T, la regione genomica compresa tra gli oligo..... Il frammento genomico è stato amplificato per PCR.

1. Preparazione del plasmide da iniettare.

Negli esperimenti di trasformazione, il DNA deve essere accuratamente purificato. Il plasmide da iniettare è stato purificato per centrifugazione in gradiente di Cloruro di Cesio-Bromuro d'etidio. Si precipita il DNA con etanolo 95%, si lava il pellet con etanolo 70% per rimuovere i sali e si reispone il DNA in una soluzione contenente KCl 5mM e fosfato 0.1mM ad una concentrazione finale di 0.5µg/µl.

2. Condizioni d'iniezione

L'iniezione viene condotta a 18°C, perché lo sviluppo dell'embrione è notevolmente ridotto a questa temperatura rispetto a 25°C (in modo tale da avere più tempo a disposizione per decorionare gli embrioni ed iniettarli prima che essi raggiungano lo stadio in cui si formano le cellule polari).

3. Preparazione degli aghi

Gli aghi per la microiniezione sono preparati a partire da capillari di vetro borosilicato, utilizzando un dispositivo che può essere programmato in modo da modulare le dimensioni del microago in funzione delle necessità. Da ogni capillare si ottengono due aghi, le cui estremità vengono rotte al microscopio. L'ago viene poi posizionato su di un porta capillari, a sua volta fissato a un micromanipolatore, ed è riempito con il DNA da utilizzare per la trasformazione. L'ago è collegato, tramite un tubicino di plastica, ad una siringa che permette di ottenere diversi livelli di pressione.

4. Raccolta degli embrioni e decorionamento

Per la deposizione vengono utilizzati individui adulti, provenienti dal ceppo 3664, che abbiano 2-3 giorni. Gli embrioni raccolti ad intervalli di 30 min, vengono allineati su di un lato del nastro biadesivo, fissato ad un vetrino. Con l'uso di un pennellino si elimina il corion facendo rotolare l'embrione sul nastro biadesivo. Successivamente gli embrioni privi di corion vengono allineati per permettere l'iniezione del DNA in modo che il polo che sporge sia quello vegetativo (posteriore). E' importante che il polo che viene iniettato sia il vegetativo, perché in questa regione dell'embrione si svilupperanno le cellule che daranno origine alla linea germinale. Gli embrioni vengono disidratati per trasferimento su una capsula Petri contenente CaCl₂, per renderli recettivi al DNA iniettato, quindi, vengono ricoperti con olio (Halocarbon oil 10S), che impedisce un'ulteriore disidratazione, ma che consente gli scambi gassosi.

5. Iniezione

L'iniezione viene effettuata con un microscopio ed un micromanipolatore. Vengono iniettati solo gli embrioni che non sono in uno stadio troppo avanzato dello sviluppo; gli altri sono

eliminati. La quantità ottimale di DNA da iniettare corrisponde a circa 1/100 del volume dell'uovo (Rubin and Spradling, 1982)

6. Incroci necessari per l'identificazione dei trasformanti

Gli embrioni che sono sopravvissuti al trauma della microiniezione, danno origine ad individui che sono incrociati singolarmente con moscerini appartenenti al ceppo selvatico *w* (occhio bianco). Gli individui trasformanti in G1 sono riconosciuti dal colore dell'occhio, che può variare dal giallo chiaro al rosso, in funzione del sito di inserzione del costrutto sul cromosoma. Questi individui con l'occhio pigmentato sono, a loro volta incrociati con ceppi *w* per dare origine a ceppi indipendenti di trasformazione.

Elenco delle soluzioni:

soluzioni utilizzate nelle corse elettroforetiche:

TAE 50X : 2M Tris-acetato; 0.05M EDTA pH 8.0

TBE 5X : 0.445M Tris; 0.445M Acido borico; 0.01M EDTA

RB10X (pH 7.0): 0.2M Mops; 0.05M Sodioacetato; 0.01% EDTA

SSC 20X : 3M Cloruro di sodio; 0,3M Sodio citrato tribasico

Soluzioni usate per la crescita dei batteri e per la purificazione del DNA plasmidico

PEG 13%-Sodio cloruro 1.6M

LB : 1% Cloruro di sodio; 1% Triptone; 0.5% Estratto di lievito

Sol A : 50mM Tris-Cl pH 7.5; 10mM EDTA

Sol B : 0,2M Idrossido di sodio ; 1% SDS

Sol C : 1,32M Potassio acetato pH 4.8

TE 10X : 0,1M Tris-Cl pH 7.0; 0.01M EDTA pH 8.0

Buffer di omogeneizzazione per l'estrazione di RNA da adulti ed embrioni:

- 50mM Tris-Cl pH 7.5;
- 60mM Sodio cloruro;
- 20mM EDTA pH 8.0
- H₂O DEPC a volume;
- 1% SDS sterile

Buffer di omogeneizzazione per l'estrazione di RNA da larve:

- 20mM Tris-Cl pH 7.5;
- 0,2M Sodio cloruro
- 20mM EDTA pH 8.0
- H₂O DEPC a volume
- 2% SDS dopo sterilizzazione

Buffer di omogeneizzazione per l'estrazione del DNA genomico:

- 10mM Tris-Cl (pH 7.5);
- 60mM Sodio cloruro;
- 10mM EDTA;
- 0.15mM Spermina
- 0.15mM Spermidina

Formammide loading buffer:

- 80%di formammide deionizzata
- EDTA 10 mM pH8.0
- xylene cyanolo FF 1mg/ml
- blu di bromofenolo 1mg/ml

Sonde utilizzate negli esperimenti di Northern Blot:

Sonda 1: regione genomica di 0.1 kb compresa tra gli oligonucleotidi 5'-CACAAAGAACTTAAGGTGTG-3' e 5'-GATGTCTTGGGCAGTGTGTACC-3'

Sonda 2: regione genomica di 0.2 kb compresa tra gli oligonucleotidi 5'-GAGGTAAATATTTAATAACTAAAAG-3' and 5'-GATTCCTGTGGCATTCAATG-3'

Sonda 3: regione genomica di 0.6 kb compresa tra gli oligonucleotidi 5'-CAAGCCTCAATCTTTTCGATTGCCTTTC-3' and 5'-GATGTCTTGGGCAGTGTGTACC-3'.

- Esperimenti di 3'-RACE:

oligoT-Adaptor (5'-GACTCGAGTCGACATCGA(T)₁₇-3').

Oligo Adaptor (5'-GACTCGAGTCGACATCG -3')

esone 5 (5'-GACCATGGTGTGGTGG-3')

esone 6 (5'-CTCTATCGCTAGTTTCTTAGGTCTTAGC-3'),

esone 8 (5'-CTTCGAATAAACATAGGAATTAAGGTAAG-3')

esone 9 (5'-GGTCATGCAATATATGGACTATAAC-3')

- Esperimenti di 5'-RACE:

oligoT-Adaptor (5'-GACTCGAGTCGACATCGA(T)₁₇-3').

Oligo Adaptor (5'-GACTCGAGTCGACATCG -3')

esone 3-9 giunzione (5'-CCTAATCAACAAATCCATATTTTCGGG-3')

esone 3 (5'-GGGCACCACGCAGCTTCTC-3')

esone 3 più interno (5'-GGGACTTCACCAGACGGG-3')

- Esperimenti di RT-PCR

esone 2 (primer P1: 5'-CTTCCAAATCAAGCCCTCCTCCAAG-3'),

esone 5 (5'-GCCAGCAGCTCAAGAAGTCTCCCC-3')

esone 6 (5'-GTGTAGAAGTGCATACAAATTAC-3')

esone 10 (5'-GATTCCTGTGGCATTCAATG-3'),

esone 11 (5'-GCAATAAAGTGTCATGCG-3')

esone 12 (primer P2: 5'-GATGTCTTGGGCAGTGTGTACC-3').

- Per confermare la presenza del trascritto derivato dall'evento di splicing alternative:

primer P1 (come sopra)

primer P2 (come sopra)

primer P3, spanning exon 3-9 junction: 5'-CCTAATCAACAAATCCATATTTTCGGG-3'.

primer P4, spanning exon 5-6 junction: 5'-GGTGCTTAACTTGCCTTGTGGG-3'.

Come controllo della quantità di RNA utilizzata in ciascuna reazione sono stati utilizzati due oligonucleotidi specifici del trascritto del gene dell'*αTub84B*

esone 1 (5'-GTGAAACTTCCAATAAAAACTCAATATG-3')

esone 2 (5'-CCAGCAGGCGTTTCCAAT-3').

- Quantitative Real Time RT-PCR (qPCR)

mfl Ex 3 forward primer: 5'-GTTGCGCGT TCGTACTGTCTAC-3';

mfl Ex 4 reverse primer 5'-CCTCGCAACTAACCCAAAAAAC-3';

mfl Ex 3/9 junction reverse primer 5'-TTGCATAACCTAATCAACAAATCCA-3';

α -*Tub84B* Ex 1/2 forward primer 5'-GTGAAACACTTCCAATAAAAACTCAATATG-3';

α -*Tub84B* Ex 2 reverse primer 5'-CCAGCAGGCGTTTCCAAT -3'.

Bibliografia

Bibliografia

- Accardo MC, Giordano E, Riccardo S, Digilio FA, Iazzetti G, Calogero RA, Furia M.
A computational search for box C/D snoRNA genes in the *Drosophila melanogaster* genome.
Bioinformatics. 2004 Dec 12;20(18):3293-301. Epub 2004 Jul 9.
- Amaldi F, Pierandrei-Amaldi P.
TOP genes: a translationally controlled class of genes including those coding for ribosomal proteins. *Prog Mol Subcell Biol*. 1997;18:1-17. Review.
- Arnez JG, Steitz TA.
Crystal structure of unmodified tRNA(Gln) complexed with glutamyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure.
Biochemistry. 1994 Jun 21;33(24):7560-7.
- Aravind L, Koonin EV.
Novel predicted RNA-binding domains associated with the translation machinery.
J Mol Evol. 1999 Mar;48(3):291-302.
- Bachellerie JP, Cavaille J, Huttenhofer A.
The expanding snoRNA world.
Biochimie. 2002 Aug;84(8):775-90. Review.
- Ballarino M, Morlando M, Pagano F, Fatica A, Bozzoni I.
The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in *Saccharomyces cerevisiae*.
Mol Cell Biol. 2005 Jul;25(13):5396-403.
- Barth S, Hury A, Liang XH, Michaeli S.
Elucidating the role of H/ACA-like RNAs in trans-splicing and rRNA processing via RNA interference silencing of the *Trypanosoma brucei* CBF5 pseudouridine synthase.
J Biol Chem. 2005 Oct 14;280(41):34558-68. Epub 2005 Aug 17.
- Baker DL, Youssef OA, Chastkofsky MI, Dy DA, Terns RM, Terns MP.
RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP.
Genes Dev. 2005 May 15;19(10):1238-48. Epub 2005 May 3.
- Bier E.
Drosophila, the golden bug, emerges as a tool for human genetics. *Nat Rev Genet*. 2005 Jan;6(1):9-23. Review.
- Bousquet-Antonelli C, Henry Y, G'elugne JP, Caizergues-Ferrer M, Kiss T.
A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs. *EMBO J*. 1997 Aug 1;16(15):4770-6.
- Caffarelli E, Fatica A, Prislei S, De Gregorio E, Fragapane P, Bozzoni I.
Processing of the intron-encoded U16 and U18 snoRNAs: the conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA.
EMBO J. 1996 Mar 1;15(5):1121-31.
- Cavaille J, Bachellerie JP.
Processing of fibrillar-associated snoRNAs from pre-mRNA introns: an exonucleolytic process exclusively directed by the common stem-box terminal structure.
Biochimie. 1996;78(6):443-56.
- Charpentier B, Muller S, Branlant C.

Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res.* 2005 Jun 2;33(10):3133-44.

Chen L, Lullo DJ, Ma E, Celniker SE, Rio DC, Doudna JA.
Identification and analysis of U5 snRNA variants in *Drosophila*.
RNA. 2005 Oct;11(10):1473-7.

Darzacq X, Kittur N, Roy S, Shav-Tal Y, Singer RH, Meier UT.
Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells.
J Cell Biol. 2006 Apr 24;173(2):207-18. Epub 2006 Apr 17.

Dez C, Henras A, Faucon B, Lafontaine DBJ, Caizergues-Ferrer M and Yves H.

Stable expression in yeast of the mature form of human telomerase RNA depends on its association with the box H/ACA small nucleolar RNP proteins Cbf5p, Nhp2p and Nop10p. *Nucleic Acids Research.* 2001. 29: 598-603.

de Turrís V, Di Leva G, Caldarola S, Loreni F, Amaldi F, Bozzoni I. TOP promoter elements control the relative ratio of intron-encoded snoRNA versus spliced mRNA biosynthesis. *J Mol Biol.* 2004 Nov 19;344(2):383-94.

Donmez G, Hartmuth K, Luhrmann R. Modified nucleotides at the 5' end of human U2 snRNA are required for spliceosomal E-complex formation. *RNA.* 2004 Dec;10(12):1925-33. Epub 2004 Nov 3.

Dragon F, Pogacic V, Filipowicz W.
In vitro assembly of human H/ACA small nucleolar RNPs reveals unique features of U17 and telomerase RNAs. *Mol Cell Biol.* 2000 May;20(9):3037-48.

Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS.
MicroRNA targets in *Drosophila*. *Genome Biol.* 2003;5(1):R1. Epub 2003 Dec 12

Fatica A, Dlakic M, Tollervey D.
Naf1 p is a box H/ACA snoRNP assembly factor.
RNA. 2002 Dec;8(12):1502-14.

Ganot P, Bortolin ML, Kiss T.
Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell.* 1997 May 30;89(5):799-809.

Giordano E, Peluso I, Senger S, and Furia M. 1999. *minifly*, a *Drosophila* gene required for ribosome biogenesis. *J Cell Biol.* 144: 1123-113.

Giordano E, Rendina R, Peluso I, Furia M.
RNAi triggered by symmetrically transcribed transgenes in *Drosophila melanogaster*.
Genetics. 2002 Feb;160(2):637-48.

Godfrey AC, Kupsco JM, Burch BD, Zimmerman RM, Dominski Z, Marzluff WF, Duronio RJ. U7 snRNA mutations in *Drosophila* block histone pre-mRNA processing and disrupt oogenesis. *RNA.* 2006 Mar;12(3):396-409.

Grewal SS, Li L, Orian A, Eisenman RN, Edgar BA.
Myc-dependent regulation of ribosomal RNA synthesis during *Drosophila* development.
Nat Cell Biol. 2005 Mar;7(3):295-302. Epub 2005 Feb 20.

Gu X, Liu Y, Santi DV.

- The mechanism of pseudouridine synthase I as deduced from its interaction with 5-fluorouracil-tRNA. *Proc Natl Acad Sci U S A*. 1999 Dec 7;96(25):14270-5.
- Hamma T, Reichow SL, Varani G, Ferre-D'Amare AR.
The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat Struct Mol Biol*. 2005. 12(12):1101-7.
- He, J., Navarrete, S., Jasinski, M., Vulliamy, T., Bessler, M., and Mason, P.J. Targeted disruption of DKC1, the gene mutated in X-linked dyskeratosis congenita, causes embryonic lethality in mice. *Oncogene*. 2002. 21: 7740-4.
- Heiss, N. S., Knight, S. W., Vulliamy, T. J., Klauck, S. M., Wiemann, S., Mason, P.J., Poustka, A., and Dokal, I. X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet*. 1998 19: 32-8.
- Heiss, N.S., Girod, A., Salowsky, R., Wiemann, S., Pepperkok, R., and Poustka, A. 1999. Dyskerin localizes to the nucleolus and its mislocalization is unlikely to play a role in the pathogenesis of dyskeratosis congenita. *Hum Mol Genet*. 8: 2515-24.
- Heiss NS, Bachner D, Salowsky R, Kolb A, Kioschis P, Poustka A.
Gene structure and expression of the mouse dyskeratosis congenita gene, *dkc1*. *Genomics*. 2000 Jul 15;67(2):153-63.
- Henras A, Henry Y, Bousquet-Antonelli C, Noaillac-Depeyre J, Gelugne JP, Caizergues-Ferrer M. Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *EMBO J*. 1998 Dec 1;17(23):7078-90.
- Henras AK, Dez C, Henry Y.
RNA structure and function in C/D and H/ACA s(no)RNPs. *Curr Opin Struct Biol*. 2004 Jun;14(3):335-43. Review.
- Hirose T, and Steitz J.A.
Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. *Proc Natl Acad Sci U S A* 2001. 98: 12914-9.
- Hoang C, Ferre-D'Amare AR.
Cocrystal structure of a tRNA^{Psi55} pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell*. 2001 Dec 28;107(7):929-39.
- Holcik M, Lefebvre C, Yeh C, Chow T, Korneluk RG.
A new internal-ribosome-entry-site motif potentiates XIAP-mediated cytoprotection. *Nat Cell Biol*. 1999 Jul;1(3):190-2.
- Holcik M, Sonenberg N.
Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol*. 2005 Apr;6(4):318-27. Review.
- Huang L, Pookanjanatavip M, Gu X, Santi DV.
A conserved aspartate of tRNA pseudouridine synthase is essential for activity and a probable nucleophilic catalyst. *Biochemistry*. 1998 Jan 6;37(1):344-51.
- Huang ZP, Zhou H, He HL, Chen CL, Liang D, Qu LH.
Genome-wide analyses of two families of snoRNA genes from *Drosophila melanogaster*, demonstrating the extensive utilization of introns for coding of snoRNAs. *RNA*. 2005 11(8):1303-16.
- Hughes TA.

- Regulation of gene expression by alternative untranslated regions.
Trends Genet. 2006 Mar;22(3):119-22. Epub 2006 Jan 23. Review.
- Huttenhofer A, Kiefmann M, Meier-Ewert S, O'Brien J, Lehrach H, Bachellerie JP, Brosius J.
RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. EMBO J. 2001 Jun 1;20(11):2943-53.
- Huttenhofer A., Brosius J., and Bachellerie J.P.
RNomics: identification and function of small, non-messenger RNAs. Curr Opin Chem Biol. 6: 835-43.
- Jiang W., Middleton K., Yoon H.J., Fouquet C., and Carbon J.
An essential yeast protein, CBF5p, binds in vitro to centromeres and microtubules. Mol Cell Biol. 1993.13: 4884-93.
- Kai T., Williams D, Spradling AC.
The expression profile of purified Drosophila germline cells Dev. Biol. 2005. 283:486-502
- Kendall A, Hull MW, Bertrand E, Good PD, Singer RH, Engelke DR.
A CBF5 mutation that disrupts nucleolar localization of early tRNA biosynthesis in yeast also suppresses tRNA gene-mediated transcriptional silencing. 2000 Proc Natl Acad Sci U S A. 97:13108-13.
- Kiss T, Filipowicz W.
Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns.
Genes Dev. 1995 Jun 1;9(11):1411-24.
- Kiss-Laszlo Z, Henry Y, Bachellerie JP, Caizergues-Ferrer M, Kiss T.
Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs.
Cell. 1996 Jun 28;85(7):1077-88.
- Kiss-Laszlo Z, Henry Y, Kiss T.
Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA.
EMBO J. 1998 Feb 2;17(3):797-807.
- Kiss T.
Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs.
EMBO J. 2001 Jul 16;20(14):3617-22. Review.
- Kiss T.
Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions.
Cell. 2002. 19: 145-8
- Kiss T.
Biogenesis of small nuclear RNPs.
J Cell Sci. 2004 Dec 1;117(Pt 25):5949-51. Review.
- Kiss T.
SnoRNP biogenesis meets Pre-mRNA splicing.
Mol Cell. 2006 Sep 15;23(6):775-6. Review
- Knight, S. W., Heiss, N.S., Vulliamy, T.J., Aalfs, C.M., McMahon, C., Richmond, P., Jones, A., Hennekam, R.C.M., Poustka, A., Mason, P.J., and Dokal, I. 1999. Unexplained aplastic anaemia, immunodeficiency, and cerebellar hypoplasia (Hoyeraal-Hreidarsson syndrome) due to mutations in the dyskeratosis congenita gene, DKC1. Brit. J. Haemat. 107: 335-339.

- Koonin EV.
Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases.
Nucleic Acids Res. 1996 Jun 15;24(12):2411-5.
- Kullmann M, Gopfert U, Siewe B, Hengst L.
ELAV/Hu proteins inhibit p27 translation via an IRES element in the p27 5'UTR.
Genes Dev. 2002 Dec 1;16(23):3087-99.
- Lafontaine DL, Bousquet-Antonelli C, Henry Y, Caizergues-Ferrer M, Tollervey D.
The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase.
Genes Dev. 1998 Feb 15;12(4):527-37.
- Lai EC.
miRNAs: whys and wherefores of miRNA-mediated regulation.
Curr Biol. 2005 Jun 21;15(12):R458-60. Review.
- Ledda M, Di Croce M, Bedini B, Wannenes F, Corvaro M, Boyl PP, Caldarola S, Loreni F, Amaldi F. 2005. Effect of 3'UTR length on the translational regulation of 5'-terminal oligopyrimidine mRNAs. *Gene.* 344: 213-20.
- Maden B.E.
Mapping 2'-O-methyl groups in ribosomal RNA. *Methods.* 2001. 25: 374-378
- Manival X, Charron C, Fourmann JB, Godard F, Charpentier B, Branlant C.
Crystal structure determination and site-directed mutagenesis of the *Pyrococcus abyssi* aCBF5-aNOP10 complex reveal crucial roles of the C-terminal domains of both proteins in H/ACA sRNP activity. *Nucleic Acids Res.* 2006. 34(3):826-39.
- Mason PJ.
Stem cells, telomerase and dyskeratosis congenita.
Bioessays. 2003. 25:126-33
- Meier U.T., and Blobel G.
NAP57, a mammalian nucleolar protein with a putative homolog in yeast and bacteria. *J Cell Biol.* 1994. 127: 1505-14.
- Meier UT.
Dissecting dyskeratosis.
Nat Genet. 2003 Feb;33(2):116-7.
- Mignone F, Grillo G, Licciulli F, Iacono M, Liuni S, Kersey PJ, Duarte J, Saccone C, Pesole G. UTRdb and UTRsite: a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. 2005. *Nucleic Acids Res.* 33 (Database issue)
- Mitchell JR, Wood E, Collins K.
A telomerase component is defective in the human disease dyskeratosis congenita.
Nature. 1999 Dec 2;402(6761):551-5.
- Mochizuki Y., He J., Kulkarni S., Bessler M., and Mason P.J.
Mouse dyskerin mutations affect accumulation of telomerase RNA and small nucleolar RNA, telomerase activity, and ribosomal RNA processing. *Proc. Natl. Acad. Sci U SA.* 2004. 101: 10756-61.
- Motorin Y, Keith G, Simon C, Foiret D, Simos G, Hurt E, Grosjean H.
The yeast tRNA:pseudouridine synthase Pus1p displays a multisite substrate specificity.
RNA. 1998 Jul;4(7):856-69.

- Muller B, Schumperli D.
The U7 snRNP and the hairpin binding protein: Key players in histone mRNA metabolism. *Semin Cell Dev Biol.* 1997 Dec;8(6):567-76.
- Nesic D, Kramer A.
Domains in human splicing factors SF3a60 and SF3a66 required for binding to SF3a120, assembly of the 17S U2 snRNP, and prespliceosome formation. *Mol Cell Biol.* 2001 Oct;21(19):6406-17
- Newby MI, Greenbaum NL.
Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nat Struct Biol.* 2002 Dec;9(12):958-65.
- Ni J, Tien AL, Fournier MJ.
Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell.* 1997 May 16;89(4):565-73.
- Nicoloso M, Qu LH, Michot B, Bachellerie JP.
Intron-encoded, antisense small nucleolar RNAs: the characterization of nine novel species points to their direct role as guides for the 2'-O-ribose methylation of rRNAs. *J Mol Biol.* 1996 Jul 12;260(2):178-95.
- Normand C, Capeyrou R, Quevillon-Cheruel S, Mougou A, Henry Y, Caizergues-Ferrer M.
Analysis of the binding of the N-terminal conserved domain of yeast Cbf5p to a box H/ACA snoRNA. *RNA.* 2006 Oct;12(10):1868-82. Epub 2006 Aug 24.
- Nurse K, Wrzesinski J, Bakin A, Lane BG, Ofengand J.
Purification, cloning, and properties of the tRNA psi 55 synthase from *Escherichia coli*. *RNA.* 1995 Mar;1(1):102-12.
- Oskarsson T, Trumpp A.
The Myc trilogy: lord of RNA polymerases. *Nat Cell Biol.* 2005 7(3):215-7.
- Phillips, B., Billin, A.N., Cadwell, C., Buchholz, R., Erickson, C., Merriam, J.R., Carbon, J., and Poole, S.J. The Nop60B gene of *Drosophila* encodes an essential nucleolar protein that functions in yeast. *Mol Gen. Genet.* 1998. 260: 20-9.
- Rashid R, Liang B, Baker DL, Youssef OA, He Y, Phipps K, Terns RM, Terns MP, Li H.
Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol Cell.* 2006. 21(2):249-60
- Richard P, Kiss T.
Integrating snoRNP assembly with mRNA biogenesis. *EMBO Rep.* 2006 Jun;7(6):590-2.
- Rehwinkel J, Letunic I, Raes J, Bork P, Izaurralde E.
Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. *RNA.* 2005 Oct;11(10):1530-44.
- Ronshaugen M, Biemar F, Piel J, Levine M, Lai EC.
The *Drosophila* microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. *Genes Dev.* 2005 Dec 15;19(24):2947-52.
- Ruggero D, Grisendi S, Piazza F, Rego E, Mari F, Rao PH, Cordon-Cardo C, and Pandolfi PP.
Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. *Science.* 2003. 299: 259-62.
- Rusinov V, Baev V, Minkov IN, Tabler M.

- MicroInspector: a web tool for detection of miRNA binding sites in an RNA sequence. *Nucleic Acids Res.* 2005 Jul 1;33(Web Server issue):W696-700.
- Salowsky R, Heiss NS, Benner A, Wittig R, Poustka A.
Basal transcription activity of the dyskeratosis congenita gene is mediated by Sp1 and Sp3 and a patient mutation in a Sp1 binding site is associated with decreased promoter activity. *Gene.* 2002 Jun 26;293(1-2):9-19.
- Samarsky DA, Fournier MJ, Singer RH, Bertrand E.
The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. *EMBO J.* 1998 Jul 1;17(13):3747-57.
- Sambrook, J., and Russell, D.W. 2001. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schreiber-Agus N, Stein D, Chen K, Goltz JS, Stevens L, DePinho RA. Drosophila Myc is oncogenic in mammalian cells and plays a role in the diminutive phenotype. *Proc Natl Acad Sci U S A.* 1997;94(4):1235-40 Seitz H, Royo H, Lin SP
- Schuler GD, Altschul SF, Lipman DJ.
A workbench for multiple alignment construction and analysis. *Proteins.* 1991;9(3):180-90.
- Smith CM, Steitz JA.
Sno storm in the nucleolus: new roles for myriad small RNPs. *Cell.* 1997 May 30;89(5):669-72. Review.
- Smith CM, Steitz JA.
Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. *Mol Cell Biol.* 1998 Dec;18(12):6897-909.
- Spradling, A. 1993. Developmental genetics of oogenesis . In "Drosophila development" (M. Bate and A. Martinez Arias Eds), pp1-70. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Tycowski KT, Shu MD, Steitz JA.
A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. *Genes Dev.* 1993 Jul;7(7A):1176-90.
- Tycowski KT, Shu MD, Steitz JA.
A mammalian gene with introns instead of exons generating stable RNA products. *Nature.* 1996 Feb 1;379(6564):464-6.
- Tycowski KT, You ZH, Graham PJ, Steitz JA.
Modification of U6 spliceosomal RNA is guided by other small RNAs. *Mol Cell.* 1998 Nov;2(5):629-38.
- Tycowski, K.T., and Steitz, J.A. 2001. Non-coding snoRNA host genes in Drosophila: expression strategies for modification guide snoRNAs. *Eur J Cell Biol.* 80: 119-25.
- Valadkhan S, Manley JL.
Characterization of the catalytic activity of U2 and U6 snRNAs. *RNA.* 2003 Jul;9(7):892-904.
- Villa T, Ceradini F, Bozzoni I.
Identification of a novel element required for processing of intron-encoded box C/D small nucleolar RNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 2000 Feb;20(4):1311-20.

- Vitali P, Basyuk E, Le Meur E, Bertrand E, Muscatelli F, Cavaille J, Huttenhofer A.
ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs. *J Cell Biol.* 2005 Jun 6;169(5):745-53.
- Vulliamy TJ, Knight SW, Heiss NS, Smith OP, Poustka A, Dokal I, Mason PJ.
Dyskeratosis congenita caused by a 3' deletion: germline and somatic mosaicism in a female carrier. *Blood.* 1999. 94:1254-60
- Vulliamy TJ, Knight SW, Mason PJ, Dokal I.
Very short telomeres in the peripheral blood of patients with X-linked and autosomal dyskeratosis congenita. *Blood Cells Mol Dis.* 2001 Mar-Apr;27(2):353-7.
- Vulliamy TJ, Marrone A, Knight SW, Walne A, Mason PJ, Dokal I.
Mutations in dyskeratosis congenita: their impact on telomere length and the diversity of clinical presentation. *Blood.* 2006. 107:2680-5.
- Wang C, Meier UT.
Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *EMBO J.* 2004 Apr 21;23(8):1857-67. Epub 2004 Mar 25.
- Wrzesinski J, Bakin A, Nurse K, Lane BG, Ofengand J.
Purification, cloning, and properties of the 16S RNA pseudouridine 516 synthase from *Escherichia coli*.
Biochemistry. 1995 Jul 11;34(27):8904-13.
- Watkins NJ, Gottschalk A, Neubauer G, Kastner B, Fabrizio P, Mann M, Luhrmann R.
Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA.* 1998 Dec;4(12):1549-68.
- Weinstein LB, Steitz JA.
Guided tours: from precursor snoRNA to functional snoRNP.
Curr Opin Cell Biol. 1999 Jun;11(3):378-84. Review.
- Wong JM, Kyasa MJ, Hutchins L, Collins K.
Telomerase RNA deficiency in peripheral blood mononuclear cells in X-linked dyskeratosis congenita. *Hum Genet.* 2004 Oct;115(5):448-55. Epub 2004 Sep 3.
- Yaghmai R, Kimyai-Asadi A, Rostamiani K, Heiss NS, Poustka A, Eyaid W, Bodurtha J, Nousari HC, Hamosh A, Metzberg A. Overlap of dyskeratosis congenita with the Hoyeraal-Hreidarsson syndrome. *J Pediatr.* 2000 Mar;136(3):390-3.
- Yang PK, Rotondo G, Porras T, Legrain P, Chanfreau G.
The Shq1p.Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis. *J Biol Chem.* 2002 Nov 22;277(47):45235-42. Epub 2002 Sep 11.
- Yang C, McPheeters DS, Yu YT.
Psi35 in the branch site recognition region of U2 small nuclear RNA is important for pre-mRNA splicing in *Saccharomyces cerevisiae*. *J Biol Chem.* 2005 Feb 25;280(8):6655-62. Epub 2004 Dec 20
- Yang, P.K., Hoareau, C., Froment, C., Monsarrat, B., Henry, Y., and Chanfreau, G. Cotranscriptional recruitment of the pseudouridylsynthetase Cbf5p and of the RNA binding protein Naf1p during H/ACA snoRNP assembly. *Mol. Cell. Biol.* 2005. 25: 3295-304.
- Yang JH, Zhang XC, Huang ZP, Zhou H, Huang MB, Zhang S, Chen YQ, Qu LH.

snoSeeker: an advanced computational package for screening of guide and orphan snoRNA genes in the human genome. *Nucleic Acids Res.* 2006;34(18):5112-23.

Yarian CS, Basti MM, Cain RJ, Ansari G, Guenther RH, Sochacka E, Czerwinska G, Malkiewicz A, Agris PF. Structural and functional roles of the N1- and N3-protons of psi at tRNA's position 39. *Nucleic Acids Res.* 1999 Sep 1;27(17):3543-9.

Yoon A, Peng G, Brandenburg Y, Zollo O, Xu W, Rego E, Ruggero D. Impaired control of IRES-mediated translation in X-linked dyskeratosis congenita. *Science* 2006. 312(5775):902-6.

Zhang B, Wang Q, Pan X.
MicroRNAs and their regulatory roles in animals and plants.
J Cell Physiol. 2006 Nov 9;210(2):279-289

Zhao X, Yu YT.
Pseudouridines in and near the branch site recognition region of U2 snRNA are required for snRNP biogenesis and pre-mRNA splicing in *Xenopus* oocytes. *RNA.* 2004 Apr;10(4):681-90.