



UNIVERSITÀ DEGLI STUDI DI PARMA
Dipartimento di Biochimica e Biologia Molecolare

DOTTORATO DI RICERCA IN BIOCHIMICA E BIOLOGIA MOLECOLARE
XX CICLO

**BIOGENESI DEGLI snoRNA IN *Saccharomyces cerevisiae*:
IDENTIFICAZIONE DEL PROMOTORE TRASCRIZIONALE E DI
FATTORI RICHIESTI PER IL PROCESSAMENTO**

Coordinatore del Corso:
Prof. Gian Luigi Rossi

Tutore:
Prof. Giorgio Dieci

Dottoranda:
Milena Preti

1. INTRODUZIONE.....	2
1.1 RNA non codificanti	2
1.1.1 RNA “housekeeping”	3
1.1.2 RNA di regolazione.....	5
1.2 Small nucleolar RNA	6
1.3 snoRNA C/D box	7
1.4 snoRNA H/ACA box	8
1.5 Le “core proteins”	9
1.5.1 Proteine della snoRNP C/D box	9
1.5.2 Proteine della snoRNP H/ACA box	11
1.6 Le funzioni degli snoRNA.....	12
1.6.1 Processamento del pre-rRNA	12
1.6.2 Modificazioni chimiche dell’rRNA.....	15
1.6.3 Modificazioni chimiche di altri RNA.....	17
1.6.4 snoRNA “orfani”	17
1.7 Rilascio degli snoRNA dal pre-rRNA.....	17
1.8 snoRNA tessuto-specifici e “imprinting”	18
1.9 Small Cajal body RNAs (scaRNAs).....	18
1.10 Gli sRNA degli Archaea	19
1.11 Organizzazione genomica dei geni per snoRNA.....	20
1.12 Trascrizione degli snoRNA.....	24
1.13 Processamento e maturazione degli snoRNA	27
1.14 Assemblaggio della snoRNP	31
1.14.1 Fattori richiesti per l’assemblaggio	32
1.15 Localizzazione degli snoRNA	33
2. SCOPO DELLA RICERCA.....	35
3. MATERIALI E METODI	37
3.1 Ceppi	37
3.2 PCR	38
3.2.1 Costruzione del mutante <i>SNR52 Cdown</i>	38
3.2.2 Costruzione dei mutanti <i>SNR52 Ddown</i> e <i>SNR52 D less</i>	40
3.2.3 Costruzione dei mutanti per lo studio della regione a monte dei geni per snoRNA.....	41
3.3 Clonaggi	44
3.4 Trasformazione di lievito (“quick yeast transformation”)	45

3.5 Estrazione dell'RNA	45
3.6 Northern blot.....	46
3.7 Deplezione di Nop1p, Nop58p, Snu13p	47
3.8 Chromatin immunoprecipitation (ChIP).....	48
3.9 Deplezione d'azoto	50
3.10 Induzione da glucosio.....	51
3.11 Primer extension	51
3.11.1 Sequenziamento	51
3.12 Isolamento dell'RNA poli A+	52
3.13 Trascrizione in vitro.....	53
3.14 Analisi computazionale.....	53
4. RISULTATI E DISCUSSIONE	54
4.1 Analisi delle regioni a monte dei geni per snoRNA	54
4.1.1 Analisi computazionale	54
4.1.2 Analisi sperimentale: scelta dei geni	59
4.1.3 Analisi della regione a monte di <i>SNR31</i>	60
4.1.4 Analisi della regione a monte di <i>SNR32</i>	62
4.1.5 Analisi della regione a monte di <i>SNR64</i>	65
4.1.6 Analisi della regione a monte di <i>SNR65</i>	68
4.1.7 Legame di proteine a monte dei geni per snoRNA	71
4.1.8 Espressione degli snoRNA in particolari condizioni di crescita.....	74
4.2 Maturazione dello snoRNA snR52	77
4.2.1 Il gene <i>SNR52</i>	77
4.2.2 Ruolo della C box e della D box nella biosintesi di snR52	78
4.2.3 Ruolo delle proteine della snoRNP nella maturazione di snR52.....	81
4.2.4 Maturazione dell'estremità 3' di snR52	84
5. BIBLIOGRAFIA	87
6. APPENDICE	108
RINGRAZIAMENTI	124

1. INTRODUZIONE

1.1 RNA non codificanti

Il termine RNA non codificante (ncRNA) viene utilizzato per indicare una molecola di RNA che non codifica per una proteina ed è ormai chiaro che molecole di questo tipo rivestono un importante ruolo nella regolazione di diverse attività cellulari.

La porzione di genoma che codifica per i ncRNA è ridotta nei Procarioti (meno del 25%), mentre negli Eucarioti, e in particolare negli Eucarioti superiori, la gran parte del trascrittoma è rappresentato da RNA non codificanti per proteine (25-50% del genoma negli Eucarioti inferiori e più del 50% in quelli superiori) (Mattick, 2004).

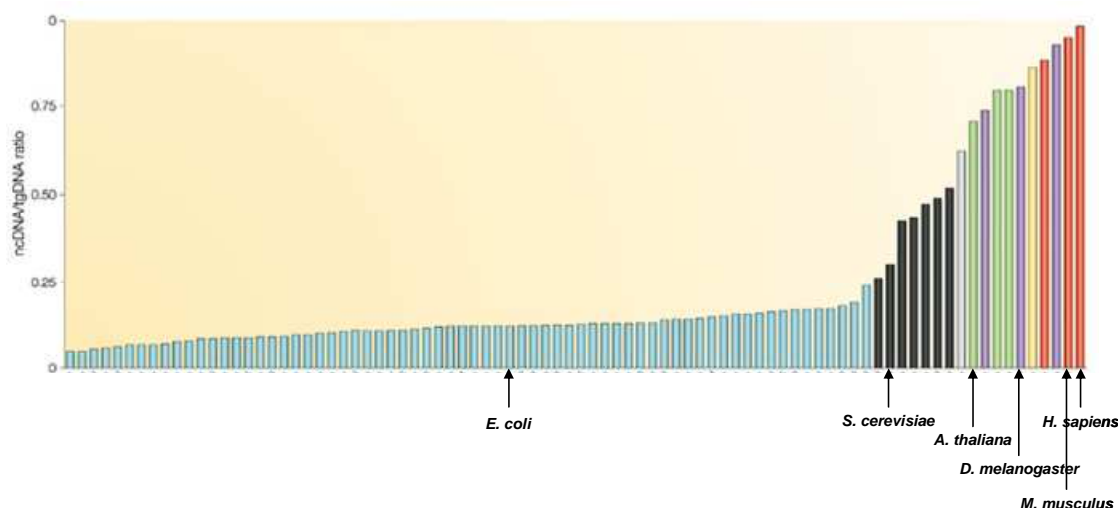


Figura 1.1: il rapporto DNA non codificante/DNA totale (riportato in ordinata) aumenta all'aumentare della complessità degli organismi. In azzurro sono rappresentati i Procarioti, in nero gli Eucarioti inferiori, in grigio *N. crassa*, in verde le piante, in viola gli invertebrati (nematodi, insetti), in giallo *C. intestinalis* (urocordato) e in rosso i vertebrati (adattato da Mattick, 2004).

I primi ncRNA vennero identificati negli anni Sessanta grazie al loro alto livello di espressione; successivamente altri furono scoperti mediante subfrazionamento di estratti nucleari, altri grazie all'associazione con proteine specifiche, alcuni grazie a mutazioni o particolari fenotipi dovuti alla sovraespressione. Negli ultimi anni sono state effettuate ricerche bioinformatiche mirate a individuare brevi motivi di sequenza

caratteristici di alcuni ncRNA o altre caratteristiche quali sequenze conservate in regioni intergeniche in specie correlate oppure la presenza di promotori o terminatori orfani (Storz, 2002).

I ncRNA possono essere divisi in due classi: gli RNA “housekeeping”, costitutivamente espressi ed essenziali per la vitalità della cellula, e gli RNA di regolazione, coinvolti nella regolazione di vari aspetti dell’espressione genica (Morey and Avner, 2004).

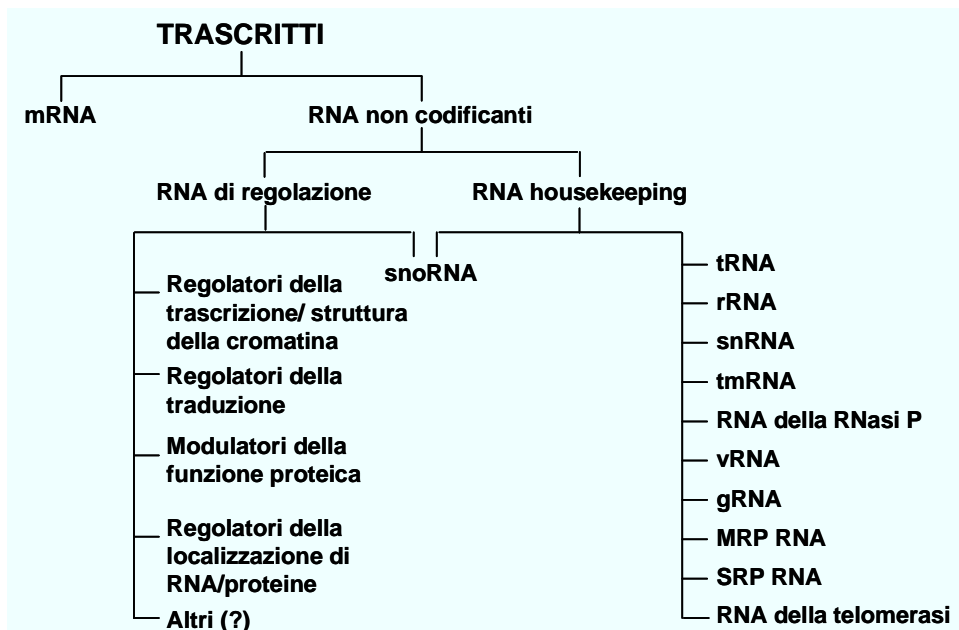


Figura 1.2: rappresentazione del trascrittoma. Sono elencati sia RNA presenti in tutti gli organismi (es.: tRNA), sia RNA comuni ai soli Eucarioti (es.: snRNA), sia RNA esclusivamente batterici (es.: tmRNA). Per altri dettagli vedi testo.

1.1.1 RNA “housekeeping”

Di questa classe fanno parte gli RNA coinvolti nel processamento e nella modificazione di trascritti primari, nella traduzione e nel controllo di qualità di quest’ultima, nonché l’RNA della telomerasi, l’RNA della SRP e i “vault RNA”.

Tra i primi ritroviamo gli “small nuclear RNA” (snRNA), gli “small nucleolar RNA” (snoRNA), i “guide RNA (gRNA) e l’RNA della RNasi P.

- **snRNA:** questi RNA abbondanti e non poliadenilati agiscono nel nucleoplasma; formano complessi ribonucleoproteici coinvolti nello “splicing” del pre-mRNA, in cui il ruolo della componente ribonucleica è quello di appaiarsi alla giunzione

introne-esone. Durante la reazione di “splicing” avvengono inoltre appaiamenti tra gli snRNA stessi. Nei Metazoi è presente una snRNP (U7 snRNP) non coinvolta nello “splicing” dell’ mRNA ma nella formazione dell’estremità 3’ dei pre-mRNA degli istoni (Matera *et al.*, 2007). Gli snRNA vengono trascritti dalla RNA polimerasi II, tranne U6, trascritto dalla RNA polimerasi III.

- **snoRNA:** di questo gruppo si parlerà in modo specifico in seguito.
- **gRNA:** questa classe di RNA guida la modificazione dell’estremità 5’ dei pre-mRNA (“RNA editing”) nei Tripanosomatidi.
- **RNasi P RNA:** la RNasi P è un complesso ribonucleoproteico richiesto per la formazione dell’estremità 5’ dei tRNA; la componente catalitica è costituita dall’RNA. In *S. cerevisiae* le prime fasi del processamento dei tRNA avvengono nel nucleolo, in cui si è osservato un arricchimento in RNA della RNasi P (Bertrand *et al.*, 1998). Correlato a questo RNA è anche l’RNA della RNasi MRP (vedi paragrafo 1.6.1).

Gli RNA coinvolti nella traduzione sono i “transfer RNA” (tRNA) e gli RNA ribosomiali (rRNA), mentre il tmRNA è implicato nel controllo di qualità.

- **tRNA:** questi piccoli RNA legano e trasportano uno specifico amminoacido e riconoscono, mediante l’appaiamento di basi, uno specifico codone sull’ mRNA, durante la traduzione. Presentano una caratteristica struttura secondaria a trifoglio e anche la struttura terziaria è conservata; nei tRNA maturi sono presenti particolari basi, generate per modificazione chimica, dopo la sintesi. Negli Eucarioti i tRNA vengono sintetizzati dalla RNA polimerasi III.
- **rRNA:** i ribosomi sono particelle ribonucleoproteiche costituite da due subunità. Nei Procarioti la subunità più grande (50S) comprende 31 proteine e due specie di RNA, il 23S e il 5S, mentre la più piccola (30S) comprende 21 proteine e l’rRNA 16S: quest’ultimo si appaia alla sequenza a monte del codone di inizio dell’ mRNA. Nei mammiferi la subunità grande (60S) è formata da 49 proteine e tre specie di RNA, 25S (28S nei Mammiferi), 5.8S e 5S, la piccola (40S) è costituita da 33 proteine e l’rRNA 18S. I geni per gli rRNA 25S, 5.8S e 16S sono riuniti in gruppi: la trascrizione da parte della RNA polimerasi I genera un unico RNA precursore, successivamente processato e modificato. L’rRNA 5S è trascritto invece dalla RNA polimerasi III. Sia nei Procarioti che

negli Eucarioti gli rRNA maggiori presentano una struttura secondaria complessa; è provato che l'rRNA interagisce con mRNA e tRNA durante la traduzione e attualmente si pensa che l'attività peptidil transferasica del ribosoma sia dovuta all'rRNA della subunità grande. Gli rRNA presentano basi modificate chimicamente che conferiscono stabilità alla struttura e si pensa che possano essere importanti per la funzione dell'RNA. Negli Eucarioti queste modificazioni sono prevalentemente guidate dagli snoRNA.

- **tmRNA:** questo RNA presente nei batteri è richiesto per il rilascio del ribosoma nel caso di blocco della traduzione. È costituito da una porzione “tRNA-like” amminoacilata che permette l'ingresso nel ribosoma e a cui viene trasferita la catena polipeptidica; l'mRNA in traduzione è rimpiazzato quindi dalla porzione “mRNA-like” codificante per un segnale di degradazione della proteina.
- **RNA della telomerasi:** la funzione di questo RNA è di fornire un sito di aggancio per la trascrittasi inversa, oltre a essere il template per la replicazione dei telomeri. Nei vertebrati l'RNA della telomerasi appartiene alla famiglia degli snoRNA H/ACA box e si localizza nel nucleolo tranne che nel momento della replicazione dei telomeri, mentre in lievito questo RNA appartiene alla famiglia degli snRNA leganti le proteine Sm.
- **SRP RNA:** questo RNA fornisce il sito di aggancio per le proteine che formano il complesso della SRP, che riconosce il peptide segnale presente all'N-terminale delle proteine destinate al reticolo endoplasmatico. Il legame di SRP al segnale avviene durante la traduzione della proteina.
- **vRNA:** i “vaults” sono complessi ribonucleoproteici citoplasmatici con funzione sconosciuta, presenti nei Vertebrati. Sono costituiti da tre proteine e da uno o più RNA la cui lunghezza è variabile; per gli RNA la sequenza primaria non è altamente conservata mentre è simile la struttura secondaria predetta. Questi RNA sono trascritti dalla RNA polimerasi III (Kickhoefer *et al.*, 1993).

1.1.2 RNA di regolazione

Questi RNA regolano aspetti dell'espressione genica sia nei Procarioti che negli Eucarioti, agendo a vari livelli, dalla regolazione della trascrizione al controllo della traduzione, e con diversi meccanismi di azione. In questo gruppo si trovano i “micro

RNA” (miRNA), gli RNA coinvolti nel “silencing” (p. es. *Xist* in *H. sapiens*) e nella attivazione di alcuni geni (p. es. *roX1* e *roX2* in *Drosophila*).

I miRNA possono indurre il silenziamento post-trascrizionale di un gene mediante repressione della traduzione dell'mRNA; sono coinvolti nel controllo dello sviluppo e/o nelle funzioni tessuto-specifiche. I primi miRNA sono stati individuati in *C. elegans* ma sono presenti anche in altri organismi (Zhao and Srivastava, 2007).

Tra gli altri RNA di regolazione ricordiamo *Xist*, la cui espressione e il conseguente accumulo sono richiesti per il silenziamento di uno dei cromosomi X nelle femmine di mammifero: *Xist* si propaga lungo il cromosoma che lo ha prodotto, inattivandolo. L'espressione di *Xist* è controllato da un altro ncRNA, *Tsix* (RNA antisenso rispetto a *Xist*), che si sovrappone al locus *Xist*; *Tsix* è espresso in modo continuo dal cromosoma attivo mentre è represso sul cromosoma inattivo.

1.2 Small nucleolar RNA

Gli snoRNA sono componenti di particelle ribonucleoproteiche (snoRNP) coinvolte nella maturazione non solo degli rRNA ma anche di altri RNA, nelle cellule eucariotiche; la loro funzione viene svolta nel nucleolo, da cui il nome.

Una snoRNP è costituita da una singola molecola di RNA, un ristretto numero di “core proteins” e un numero variabile di altre proteine caratteristiche di ciascuna particolare snoRNP. Il ruolo dell'RNA è quello di guidare la modificazione della molecola di RNA bersaglio mediante appaiamento delle basi, mentre l'attività catalitica è fornita da una delle proteine del “core”.

Le snoRNP hanno un'origine antica in quanto sono presenti ortologhi nei Protisti e negli *Archaea*: poiché si pensa che non esista un vero e proprio nucleolo in questi ultimi, gli snoRNA sono chiamati “small RNA” (sRNA) (Gaspin *et al.*, 2000; Omer *et al.*, 2000).

Un ristretto numero di snoRNA è richiesto per il processamento endonucleolitico del pre-rRNA mentre numerosi sono necessari per la formazione di nucleotidi modificati: pseudouridine e nucleotidi 2'-O-metilati.

Gli snoRNA sono divisi in due grandi famiglie, definite in base alla presenza di elementi di sequenza conservati: gli snoRNA H/ACA box che guidano le pseudouridilazioni e gli snoRNA C/D box che guidano le metilazioni sul ribosio. Gli

snoRNA coinvolti nel processamento del pre-rRNA possono appartenere a una famiglia oppure all'altra.

1.3 snoRNA C/D box

Questa classe di snoRNA presenta due elementi caratteristici: la box C (RUGAUGA) e la box D (UCUGA) posizionate rispettivamente vicino all'estremità 5' e 3' del trascritto maturo; frequentemente sono presenti anche altri due elementi degenerati, indicati come box C' e D' (Kiss-Laszlo *et al.*, 1998). Il motivo C/D box è coinvolto nel legame delle proteine del "core" e in altri processi probabilmente dipendenti da questo legame, quali la maturazione dello snoRNA stesso (se questo è trascritto in forma di precursore), la eventuale ipermetilazione del "cap" al 5' (Speckmann *et al.*, 2000) e la localizzazione, nonché la (eventuale) attività di metilazione.

Sono presenti una o due sequenze guida situate 1-2 nucleotidi al 5' rispetto alla box D o D', lunghe 9-21 nt e complementari alla regione da modificare. La metilazione avviene sul nucleotide posizionato 5 (a volte 6) nucleotidi a monte della box D o D'. Anche alcuni snoRNA richiesti nel processamento possiedono elementi guida, ad esempio la specie U14 che è richiesta sia per il processamento che per la metilazione dell'rRNA 18S (Li *et al.*, 1990; Dumbar and Baserga, 1998).

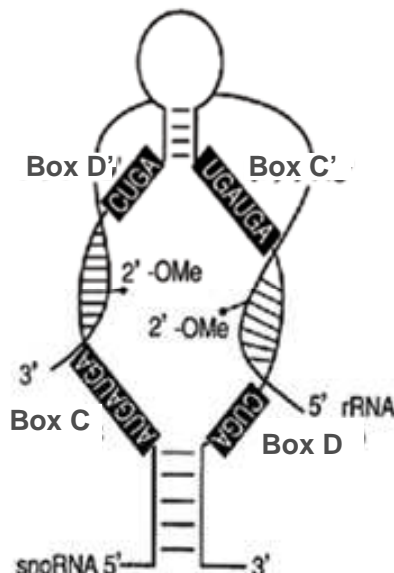


Figura 1.3: struttura di un tipico snoRNA C/D box. È mostrato anche l'appaiamento con l'rRNA substrato.

A questa famiglia appartiene anche il primo snoRNA identificato, nonché il meglio conosciuto: U3. Questo snoRNA è richiesto, assieme alle proteine con cui è associato,

per il taglio del pre-rRNA ai siti A₀, A₁ e A₂ in lievito (Beltrame *et al.*, 1994) (figura 1.8) e ai siti A0, 1, 2, 3 in *X. laevis*. (Borovjagin and Gerbi, 1999).

U3 presenta molte caratteristiche che lo distinguono dagli altri snoRNA C/D box: possiede sei brevi motivi di sequenza conservati, chiamati box A', A, C', B, C e D; è presente una regione detta “cerniera”, non conservata, ma importante, insieme agli elementi in 5' (box A' e A), per l'appaiamento con il pre-rRNA (Beltrame and Tollervey, 1995). I motivi formati dalle altre box (motivi B/C e C'/D) si trovano nella porzione 3' della molecola e sono richiesti per il legame delle proteine “core” (Samarsky and Fournier, 1998). L'ordine di successione degli elementi è invariante in organismi diversi.

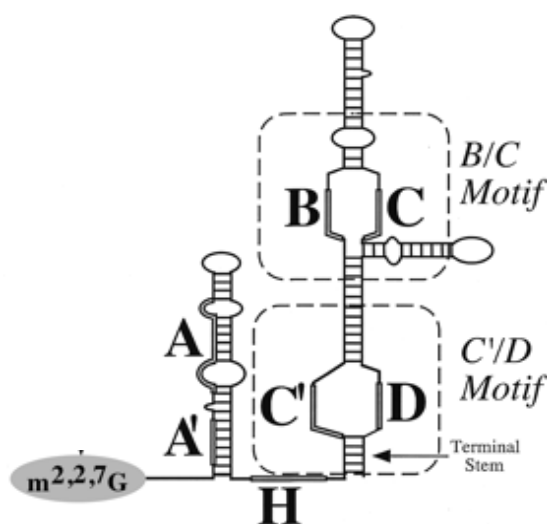


Figura 1.4: struttura dello snoRNA U3 di *X. laevis*. Sono evidenziate le sei box, la regione cerniera (H) il braccio terminale al 3' e il “cap” al 5'. I motivi B/C e C'/D sono racchiusi da linee tratteggiate

1.4 snoRNA H/ACA box

Anche questa famiglia è caratterizzata dalla presenza di due elementi di sequenza, la cui posizione nella struttura secondaria è conservata: la box H (ANANNA) e la box ACA (ACA); questi snoRNA hanno una tipica struttura secondaria composta da due forcine separate da una regione cerniera in cui si trova la box H, mentre la box ACA è posizionata all'estremità 3' (Ganot *et al.*, 1997). Nei Tripanosomatidi e in *E. gracilis*, al contrario, questo gruppo di snoRNA presenta una sola forcina e una box AGA, invece della box ACA (Liang *et al.*, 2004; Russell *et al.*, 2004).

Le due box sono coinvolte nel legame delle proteine del “core” e da questo dipendono la stabilità dell’RNA, la corretta localizzazione, il corretto processamento dello snoRNA (se necessario) e la eventuale attività di pseudouridilazione.

Le sequenze guida (una o due) sono contenute nelle regioni non appaiate che interrompono le forcine: l’uridina da modificare viene a trovarsi a una distanza di 14-16 nt dalla box H o ACA (figura 1.5).

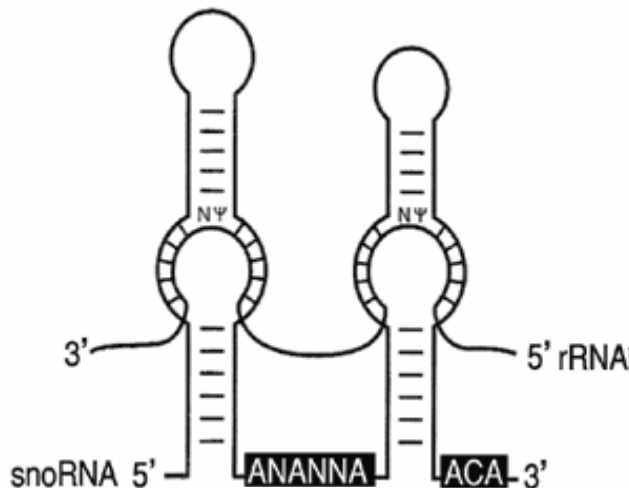


Figura 1.5: struttura tipica di uno snoRNA H/ACA box. È mostrata l’interazione con l’rRNA substrato

1.5 Le “core proteins”

Gli snoRNA di entrambe le famiglie si associano stabilmente a quattro proteine dette “core proteins”, diverse a seconda che lo snoRNA sia C/D o H/ACA box; si pensa che queste proteine siano universalmente presenti tra gli Eucarioti e sono stati inoltre identificati degli ortologi tra gli *Archaea*.

1.5.1 Proteine della snoRNP C/D box

Il nome indicato è quello della proteina di *S. cerevisiae*, tra parentesi è riportato il nome dell’ortologo presente in *H. sapiens*. Queste proteine sono tutte essenziali.

- **Snu13p (15.5K protein):** è la prima proteina che si lega allo snoRNA ed è anche presente nella snRNP U4. In entrambi i casi Snu13p interagisce con il motivo di struttura detto “kink turn”; è probabile che il legame di questa proteina provochi un riarrangiamento della struttura dell’RNA in modo da creare un sito di legame per le altre proteine. Snu13p è essenziale per la stabilità di tutti gli snoRNA C/D box e la sua deplezione, così come la deplezione delle

altre tre proteine del “core”, inibisce la localizzazione nucleolare della snoRNP (Watkins *et al.*, 2002). In *S. cerevisiae* questa proteina mostra una significativa somiglianza con la proteina ribosomiale L30 (Bachelierie *et al.*, 2002).

- **Nop1p (fibrillarina):** si suppone che questa sia la subunità catalitica della snoRNP C/D box, anche se non è ancora stato dimostrato direttamente; l'ipotesi è fondata sulla presenza di quattro brevi elementi di sequenza conservati in diversi tipi di metilasi che utilizzano S-adenosilmetionina come donatore di gruppi metile e sul fatto che mutazioni nella regione contenente questi motivi provoca il blocco della metilazione dell'rRNA in modo globale (Tollervey *et al.*, 1993). Inoltre snoRNP C/D box purificate presentano attività metilasica sito-specifica *in vitro* e Nop1p, presente in questi complessi, è in grado di legare S-adenosilmetionina (Galardi *et al.*, 2002). Nop1p si lega alle box D e D' (Cahill *et al.*, 2002), in lievito è essenziale per la stabilità solo di alcuni snoRNA C/D box (policistronici e intronici) e probabilmente è in grado di interagire con l'esosoma, responsabile del processamento in 3' degli snoRNA (Lafontaine and Tollervey, 2000). Anticorpi diretti contro la fibrillarina caratterizzano la malattia autoimmune umana scleroderma (Ochs *et al.*, 1985).
- **Nop58p (hNop58p):** è essenziale per l'accumulo degli snoRNA C/D box e si lega alla box C (Cahill *et al.*, 2002). In lievito interagisce con una proteina non essenziale, Nop17p, richiesta per la ritenzione nel nucleolo delle proteine del “core” e per la corretta formazione del complesso (Gonzales *et al.*, 2004).
- **Nop56p (hNop56p):** è correlata a Nop58p, essendo le due proteine identiche per il 45%, compreso il dominio carbossiterminale ricco in residui di lisina, aspartato e glutammato (dominio KKD/E) che comunque non è richiesto per il legame all'RNA (Lafontaine and Tollervey, 1999); entrambe sono simili a Prp31p (61-kDa protein), un componente del complesso U4-U6/U5 snRNP (Watkins *et al.*, 2000). Al contrario di Nop58p, richiede Nop1p per l'associazione allo snoRNA (Lafontaine and Tollervey, 2000) e si è inoltre osservato che, nonostante la somiglianza tra le proteine, nessuno dei due geni può funzionare come soppressore multicopia nel caso in cui sia distrutto l'altro (Gautier *et al.*, 1997).

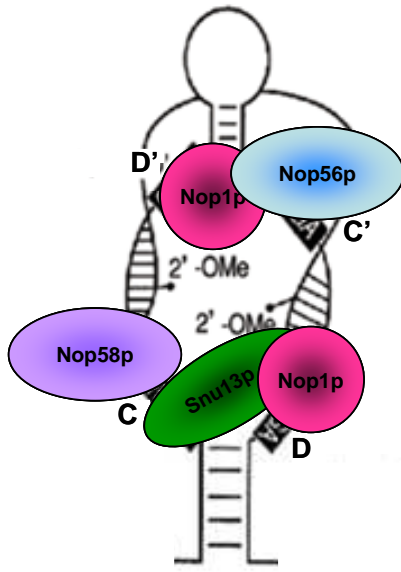


Figura 1.6: la snoRNP C/D box. Sono mostrate le quattro “core proteins”. Nop1p interagisce sia con la box D che D' ma Nop56p si lega solo a livello del motivo C/D'. Snu13p interagisce solo a livello del motivo C/D.

1.5.2 Proteine della snoRNP H/ACA box

Così come le proteine della snoRNP C/D box, anche queste sono tutte essenziali.

- **Cbf5p (diskerina):** è quasi certamente la componente catalitica del complesso, in quanto possiede tre sequenze tipiche delle pseudouridina sintasi. Mutazioni puntiformi introdotte in questi elementi provocano la distruzione globale delle pseudouridilazioni nell'rRNA di lievito (Zebarjadian *et al.*, 1999). Come Nop58p e Nop56p presenta un dominio KKE/D (Lafontaine and Tollervey, 1999). Mutazioni nella diskarina umana sono state correlate alla diskerosi congenita, una malattia rara che colpisce il midollo osseo, la pelle e le mucose.
- **Gar1p (hGar1p):** è in grado di interagire direttamente con lo snoRNA. Non è richiesta per la stabilità dell'RNA, al contrario delle altre tre proteine “core”, mentre è necessaria per la corretta funzione della snoRNP; inoltre è richiesta per la coesione della snoRNP (Henras *et al.*, 2004). Come Nop1p, presenta delle ripetizioni ricche in glicina e arginina (Lafontaine and Tollervey, 1999). Per la sua importazione nel nucleo è richiesta l'interazione con l'endonucleasi Rnt1p e la mancata formazione del complesso inibisce anche l'importazione di Cbf5p e Nhp2p (Tremblay *et al.*, 2002).
- **Nhp2p (hNhp2p):** anche questa proteina è in grado di interagire direttamente con lo snoRNA ed è correlata a Snu13p (Bachelierie *et al.*, 2002) (in lievito le due proteine sono identiche al 38% e simili al 61%); tuttavia non è stato

individuato nessun “kink turn” negli snoRNA H/ACA e il sito di legame per Nhp2p non è stato definito precisamente. La funzione di questa proteina, negli *Archaea*, è svolta dalla proteina ribosomiale L7Ae che lega anche gli sRNA C/D box e probabilmente fa parte della sRNP H/ACA box (Rozhdestvensky *et al.*, 2003).

- **Nop10p (hNop10p):** è richiesta per la stabilità dello snoRNA; può interagire direttamente con Cbf5, anche in assenza delle altre componenti della snoRNP (Henras *et al.*, 2004). Il legame di Nop10p a Cbf5p estende e rafforza il sito attivo di quest’ultima (Hamma *et al.*, 2005).

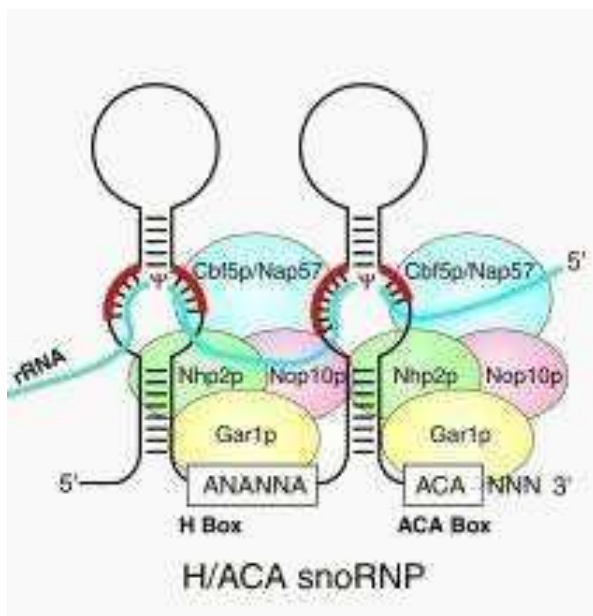


Figura 1.7: la snoRNP H/ACA box.

1.6 Le funzioni degli snoRNA

1.6.1 Processamento del pre-rRNA

La trascrizione dei geni per rRNA da parte della RNA polimerasi I genera un unico trascritto (pre-rRNA 35S), che successivamente viene processato per dare origine agli rRNA 18S, 5.8S e 25S.

In *Saccharomyces cerevisiae* sono richieste le snoRNP C/D U3 (siti di taglio A₀, A₁, A₂) e U14 (siti A₁, A₂), le snoRNP H/ACA snR30 e snR10 (siti A₁, A₂), oltre a numerose altre proteine quali Rnt1p, Rat1p, Xrn1p, il complesso dell’esosoma (coinvolti anche nella maturazione degli snoRNA) e la RNasi MRP (figura 1.8).

Quest'ultima è un complesso ribonucleoproteico, presente solo negli eucarioti, la cui prima funzione riportata è stata quella di taglio dei piccoli RNA che servono come "primer" nella replicazione del DNA mitocondriale, funzione da cui deriva il nome ("Mitochondrial RNA-Processing"). Successivamente si è osservata la localizzazione nucleolare e la richiesta per il taglio al sito A₃ (Schmitt and Clayton, 1993). La struttura secondaria dell'RNA MRP (che attualmente si tende a classificare come snoRNA, non appartenente però a nessuna delle due famiglie) è simile a quella dell'RNA dell'RNasi P, otto delle nove proteine della RNasi P fanno parte della RNasi MRP e inoltre dall'allineamento dei pre-rRNA batterici ed eucariotici è emerso che i siti di taglio della RNasi P batterica e della RNasi MRP eucariotica hanno posizione simile: questi dati hanno permesso di ipotizzare che il gene per l'RNA MRP derivi dalla duplicazione del gene per l'RNA della RNasi P (Morrissey and Tollervey, 1995; Xiao *et al.*, 2002).

Si ritiene che le snoRNP U3 e U14 siano universali tra gli Eucarioti e che U3 sia la prima a legarsi al pre-rRNA nascente, vicino al 5'; si pensa che questa snoRNP sia fondamentale per la formazione di un complesso attivo necessario per il processamento, di cui fanno parte almeno 40 proteine, lo "small-subunit (SSU) processome" (Dragon *et al.*, 2002; Bernstein *et al.*, 2004). Nei Vertebrati sono presenti altre due snoRNP C/D conservate, necessarie per il processamento del pre-rRNA, U8 e U22.

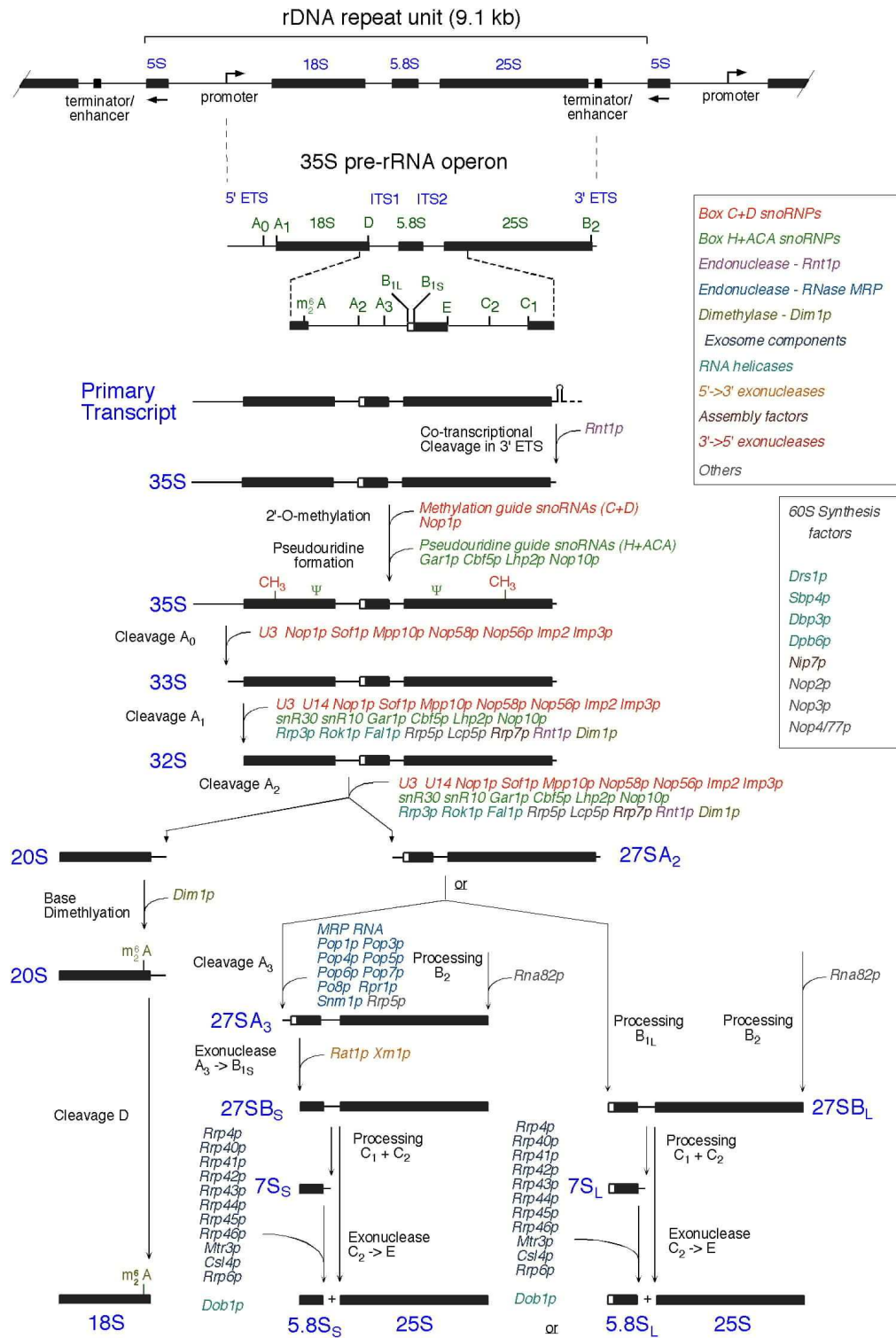


Figura 1.8: in alto è riportata l'organizzazione dei geni per rRNA in *S. cerevisiae*: l'rRNA 5S è trascritto dalla RNA polimerasi III in senso opposto rispetto agli altri rRNA. Sotto è riportato il processamento del pre-rRNA 35S: sono indicate tutte le componenti richieste per ogni singolo passaggio. Il precursore 27SA₂ è tagliato preferenzialmente al sito A₃ (dal sito <http://homepages.ed.ac.uk/dtoller/>).

1.6.2 Modificazioni chimiche dell'rRNA

Come già detto gli snoRNA guidano due tipi di modificazioni chimiche: la metilazione del ribosio e la pseudouridilazione. La metilazione aumenta la superficie idrofobica e blocca gruppi chimici del ribosio che potenzialmente possono formare ponti idrogeno; anche la formazione di pseudouridine altera le possibilità di formazione di legami idrogeno, mediante l'aggiunta di un donatore di protoni. Queste modificazioni avvengono principalmente sul pre-rRNA 35S, subito dopo il taglio all'estremità 3' (Venema and Tollervey, 1999).

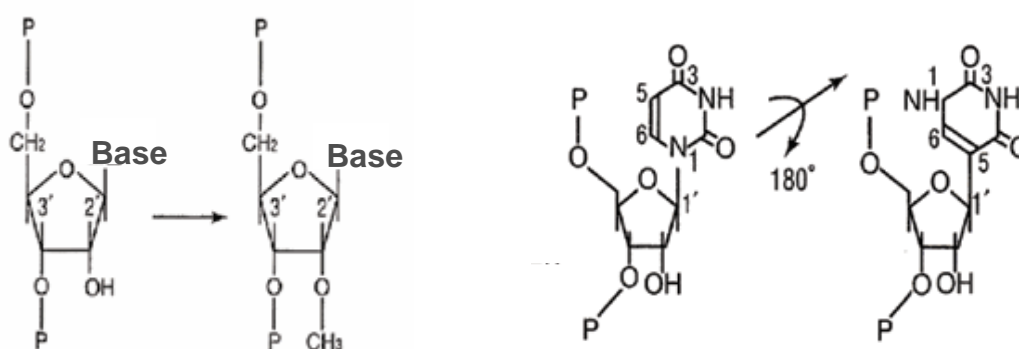


Figura 1.9: sulla sinistra è mostrata la reazione di metilazione, sulla destra la reazione di pseudouridilazione.

Una modificazione, o un gruppo di modificazioni, possono avere effetti su varie caratteristiche dell'rRNA, quali il ripiegamento, la stabilità, l'interazione con proteine ribosomiali. Nessuna singola modificazione risulta essere essenziale, mentre in lievito l'eliminazione globale di un gruppo di modificazioni, conseguente a mutazioni puntiformi nelle componenti enzimatiche, provoca gravi effetti sulla crescita (Tollervey *et al.*, 1993). Mappe tridimensionali mostrano che regioni del ribosoma importanti per la funzione sono ricche in modificazioni, ma lo sono anche regioni non direttamente coinvolte nella traduzione (Decatur and Fournier, 2002); è ragionevole ipotizzare che nucleotidi modificati presenti nel centro con attività peptidil-transferasica del ribosoma siano coinvolti nel riconoscimento e nel legame dell'amminoacil-tRNA, nella reazione di formazione del legame peptidico, nella traslocazione del tRNA dal sito A al sito P e nell'uscita del tRNA dal ribosoma.

In *S. cerevisiae* vengono metilati 55 siti e 44 sono pseudouridilati: gli snoRNA che guidano le pseudouridilazioni sono stati tutti identificati e in totale sono 28 (Schattner *et al.*, 2004; Torchet *et al.*, 2005), mentre il numero di snoRNA C/D box conosciuti è 46: il coinvolgimento nell'attività di metilazione o nel taglio endonucleolitico è stata verificata solo per 43 di questi (Lowe and Eddy, 1999; Davis and Ares, 2006) mentre per uno (snR190) la funzione di guida nella metilazione è stata solo predetta e per gli altri due (snR4 e snR45) la funzione è attualmente ignota (Piekna-Przybylska *et al.*, 2007). Al momento non si è ancora riusciti a localizzare precisamente uno dei siti metilati presenti nella subunità piccola del ribosoma (Lowe and Eddy, 1999), mentre una delle 55 metilazioni (G_m2922 nell'rRNA 25S) è catalizzata da una proteina essenziale, Spb1p. Questa proteina agisce in una fase tardiva del processamento dell'rRNA, durante la maturazione del pre-rRNA 27S, al contrario delle snoRNP; il sito modificato fa parte del sito accettore del tRNA nella regione con attività peptidil-transferasica della subunità grande del ribosoma, è universalmente conservato e sembra svolgere un ruolo preminente nella biogenesi e nella funzione del ribosoma. Inoltre Spb1p è in grado di metilare anche U_m2921 , in mancanza di snR52 (Bonnerot *et al.*, 2003; Lapeyre and Purushothaman, 2004). Anche l'unica pseudouridilazione dell'rRNA 5S attualmente nota ($\psi50$) è catalizzata da una proteina, Pus7p. In *H. sapiens* invece i siti metilati attualmente conosciuti sono 105-107 (Gonzales *et al.*, 2005) e quelli pseudouridilati sono circa 91 (Ofengand and Bakin, 1997).

Anche alcune basi dell'rRNA risultano essere modificate, ma si ritiene che le reazioni siano catalizzate da proteine piuttosto che da complessi RNA-proteina: attualmente, in *S. cerevisiae*, sono note Dim1p (dimetilazione in posizione 6 di A1781 e A1782 nell'rRNA 18S) e Bud23p (metilazione in posizione 7 di G1575 dell'rRNA 18S). La dimetilazione dell'rRNA 18S è una delle poche modificazioni conservate tra batteri ed eucarioti, ma tale modificazione non risulta essere essenziale in lievito, mentre si ipotizza che Dim1p (una proteina essenziale) sia coinvolta in un meccanismo di controllo della biosintesi del ribosoma (Lafontaine *et al.*, 1998). La modificazione delle basi è la sola modificazione che risulti più frequente nei Procarioti (24 su 38 modificazioni totali) che negli Eucarioti (in *S. cerevisiae* 10 su 112 modificazioni totali).

1.6.3 Modificazioni chimiche di altri RNA

Il primo RNA bersaglio degli snoRNA, differente dagli rRNA, che è stato individuato è lo snRNA U6 (Tycowski *et al.*, 1998): questo snRNA è metilato in organismi pluricellulari, ma anche in *S. pombe* e sembra che questa modificazione sia coinvolta nella regolazione fine dello “splicing” (Zhou *et al.*, 2002). Successivamente si è visto che anche gli snRNA U2 e U4 in topo sono modificati (Huttenhofer *et al.*, 2001). Recentemente è stato osservato che in *S. cerevisiae* una delle tre pseudouridilazioni dello snRNA U2 è guidata da snR81 (Ma *et al.*, 2005), che comunque guida anche una modificazione dell’rRNA.

In *C. elegans* sono stati individuati potenziali siti bersaglio in vari tRNA, caso per ora unico tra gli Eucarioti; inoltre è stata osservata la presenza, in alcuni snoRNA, di due elementi antisenso diretti verso molecole di RNA bersaglio con differente localizzazione (ad esempio l’rRNA nel nucleolo e uno snRNA nei “Cajal bodies”) (Zemann *et al.*, 2006).

1.6.4 snoRNA “orfani”

I bersagli di alcuni snoRNA risultano essere ancora sconosciuti e tali snoRNA vengono chiamati “orfani”. Il loro ruolo potrebbe essere legato al taglio dell’rRNA oppure potrebbero guidare svariati tipi di RNA al nucleolo come, ad esempio, alcuni mRNA che vi sono localizzati temporaneamente, o alcuni pre-tRNA, il cui processamento in 5’ avviene nel nucleolo (Bachellerie *et al.*, 2002).

1.7 Rilascio degli snoRNA dal pre-rRNA

Il meccanismo mediante il quale le snoRNP si appaiano e si dissociano dall’rRNA è ancora sconosciuto, ma l’associazione e la rottura dei legami tra snoRNA e rRNA implica cambiamenti sia energetici sia topologici, in particolare nel caso degli snoRNA che guidano metilazioni, per i quali esiste un appaiamento minimo di 12 bp (13 in lievito). Da uno studio su vari organismi sono stati individuati, all’interno dei duplex che si formano tra rRNA e snoRNA C/D box, nove appaiamenti per i quali si osserva fortemente il mantenimento dell’appaiamento mediante mutazioni compensatrici (Chen *et al.*, 2007). In lievito sono state identificate due elicasi DEAD box coinvolte nel rilascio degli snoRNA: Dbp4p (Kos and Tollervey, 2005) e Has1p (Liang and Fournier,

2006). In mancanza di queste proteine gli snoRNA rimangono appaiati all'rRNA, e risultati simili sono stati osservati anche per altre due proteine prive di attività elicastica, Esf1p (Peng *et al.*, 2004) e Esf2p (Hoang *et al.*, 2005), indicando che il processo di rilascio è qualcosa di complesso che richiede l'azione coordinata di svariati fattori.

1.8 snoRNA tessuto-specifici e "imprinting"

Negli Eucarioti superiori sono stati identificati numerosi snoRNA, gran parte dei quali C/D box, la cui espressione è tessuto-specifica: molti di questi sono espressi esclusivamente nel cervello o in particolari regioni di esso (Cavaillé *et al.*, 2000). Una caratteristica dei geni per questi snoRNA è l'essere soggetti al fenomeno detto "imprinting" per cui l'espressione del gene è ristretta all'allele di origine materna o paterna; un'eccezione è costituita dallo snoRNA HBI-36 (in uomo; in topo MBI-36) il cui gene si trova in un introne del gene per il recettore della serotonina 5-HT_{2c}, collocato sul cromosoma X; questo è al momento l'unico snoRNA H/ACA cervello-specifico individuato.

In alcuni casi più copie di geni per uno snoRNA sono ripetute in tandem e risultano essere molto conservate, al contrario delle sequenze circostanti; non è tuttavia noto se tutte le copie vengano espresse (Cavaillé *et al.*, 2000).

La funzione di questi snoRNA non è ancora stata chiarita: forse sono coinvolti nei meccanismi di "imprinting", guidando particolari proteine in regioni cromosomiche soggette a questo fenomeno, oppure possono appaiarsi a RNA bersaglio non ancora definiti (Bachelier *et al.*, 2002). In uno snoRNA (HBII-52) è stato tuttavia individuato un elemento antisense complementare all'mRNA per il recettore della serotonina 2C: pare che questo snoRNA influenzi lo "splicing" alternativo dell'mRNA (Kishore and Stamm, 2006). I geni per questo snoRNA si trovano in una regione genomica coinvolta nelle sindromi di Prader-Willi e di Angelman, ma recentemente è stato escluso un ruolo preminente per questo "cluster" nella sindrome di Prader-Willi (Runte *et al.*, 2005).

1.9 Small Cajal body RNAs (scaRNAs)

I "Cajal bodies" (CB) sono organelli nucleari spesso localizzati vicino ai nucleoli, coinvolti nella biogenesi di snRNA e snoRNA; non sono presenti in *S. cerevisiae*.

Particelle altamente correlate alle snoRNP, le scaRNP, risiedono in queste strutture, dove sono coinvolte nella modificazione degli snRNA. Il primo scaRNA individuato è stato U85 (Jady and Kiss, 2001), che presenta sia le box C/D che H/ACA e dirige la metilazione e la pseudouridilazione dello snRNA U5. Tutti gli scaRNA identificati finora possiedono le box di una classe e/o dell'altra e inoltre, nelle molecole che presentano le box H/ACA, sono presenti due copie di un segnale di localizzazione nei "Cajal bodies", le "CAB box" (consenso ugAG): le box funzionano in modo additivo dal momento che la distruzione di una provoca un parziale rilascio verso il nucleolo; è inoltre importante il contesto strutturale delle box e l'integrità delle box H e ACA. Probabilmente le CAB box costituiscono un sito di legame per proteine che mediano la ritenzione dell'RNA nei "Cajal bodies" (Henras *et al.*, 2004); non è chiaro come gli scaRNA C/D box, che mancano delle CAB box, siano trattenuti nei "Cajal bodies".

1.10 Gli sRNA degli Archaea

Negli *Archaea*, gli sRNA sono generalmente più piccoli rispetto agli snoRNA eucariotici. Gli sRNA C/D box possiedono tipicamente sia la box C' che D': entrambe differiscono di poco rispetto alle box C e D; questi sRNA si associano a tre tipi di "core proteins": la proteina ribosomiale L7Ae (omologo di Snu13p), la fibrillarina e Nop56/58 (omologo sia di Nop56p che di Nop58p). A differenza della snoRNP eucariotica, la sRNP degli *Archaea* è simmetrica; la fibrillarina si lega allo sRNA mediante Nop56/58.

In *P. furiosus* è stata rilevata la presenza di sRNA C/D box, legati alle proteine caratteristiche di questa classe, ma circolari (Starostina *et al.*, 2004); RNA circolari sono già stati descritti in altri organismi, ma in nessuno di questi casi è nota una funzione in processi cellulari (se si escludono naturalmente i genomi a RNA di alcuni virus). Le sRNP C/D box possono metilare gli rRNA ma anche tRNA: in *H. volcanii* lo sRNA che guida la metilazione di due nucleotidi nel pre-tRNA^{Trp} (modificazione necessaria per lo "splicing") è contenuto nell'introne dello stesso tRNA ed agisce in *trans*, cioè la metilazione è guidata da una sRNP già esistente e autonoma (Singh *et al.*, 2004); inoltre è stato dimostrato che la funzione guida può essere svolta dall'introne exciso sia nella forma lineare che circolare.

Gli sRNA H/ACA box possiedono una unica forcina seguita dalla box ACA e lo “stem” superiore forma un motivo “K-turn” non presente negli Eucarioti; le “core proteins” legate sono quattro: Cbf5 (che pare prendere contatto direttamente con lo sRNA), L7Ae (omologo di Nhp2p), Gar1 e Nop10. Ad oggi sono pochi gli sRNA H/ACA box identificati, in accordo con il fatto che, negli *Archaea*, il contenuto di pseudouridine nell’rRNA è basso.

Per quanto riguarda l’organizzazione genomica, i geni per sRNA C/D box sono prevalentemente localizzati in regioni intergeniche e in alcuni casi si sovrappongono al 3’ o al 5’ delle ORF fiancheggianti (Bachelierie *et al.*, 2002); non sono ancora stati individuati promotori indipendenti.

Per il fatto che proteina L7Ae faccia parte sia delle sRNP sia del ribosoma (e anche per il fatto che Snu13p e Nhp2p di lievito presentano somiglianza con una proteina ribosomiale), è stato ipotizzato che i complessi che guidano modificazioni su nucleotidi grazie a RNA guida si siano evoluti dall’apparato di traduzione stesso (Kiss, 2002).

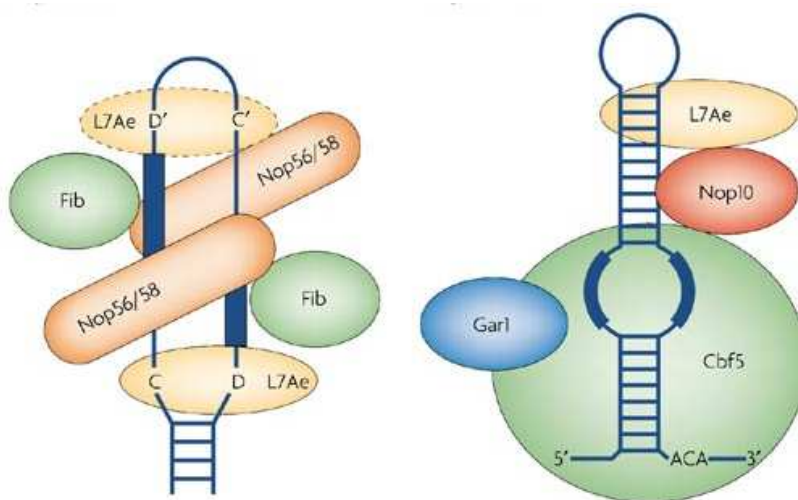


Figura 1.10: le sRNP C/D box (sulla sinistra) e H/ACA box (sulla destra).

1.11 Organizzazione genomica dei geni per snoRNA

Inizialmente sono stati individuati tre diversi tipi di organizzazione genica: l’unità trascrizionale indipendente monocistronica, l’unità trascrizionale policistronica e l’unità codificante intronica (figura 1.11). La prima è l’organizzazione più comune in alcuni Eucarioti inferiori come i lieviti; i “clusters” policistronici sono frequenti nelle piante ma anche in organismi quali *D. discoideum* (Aspegren *et al.*, 2004), mentre le

unità introniche rappresentano l'organizzazione tipica degli Eucarioti superiori, compresa *Drosophila*.

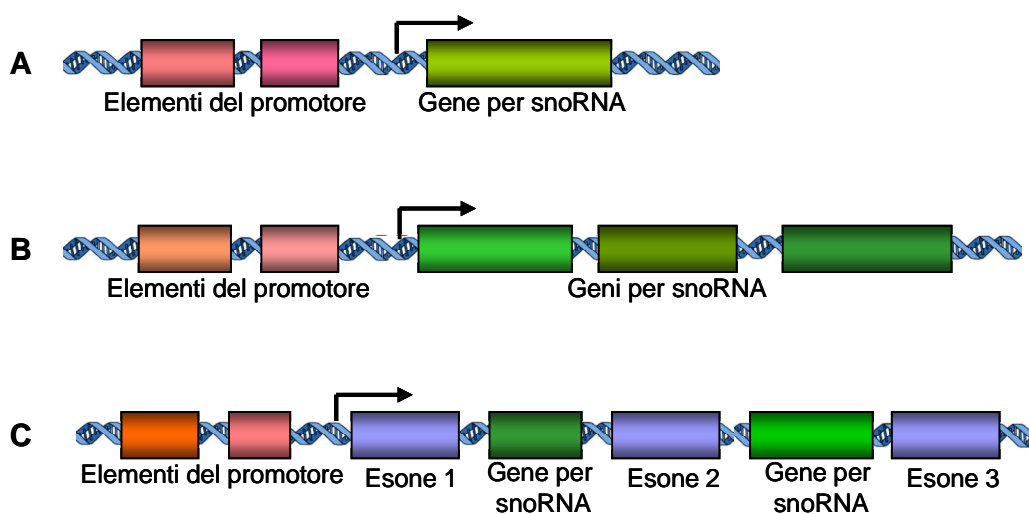


Figura 1.11: organizzazione dei geni per snoRNA. **A:** unità indipendente; **B:** unità policistronica; **C:** unità introniche.

Nel caso delle unità indipendenti un singolo gene, posto in una regione intergenica, è sotto il controllo un proprio promotore; nei “clusters” policistronici invece più geni sono sotto il controllo dello stesso promotore: la trascrizione dà luogo ad un unico RNA da cui poi vengono rilasciati i singoli snoRNA (vedi paragrafo 1.12). La trascrizione dei geni intronici è invece sotto il controllo del promotore del gene ospite, anche se in *C. elegans* è stata ipotizzata la trascrizione indipendente per alcuni geni intronici (vedi paragrafo 1.12) (He *et al.*, 2006); i geni ospiti sono frequentemente geni per proteine ribosomiali o comunque geni i cui prodotti sono coinvolti nella traduzione. Un gene ospite può contenere un unico gene per snoRNA oppure un gene in quasi tutti gli introni; mediante analisi comparativa nei nematodi *C. elegans* e *C. briggsae* è stato osservato come alcuni geni per snoRNA H/ACA presenti in introni vicini siano stati originati mediante un processo di *cis*-duplicazione (duplicazione e inserzione vicino al locus di origine) (Zemann *et al.*, 2006). Nei Mammiferi e in *Drosophila* (Huang *et al.*, 2005) sono stati individuati alcuni geni nei cui introni sono contenuti singoli geni per snoRNA, ma sembra improbabile che gli esoni possano codificare per una proteina: questi geni sono chiamati *UHG* (“U22 host gene”, dal momento che il primo gene di questo tipo individuato è stato il gene ospite di U22 in *H. sapiens*) (Tycowski *et al.*,

1994); in *Drosophila* i geni per snoRNA C/D box sono principalmente contenuti in geni ospiti di questo tipo (Huang *et al.*, 2005). Gli esoni di questi geni sono poco conservati, ad eccezione di quelli del gene ospite *U87HG*, cosa che suggerisce altre funzioni oltre alla produzione di U87 (Marakova and Kramerov, 2005).

È stato osservato come alcuni geni ospiti, sia codificanti che non, siano sintenici in uomo e topo (cioè disposti, insieme ad altri geni ortologhi, nello stesso ordine su porzioni equivalenti di cromosomi), e ciò rafforza l'idea che anche i geni non codificanti per proteine derivino da un gene ancestrale comune (Tanaka-Fujita *et al.*, 2007).

Esaminando le sequenze attorno a paraloghi di geni per snoRNA recentemente identificati in genomi di mammiferi (spesso localizzati all'interno di introni, in orientamento "senso"), è stato possibile ipotizzare che questi geni siano parte di retrotrasposoni (Weber, 2006; Luo and Li, 2007): la retrotrasposizione è in grado di creare copie funzionali del gene originale, se l'inserzione avviene all'interno di un introne, con orientamento "senso", in modo da poter sfruttare il promotore del gene ospite per l'espressione; se il gene originale continua ad essere funzionale, le nuove copie possono anche dare origine, durante l'evoluzione, a geni codificanti per snoRNA con bersagli differenti da quello iniziale (Weber, 2006). Non è attualmente noto se tutte le copie potenzialmente funzionali siano realmente espresse nei tessuti. Per la retrotrasposizione, almeno in alcuni casi, sembra che venga sfruttato il macchinario di integrazione utilizzato da altri tipi di retrotrasposoni (Luo and Li, 2007).

Recentemente sono state osservate nuove varianti nell'organizzazione dei geni per snoRNA: nelle piante *Arabidopsis thaliana* e *Oryza sativa* sono stati individuati geni dicistronici tRNA-snoRNA (figura 1.12 A), il cui trascritto è successivamente processato, probabilmente dalla RNasi Z (Kruszka *et al.*, 2003), coinvolta nella maturazione dell'estremità 3' dei tRNA.

Un'altra tipologia è rappresentata dal "cluster" intronico (figura 1.12 B) che è stato scoperto in *O. sativa* (Liang *et al.*, 2002), dove è un tipo di organizzazione molto frequente tra i geni per snoRNA e spesso i "clusters" sono formati da copie multiple della stessa specie di gene per snoRNA; in altre piante come *A. thaliana* è invece una tipologia rara (tre "clusters" su 35 identificati). Degno di nota è il fatto che nella sottospecie *O. sativa* L. ssp. *japonica* è stato identificato un "cluster" composto da 42

geni il quale deriva chiaramente da cinque duplicazioni in tandem di un “cluster” singolo presente in un’altra sottospecie, *O. sativa* L. ssp. *indica* (Chen *et al.*, 2003).

Il “cluster” intronico è il tipo di organizzazione principale degli snoRNA H/ACA box in *Drosophila* (Huang *et al.*, 2005), i quali presentano molte più isoforme rispetto agli snoRNA C/D box. Ad oggi solo in *Drosophila* è stata osservata una così netta diversità nell’organizzazione genomica delle due famiglie. *Drosophila* è, per il momento, l’unico caso tra i Metazoi in cui è stata osservata la presenza di “clusters” policistronici.

In alcuni organismi (*C.elegans*, *H. sapiens*, *M. musculus*) sono stati identificati alcuni geni intronici il cui orientamento è però contrario al senso di trascrizione (definiti “intron antisense”): un esempio è il gene per U99 che è ospitato nel terzo introne del gene (*C11orf48*) per una proteina ipotetica (MGC2477) sia in uomo che in topo. U99 non viene immunoprecipitato con anticorpi che riconoscono in modo specifico il TMG “cap” tipico degli snoRNA trascritti da unità indipendenti e questo rafforza l’idea che tale snoRNA derivi dal processamento del trascritto antisenso di *C11orf48* (Vitali *et al.*, 2003). Ad oggi, il significato biologico di questa organizzazione genomica non è stato spiegato; tuttavia recenti osservazioni effettuate sul gene per U99 portano a ipotizzare la presenza di un gene ospite non codificante per proteine (*U99HG*) all’interno di *C11orf48*, e disposto con orientamento antisenso rispetto a quest’ultimo (Weber, 2006).

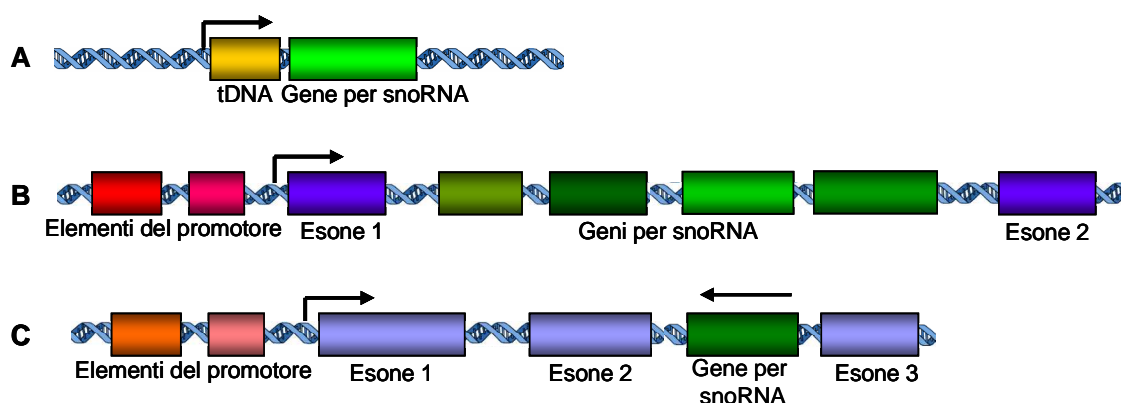


Figura 1.12: A: unità dicistronica tRNA-snoRNA; B: “cluster” intronico; C: gene “intron antisense”.

Per quanto riguarda i Tripanosomatidi, i geni per snoRNA sono organizzati in “clusters” misti (figura 1.13), ossia i gruppi comprendono sia geni per snoRNA C/D box che H/ACA box (Liang *et al.*, 2004).

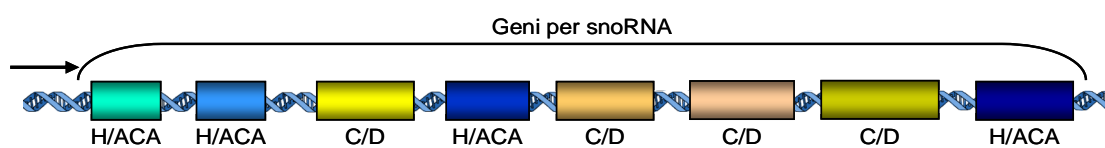


Figura 1.13: “cluster” misto tipico dei Tripanosomatidi.

Organismi	snoRNA C/D box	snoRNA H/ACA box
lieviti	indipendenti	indipendenti
piante	clusters indipendenti o intronici	clusters indipendenti o intronici
<i>C. elegans</i>	intronici	intronici
<i>D. melanogaster</i>	intronici	clusters intronici
mammiferi	intronici	intronici
tripanosomatidi	clusters misti	clusters misti

Tabella 1.1: principali tipi di organizzazione dei geni per snoRNA in diversi organismi.

1.12 Trascrizione degli snoRNA

La trascrizione degli snoRNA avviene nel nucleoplasma, ad opera della RNA polimerasi II o, in alcuni casi, della RNA polimerasi III.

Ad eccezione dei Tripanosomatidi, nel caso della trascrizione da parte della RNA polimerasi II, i geni localizzati in regioni intergeniche possiedono promotori propri: spesso è presente la TATA box e, in *S. cerevisiae*, sono state individuate una o due sequenze consenso per Rap1p a monte di alcuni geni, nonché regioni “A/T rich” e/o siti di legame per Abf1p (Qu *et al.*, 1999). Sia Rap1p che Abf1p sono proteine leganti il DNA coinvolte in svariati processi, tra cui la regolazione della trascrizione dei geni per proteine ribosomiali.

Nelle piante, oltre alla TATA box, sono presenti degli “upsteam sequence elements” (USE) (Waibel and Filipowicz, 1990), mentre nei Metazoi sono presenti “proximal sequence elements” (PSE) e “distal sequence elements” (DSE), così definiti in base alla posizione rispetto alla sequenza codificante (Brown *et al.*, 2003).

Per quanto riguarda i geni intronici, la regolazione della trascrizione dipende dagli elementi presenti nella regione a monte del gene ospite: la trascrizione di diversi geni ospiti, compresi quelli non codificanti per proteine, inizia con un residuo di citosina

seguito da un tratto di oligopirimidine: questa è una caratteristica della famiglia 5'-TOP ("5'-terminal oligopyrimidine"), che include geni per proteine ribosomiali e per alcune altre importanti proteine "housekeeping". La sequenza 5'-TOP sembra avere un ruolo nel coordinare la traduzione dell'mRNA in risposta alle condizioni di crescita (Pelczar and Filipowicz, 1998) e determina un preciso rapporto tra la produzione di mRNA "spliced" e snoRNA, rilasciato sia mediante "splicing" che mediante processamento (vedi paragrafo 1.13); la sequenza TOP potrebbe reclutare dei fattori che controllano in modo specifico il tipo di processamento post-trascrizionale a cui deve andare incontro l'mRNA, in quanto promotori pol II privi della sequenza TOP non sono compatibili con il rilascio dello snoRNA tramite la via del processamento, con un conseguente basso rapporto snoRNA/mRNA (De Turris *et al.*, 2004).

In *C. elegans* è stato individuato un tratto ricco in pirimidine attorno all'ipotetico sito d'inizio della trascrizione di molti geni per proteine ribosomiali ospiti (He *et al.*, 2006). Nello stesso organismo, a monte di alcuni geni intronici, sono stati individuati dei motivi di sequenza che potrebbero agire come promotori indipendenti (Deng *et al.*, 2006; He *et al.*, 2006), cosa che potrebbe giustificare l'elevato livello di espressione di alcuni geni intronici rispetto al gene ospite osservato in alcuni casi; tuttavia è anche possibile che la differente espressione sia dovuta alla maggiore stabilità degli snoRNA, protetti all'interno della snoRNP, rispetto agli mRNA (He *et al.*, 2006).

In topo, nel caso di snoRNA orfani, è stato scoperto che i geni ospiti non possiedono la struttura 5'-TOP e quindi seguono una via di regolazione diversa; probabilmente la funzione di questi snoRNA non è legata alla sintesi dei ribosomi. Tali geni sono tutti codificanti per proteine, tranne *Irm* (Xiao *et al.*, 2006).

La terminazione della trascrizione (da parte della polimerasi II) dei geni per snoRNA indipendenti differisce da quella dei geni per mRNA: per questi ultimi l'estremità 3' viene generata mediante un taglio endonucleolitico cotrascrizionale, seguito dalla poliadenilazione, da parte di Pap1p, dell'estremità così formata (Zhao *et al.*, 1999); nel caso degli snoRNA e degli snRNA i due eventi sono separati (Fatica *et al.*, 2000). È probabile che i geni per snoRNA abbiano, come i geni per mRNA, più siti di terminazione (Kim *et al.*, 2006). La formazione dell'estremità 3', dipendente o indipendente dalla poliadenilazione richiede una serie di fattori *trans*-agenti in parte comuni (Morlando *et al.*, 2002); per quanto riguarda la terminazione della trascrizione

dei geni per snoRNA e snRNA risultano particolarmente importanti Pti1p e Ref2p, che evitano il “read-through” nei geni a valle e si pensa che possano avere una funzione nel separare taglio e poliadenilazione (Dheur *et al.*, 2003). Per quanto riguarda i geni per snoRNA C/D box, è stato osservato che Nop1p interagisce con Ref2p e ne favorisce il legame alla regione codificante (Morlando *et al.*, 2004). Così come Nop1p sembra richiesta per il corretto processamento in 3' degli snoRNA C/D box, Cbf5p sembra coinvolta nella corretta terminazione e nella formazione dell'estremità 3' nel trascritto primario degli snoRNA H/ACA box (Ballarino *et al.*, 2005). Inoltre per la formazione dell'estremità 3' sono richieste le proteine leganti l'RNA Nrd1p e Nab3p e la presunta elicasi Sen1p: Nrd1p può interagire con Nab3p, con il dominio C-terminale della RNA polimerasi II (CTD) e con la CTD-chinasi Ctk1p (Steinmetz *et al.*, 2001). È stato osservato come Nrd1p stimoli la degradazione dell'RNA, *in vitro*, da parte dell'esosoma, con cui risulta interagire: si pensa che questa associazione possa legare la terminazione e il processamento al 3' (Vasiljeva and Buratowski, 2006). Il legame del complesso Nrd1p-Nab3p al terminatore è cruciale nella terminazione di trascritti privi della coda poli(A) e si sono riscontrati siti di legame multipli per queste proteine in parecchi terminatori di geni per snoRNA (Carroll *et al.*, 2007).

Gli snoRNA trascritti a partire da unità indipendenti, presentano all'estremità 5' un “cap” differente da quello degli mRNA: il “cap” 7-metilguanossina tipico di questi ultimi è ipermetilato per dare origine al “cap” 2-2-7-trimetilguanossina (TMG). La reazione avviene nel nucleo e più precisamente nei “Cajal bodies” in uomo e nei “nucleolar bodies” (strutture localizzate nel nucleolo) in lievito; l'enzima che catalizza l'ipermetilazione del “cap” in lievito, Tgs1p, non interagisce direttamente con l'RNA ma con le regioni basiche C-terminali di Cbf5p, Nop56p e Nop58p (Mouaikel *et al.*, 2002; Verheggen *et al.*, 2002). Nel caso di U3 in lievito, il “cap” viene ipermetilato solo quando l'estremità 3' è stata processata (Kufel *et al.*, 2000), mentre nelle piante superiori, essendo U3 trascritto da pol III, il “cap” ha una struttura differente dal trimetilguanossina; in alcuni casi è stata dimostrata la presenza di un “cap” metil-pppA (Shimba *et al.*, 1992).

Attualmente sono noti diversi geni per snoRNA trascritti dalla RNA polimerasi III: *SNR52* in *S. cerevisiae* (Harismendy *et al.*, 2003; Moqtaderi and Struhl, 2003), i geni per U3 nelle piante e nell'alga *C. reinhardtii* (Kiss *et al.*, 1991; Antal *et al.*, 2000), i

geni per snoRNA facenti parte delle unità dicistrioniche tRNA-snoRNA (Kruszka *et al.*, 2003), almeno due geni in *Drosophila* (Isogai *et al.*, 2007) e l'RNA della RNasi MRP nei Mammiferi. Anche in *C. elegans* sono stati individuati alcuni geni che potrebbero essere trascritti dalla RNA polimerasi III, in quanto presentano a monte degli elementi conservati simili alle box A e B; forse questi geni fanno parte di una unità dicistrionica tRNA-snoRNA oppure costituiscono una unità indipendente in cui l'antico promotore interno di un tDNA è diventato il promotore extragenico di un gene per snoRNA presente a valle (Deng *et al.*, 2006). Per quanto riguarda *SNR52*, sono presenti una A e una B box, localizzate nella sequenza "leader", e una TATA box a monte (Harismendy *et al.*, 2003). Nel caso dei dicistroni tRNA-snoRNA il promotore intragenico è presente nella porzione codificante per il tRNA. In *Drosophila*, dei due geni chiaramente identificati come trascritti dalla RNA polimerasi III, uno sembra presentare un promotore di tipo esclusivamente intragenico mentre nell'altro sembrano importanti sia elementi intragenici che extragenici (Isogai *et al.*, 2007).

Per quanto riguarda i Tripanosomatidi, non sono stati individuati dei promotori e la regolazione è post-trascrizionale; tuttavia è stato individuato un elemento di 700 bp, posto tra un "cluster" di snoRNA e la ORF a monte, il quale potrebbe funzionare come sequenza "promoter-like", in grado di aumentare l'espressione, in modo dipendente dall'orientamento. Questo elemento non può essere definito un vero e proprio promotore, in quanto viene trascritto, e la sua funzione potrebbe essere quella di mantenere i livelli di trascrizione lungo il cromosoma oppure di aumentare rapidamente la trascrizione in modo da sostenere la crescita (Liang *et al.*, 2004). È stato dimostrato che i geni per snoRNA nei Tripanosomatidi possono essere silenziati (Liang *et al.*, 2003).

1.13 Processamento e maturazione degli snoRNA

Una volta trascritti, gli snoRNA possono essere processati per ottenere la forma matura; entrambe le estremità dei trascritti primari possono essere il bersaglio di nucleasi (figura 1.14).

In *S. cerevisiae*, i trascritti codificati da unità monocistrioniche indipendenti possono subire un taglio in 5' da parte della endonucleasi Rnt1p; questo crea un sito per

l'ingresso delle 5'-esonucleasi che probabilmente sono Xrn1p e/o Rat1p (Chanfreau *et al.*, 1998).

Rnt1p è l'omologo della RNasi III batterica, coinvolta nella maturazione del pre-rRNA. Questa endonucleasi taglia in modo specifico l'RNA a doppio filamento che presenta un "tetraloop" con sequenza consenso AGNN ("G2 loop"), a 13-16 bp di distanza; recentemente è stato osservato che anche il "tetraloop" con sequenza AAGU ("A1 loop") può essere presente vicino al sito di taglio: in questo caso Rnt1p usa una diversa serie di interazioni per riconoscere il substrato (Ghazal *et al.*, 2006; Gaudin *et al.*, 2006). Rnt1p può tagliare anche a valle dello snoRNA, se è presente una struttura riconosciuta come substrato.

Come già detto nel paragrafo 1.5.2, Rnt1p è richiesta per l'importazione nel nucleo di Gar1p, probabilmente in forma di complesso Rnt1p-Gar1p; la dissociazione tra le due proteine avviene forse nel momento in cui Rnt1p incontra l'RNA substrato. La mancata interazione tra le due proteine riduce, ma non blocca, il processamento e la pseudouridilazione dell'rRNA, mentre la maturazione degli snoRNA non ne risente in modo significativo (Tremblay *et al.*, 2002).

L'interazione di Rnt1p con Nop1p è invece richiesta nel processamento dello snoRNA intronico U18 (Giorgi *et al.*, 2001).

La 5'-esonucleasi Rat1p è localizzata nel nucleo mentre Xrn1p è prevalentemente citoplasmatica, tuttavia per ottenere la completa inibizione dell'attività 5'-esonucleolitica nucleare è necessario inattivare entrambe le proteine, probabilmente a causa di piccola frazione di Xrn1p presente nel nucleo. In caso di mancato taglio da parte di Rnt1p, le esonucleasi possono comunque processare i trascritti primari (Lee *et al.*, 2003).

Gli snoRNA che non subiscono processamento in 5' mantengono il "TMG cap".

Rnt1p è richiesta anche per il rilascio dei singoli snoRNA dai trascritti policistronici, e a questo segue l'intervento delle esonucleasi (Chanfreau *et al.*, 1998; Petfalski *et al.*, 1998).

Il processamento in 3' richiede l'intervento di un complesso costituito da varie 3'-esonucleasi, l'esosoma. Esistono due forme di questo complesso, una nucleare e una citoplasmatica: la prima contiene due subunità specifiche e non essenziali, Rrp6p e Lrp1p, mentre le subunità comuni (tutte essenziali) sono Rrp41p, Rrp42p, Rrp43p,

Rrp45p, Rrp46p, Mtr3p, Rrp44p/Dis3p, Rrp4p, Rrp40p e Csl4p. La controparte umana dell'esosoma di lievito è il complesso PM/Scl (“polymyositis–scleroderma”) (Brouwer *et al.*, 2000). Sembra che il taglio finale, con cui si ottiene l'estremità 3' degli snoRNA maturi, sia completamente dipendente da Rrp6p (Allmang *et al.*, 1999) e in assenza di questa nucleasi si può osservare l'accumulo di precursori estesi in 3' e poliadenilati (van Hoof *et al.*, 2000); la poliadenilazione di questi RNA non codificanti non è dovuta a Pap1p, ma a due poli(A)-polimerasi di recente identificazione, Trf4p e Trf5p. Entrambe fanno parte del complesso TRAMP (Trf4p/Air2p/Mtr4p polyadenylation complex), richiesto per il processamento e/o la degradazione dei substrati da parte dell'esosoma (LaCava *et al.*, 2005). La delezione contemporanea di *TRF4* e *RRP6* abolisce l'accumulo di precursori poliadenilati (Wyers *et al.*, 2005), per cui si può dedurre che il complesso TRAMP agisca a monte di Rrp6p. In ceppi *rrp6Δ* si è osservato l'accumulo dei precursori poliadenilati in una zona discreta del nucleolo, distinta dai “nucleolar bodies”, che è stata chiamata “nucleolar poly(A) domain” e per la cui formazione è richiesta anche la presenza del complesso TRAMP. Si è ipotizzato che il raggruppamento degli RNA poliadenilati possa promuovere il riconoscimento e la degradazione da parte dell'esosoma (Carneiro *et al.*, 2007).

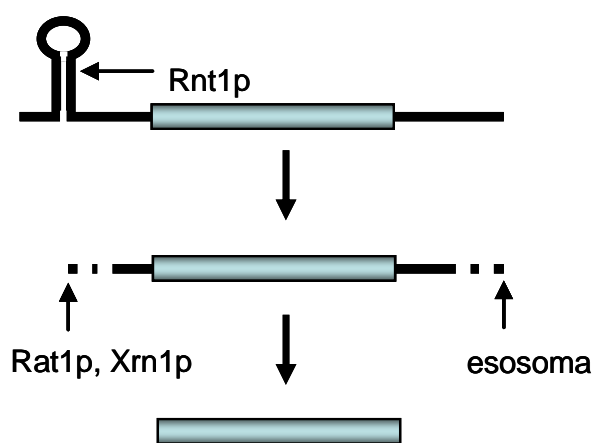


Figura 1.14: maturazione di uno snoRNA monocistronico; in azzurro è rappresentato lo snoRNA maturo.

Il processamento di U3 differisce in parte da quello degli altri snoRNA di lievito: il primo taglio in 3' è effettuato da Rnt1p, mentre il 5' rimane intatto. I tratti ricchi in U presenti nella porzione terminale sono legati da Lhp1p, una proteina richiesta per il

processamento dei tRNA: questa impedisce all'esosoma di proseguire nel taglio del 3'. A questo punto avviene lo "splicing" dell'introne, successivamente Lhp1p viene spiazzata da Nop1p, Nop56p e Nop58p e viene completato il taglio da parte dell'esosoma; nel caso di U3 il taglio finale al 3' non è effettuato da Rrp6p. Come ultimo passaggio il "cap" viene trimetilato (Kufel *et al.*, 2000).

Gli snoRNA presenti negli introni devono essere rilasciati e questo può avvenire secondo due modalità: la principale richiede il taglio e la deramificazione dell'introne, sul quale poi intervengono le esonucleasi (Tycowski *et al.*, 1993). La via alternativa invece prevede l'azione di un'endonucleasi sul pre-mRNA e il successivo intervento delle esonucleasi (Fragapane *et al.*, 1993) (figura 1.15).

Nel primo caso è molto importante, almeno per gli snoRNA C/D box, la posizione dello snoRNA all'interno dell'introne: nel caso di U24 è cruciale la distanza tra l'estremità 3' dello snoRNA e il punto di ramificazione, che non deve essere inferiore a 44 nt, ma neppure troppo ampia (Vincenti *et al.*, 2007); variazioni nella distanza di cui sopra, non solo influenzano la sintesi dello snoRNA, ma anche lo "splicing".

Nel caso di U18 è possibile anche la produzione dello snoRNA attraverso la via alternativa, che è responsabile di circa il 30% della quantità di tale snoRNA. Questa seconda possibilità dipende dalla presenza di uno "stem" nell'introne, appena al di fuori della sequenza codificante di U18; questa struttura è conservata in parecchi introni, dal lievito fino ai mammiferi, e ha un ruolo nel favorire lo "splicing" e la sintesi della snoRNP (Villa *et al.*, 2000; Giorgi *et al.*, 2001), oltre a essere il bersaglio di una o più proteine responsabili del taglio del pre-mRNA necessario alla produzione di U18 (Giorgi *et al.*, 2001). Tale "stem" solitamente è presente nel caso in cui lo snoRNA non sia correttamente posizionato all'interno dell'introne.

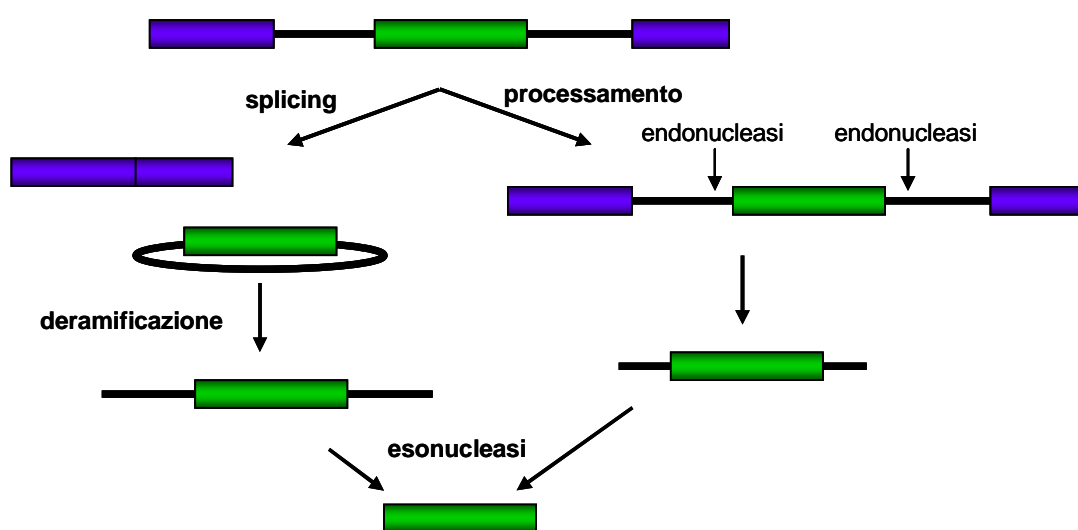


Figura 1.15: biosintesi di uno snoRNA intronico, secondo le due vie possibili (“splicing” o processamento). In viola sono rappresentati gli esoni del gene ospite, in verde lo snoRNA.

Anche nel caso dei geni per snoRNA C/D box di mammifero, la biosintesi è stata messa in relazione alla posizione all’interno dell’introne. È stato osservato come la deplezione di una proteina necessaria per lo “splicing” nei vertebrati, IBP160, provochi difetti nella formazione della snoRNP C/D box *in vitro*; questa proteina si lega in modo specifico ad una distanza fissa dal sito di ramificazione e si può ipotizzare che i geni per snoRNA si siano “spostati” nella posizione in cui IBP160 riesce a promuovere l’associazione con le proteine della snoRNP (Hirose *et al.*, 2006). Per quanto riguarda gli snoRNA H/ACA box umani, pare che non esista una localizzazione preferenziale nell’introne e la sintesi è indipendente dallo “splicing” (Richard *et al.*, 2006).

1.14 Assemblaggio della snoRNP

Le snoRNP di entrambe le classi vengono assemblate co-trascrizionalmente. Per quanto riguarda gli snoRNA C/D box, è ormai certo che la proteina che si lega per prima è Snu13p: essa riconosce e lega direttamente il motivo di struttura “K-turn” formato da appaiamenti non canonici nel motivo C/D (Watkins *et al.*, 2000), inducendo dei cambiamenti conformazionali nell’RNA. Risultati da esperimenti *in vitro* sembrano dimostrare che questa proteina è necessaria per il legame delle altre proteine “core”. Nop58p è in grado di legarsi direttamente allo snoRNA, interagendo con la box C e in

particolare, come risulta da esperimenti di “cross-linking” in oociti di *Xenopus*, con la seconda U della box (Cahill *et al.*, 2002). Nop1p è in grado di legarsi direttamente sia alla box D che D', ma risultano anche interazioni con la box C', anche se mutazioni nelle box D' e C' non sembrano influenzare le interazioni con Nop1p (Cahill *et al.*, 2002). Numerosi dati sembrano indicare l'associazione di Nop1p con i precursori degli snoRNA C/D box (Watkins *et al.*, 2002; Giorgi *et al.*, 2001), ma anche con i geni per snoRNA attivamente trascritti (Morlando *et al.*, 2004). Nop56p può legarsi allo snoRNA solo se è presente Nop1p ma comunque risulta interagire con le due U della box C' (con le quali interagisce, peraltro, anche Nop1p) (Cahill *et al.*, 2002).

Si pensa che gli snoRNA H/ACA box vengano legati da un complesso preformato, costituito da Nhp2p, Nop10p e Cbf5p, in quanto solo queste proteine sono richieste per la stabilità metabolica dell'RNA; in particolare è stato osservato che Cbf5p e Nhp2p sono associati ai geni per snoRNA H/ACA attivamente trascritti (Ballarino *et al.*, 2005).

1.14.1 Fattori richiesti per l'assemblaggio

Molte delle proteine coinvolte nell'assemblaggio della snoRNP sono necessarie per entrambe le classi. Le più conservate, presenti negli *Archaea*, in lievito e nell'uomo, sono le proteine p50/p55 (Rvb1p e Rvb2p in *S. cerevisiae*, Tip48 e Tip49 in uomo) che hanno attività DNA elicastica e sono coinvolte in numerosi processi tra cui riparo del DNA e trascrizione; non fanno parte della snoRNP matura e non sono localizzate nel nucleolo, ma sono richieste per la biogenesi delle due classi di snoRNA, con un ruolo post-trascrizionale (King *et al.*, 2001). Sembra che, nel caso della snoRNP C/D box, il complesso si formi nel momento in cui Nop58p si lega allo snoRNA.

Un'altra proteina richiesta per l'assemblaggio in entrambe le classi è SMN (“Survival of Motor Neuron protein”), conservata in invertebrati, vertebrati e in *S. pombe* ma assente in *S. cerevisiae*; tale proteina è presente in strutture nucleari dette “gems” che spesso si sovrappongono ai “Cajal bodies” e, da esperimenti *in vitro*, si è visto che lega sia la fibrillarina sia Gar1p, attraverso i loro domini contenenti le ripetizioni RG (Jones *et al.*, 2001; Pellizzoni *et al.*, 2001), e probabilmente funziona come piattaforma per permettere il legame delle proteine all'RNA. SMN è essenziale anche per

l'assemblaggio delle snRNP nel citoplasma ed è richiesta per la trascrizione e lo "splicing" del pre-mRNA (Pellizzoni *et al.*, 1998).

Essenziali per la biogenesi delle sole snoRNP H/ACA box sono invece Naf1p e Shq1p, che formano un complesso e interagiscono con Nhp2p e Cbf5p; Naf1p è in grado di legare il dominio carbossiterminale (CTD) fosforilato della RNA polimerasi II e lega l'RNA attraverso un dominio simile a quello di Nrd1p, proteina richiesta per la formazione dell'estremità 3' degli snoRNA in lievito (Yang *et al.*, 2002). Naf1p e Gar1p interagiscono con lo stesso sito di Cbf5p, in modo da escludersi a vicenda: è possibile che Naf1p permetta l'assemblaggio di una pre-snoRNP stabile, che rimane inattiva finché Gar1p non prende il posto di Naf1p (Yang *et al.*, 2005; Matera *et al.*, 2007). Inoltre questa proteina interagisce con Spt16p, Tfg1p e Sub1p, tutte associate alla RNA polimerasi II (Yang *et al.*, 2005).

Per le snoRNP C/D box è invece richiesta Bcd1p, necessaria per l'accumulo degli snoRNA C/D box (Peng *et al.*, 2003); anche questa proteina sembra interagire con l'apparato pol II e si è ipotizzato, in modo analogo a Naf1p, che il suo posto nella snoRNP matura sia preso da un'altra proteina, in questo caso Nop56p (Matera *et al.*, 2007).

1.15 Localizzazione degli snoRNA

Per la localizzazione a livello del nucleolo risultano essere importanti le box C/D e H/ACA; inoltre, la deplezione di ciascuna delle proteine "core" della snoRNP C/D box provoca una parziale localizzazione nucleoplasmatica dello snoRNA.

Nei vertebrati gli snoRNA C/D box transitano nei "Cajal bodies". Nei mammiferi, come in lievito, U3 è sintetizzato come precursore, con un una breve estensione in 3' e un "cap" monometilato. Il precursore lega solo Snu13p/15.5 kDa ed è presente nel sito di trascrizione e nei "Cajal bodies", ma non nei nucleoli, mentre il maturo è localizzato sia nei nucleoli che nei "Cajal bodies" (Verheggen *et al.*, 2002). Una conferma viene anche dagli snoRNA intronici: Snu13p/15.5 kDa è presente inizialmente nelle zone in cui avviene lo "splicing" e nei "Cajal bodies" e solo dopo 30 minuti è presente nei nucleoli, mentre la fibrillarina si distribuisce nei "Cajal bodies" e nei nucleoli, permettendo di ipotizzare che questa proteina si leghi al complesso nei "Cajal bodies" (Leung *et al.*, 2002). Al contrario, gli snoRNA H/ACA box sembrano raggiungere

contemporaneamente i nucleoli e i “Cajal bodies”: questo può essere spiegato se si ipotizza, tra i due compartimenti, un flusso troppo rapido per essere apprezzato, oppure se il “pathway” è più complesso o se non ne esiste un unico (Narayanan *et al.*, 1999).

Come già detto in precedenza, nei “Cajal bodies” avviene anche la formazione del “TMG cap”, mentre in lievito questa reazione avviene all’interno del nucleolo, nei cosiddetti “nucleolar bodies” (Verheggen *et al.*, 2002). Da queste strutture sono esclusi i fattori coinvolti nella biogenesi dell’rRNA e la separazione tra questa e la biogenesi dei piccoli RNA potrebbe essere alla base della formazione dei “Cajal bodies” come strutture indipendenti.

2. SCOPO DELLA RICERCA

Gli snoRNA fanno parte della famiglia degli RNA non codificanti per proteine; uno dei loro ruoli fondamentali è quello di partecipare alla maturazione dell'rRNA.

L'organizzazione genomica dei geni per snoRNA è estremamente variabile tra gli Eucarioti: negli Eucarioti inferiori i geni sono prevalentemente indipendenti e monocistronici, nelle piante si ritrovano soprattutto unità policistroniche, nei mammiferi quasi tutti i geni sono intronici.

La trascrizione degli snoRNA avviene, in tutti gli Eucarioti, principalmente ad opera della RNA polimerasi II, ma la regolazione dell'espressione, gli elementi presenti nella regione del promotore e i fattori richiesti per la completa attivazione della trascrizione dei geni trascritti in modo indipendente sono stati studiati pochissimo in quanto in *H. sapiens* e negli altri mammiferi l'espressione dipende dall'espressione del gene ospite. Ad oggi, quindi, quasi nulla è conosciuto della regione del promotore dei geni per snoRNA in *S. cerevisiae*.

Scopo di questa tesi è individuare, a monte dei geni per snoRNA trascritti in modo indipendente in *S. cerevisiae* (la maggioranza dei geni per snoRNA in questo organismo), gli elementi necessari per la trascrizione. La ricerca viene effettuata sia attraverso una analisi computazionale per individuare gli elementi di sequenza conservati nella regione "upstream" in specie differenti di Saccaromiceti, sia attraverso una analisi *in vivo* in cui, attraverso delezioni e mutazioni puntiformi, si ricercano le porzioni del promotore essenziali per l'espressione.

Al momento non sono note condizioni ambientali che causino una attivazione o una repressione della trascrizione degli snoRNA in lievito: in questa tesi si è anche cercato di identificare particolari condizioni di crescita in cui la trascrizione dei geni per snoRNA possa essere soggetta a regolazione, basandosi su quanto noto a riguardo della regolazione dell'espressione dei geni ospiti di geni per snoRNA e dei geni codificanti per le proteine "core" delle snoRNP in risposta ad alcune variazioni nelle condizioni ambientali.

Al contrario della regolazione della trascrizione, il processamento degli snoRNA è stato studiato in modo estensivo ed è stato dimostrato come il corretto assemblaggio della snoRNP (che necessita, oltre alla presenza delle quattro proteine del "core", delle box C

e D integre) sia essenziale per il corretto processamento in 5' (se richiesto) degli snoRNA trascritti dalla RNA polimerasi II.

Tra i geni per snoRNA trascritti dalla RNA polimerasi III ritroviamo un solo gene di *S. cerevisiae*, *SNR52*. Una parte di questo lavoro di tesi è lo studio del processamento di snR52 al fine di chiarire il coinvolgimento, in questo processo, delle box C e D, mediante mutazioni puntiformi di questi elementi, nonché il ruolo di alcune componenti proteiche della snoRNP, grazie allo spegnimento controllato dei geni codificanti per tali componenti.

Si è inoltre inteso verificare se il complesso 3'-esonucleasico dell'esosoma è coinvolto nella maturazione dell'estremità 3' di snR52, così come avviene per gli altri snoRNA trascritti dalla RNA polimerasi II, nonostante la terminazione della trascrizione da parte della RNA polimerasi III sia differente dalla terminazione da parte della RNA polimerasi II.

3. MATERIALI E METODI

3.1 Ceppi

Per gli esperimenti riguardanti lo studio di snR52, i ceppi utilizzati sono i seguenti

CEPPO	GENOTIPO	RIFERIMENTO
YM4585 (Δ SNR52)	Mat a, his3 Δ 200,lys2-800, leu2-3, 2-112, trp1-901, tyr1-501, URA3 ⁺ , ADE ⁺ , CAN ^S , Δ SNR52-HIS3	Lowe and Eddy, 1999
D254	Mat a, ura3-52, leu2-3, 112, ade1-100, his4-519	Tollervey et al., 1991
D255	Mat a, ura3-52, leu2-3, 112, ade1-100, his4-519, URA3-pGAL10::NOP1	Tollervey et al., 1991
YDL401	Mat a, ara3-52, trp (trp1-D63 o trp1-289), leu2- Δ 1, his3- Δ 200, gal2, gal Δ 208	Lafontaine and Tollervey, 1996
YDL522-20	Mat a, ara3-52, trp (trp1-D63 o trp1-289), leu2- Δ 1, his3- Δ 200, gal2, gal Δ 208, HIS3::GAL10::NOP58	Lafontaine and Tollervey, 1999
BY4741	Mat a, his 3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0	Yeast Genetic Stock Center
YGALSNU13	Mat a, his 3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, KanMX6:P _{GAL1} -SNU13	Dobbyn and O'Keefe, 2004
BY4742	Mat α , his3 Δ 1, lys2 Δ 0, leu2 Δ 0, ura3 Δ 0	Yeast Knock Out Collection (Open Biosystem)
Δ rrp6	Mat α , his3 Δ 1, lys2 Δ 0, leu2 Δ 0, ura3 Δ 0, Δ RRP6-kanMX4	Yeast Knock Out Collection (Open Biosystem)
Δ hpl1	Mat α , his3 Δ 1, lys2 Δ 0, leu2 Δ 0, ura3 Δ 0, Δ LHP1-kanMX4	Yeast Knock Out Collection (Open Biosystem)

Per lo studio dei promotori dei geni per snoRNA sono stati utilizzati i seguenti ceppi

CEPPO	GENOTIPO	RIFERIMENTO
YS669	ade2-101, trp1- Δ 901, ura3-52, leu2-3,112, his3, SNR31::URA3	Balakin et al., 1993
YS670	ade2-101, trp1- Δ 901, ura3-52, leu2-3,112, his3, SNR32::HIS3	Balakin et al., 1993
YM4585 (Δ SNR64)	Mat a, his3 Δ 200,lys2-800, leu2-3, 2-112, trp1-901, tyr1-501, URA3 ⁺ , ADE ⁺ , CAN ^S , Δ SNR64-HIS3	Lowe and Eddy, 1999
YM4585 (Δ SNR65)	Mat a, his3 Δ 200,lys2-800, leu2-3, 2-112, trp1-901, tyr1-501, URA3 ⁺ , ADE ⁺ , CAN ^S , Δ SNR65-HIS3	Lowe and Eddy, 1999

CEPPO	GENOTIPO	RIFERIMENTO
<i>rnt1Δ xrn1Δ</i> <i>rat1-1</i>		Danin-Kreiselman et al., 2003
BY4741	Mat a, his 3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Yeast TAP Fusion Collection (Open Biosystem)
RAP1-TAP	Mat a, his 3Δ1, leu2Δ0, met15Δ0, ura3Δ0, RAP1-TAP-HIS3MX6	Yeast TAP Fusion Collection (Open Biosystem)
REB1-TAP	Mat a, his 3Δ1, leu2Δ0, met15Δ0, ura3Δ0, REB1-TAP-HIS3MX6	Yeast TAP Fusion Collection (Open Biosystem)
HMO1-RAP	Mat a, his 3Δ1, leu2Δ0, met15Δ0, ura3Δ0, HMO1-TAP-HIS3MX6	Yeast TAP Fusion Collection (Open Biosystem)
NRG1-TAP	Mat a, his 3Δ1, leu2Δ0, met15Δ0, ura3Δ0, NRG1-TAP-HIS3MX6	Yeast TAP Fusion Collection (Open Biosystem)
TBF1-TAP	Mat a, his 3Δ1, leu2Δ0, met15Δ0, ura3Δ0, TBF1-TAP-HIS3MX6	Yeast TAP Fusion Collection (Open Biosystem)

3.2 PCR

3.2.1 Costruzione del mutante *SNR52 Cdown*

Si sono allestite inizialmente due PCR: in una la coppia di primer è snR52_fw e snR52_Cdown_rev, nell'altra è snR52_Cdown_fw e snR52_rev (per le sequenze vedi tabella 3.1).

1X Pfu buffer (Promega),
0.25 mM dNTP,
0.56 mM primer forward,
0.56 mM primer reverse,
2.5 U di Pfu (Promega),
20 ng di pNEB-SNR52 wt,
H₂O (Sigma) a 25 µl

94°C per 4'

94°C per 45'' }
55°C per 45'' } 25 cicli
72°C per 1' }

72°C per 5'

4°C for ever

Si è poi eseguita una nuova PCR usando i due precedenti ampliconi come template.

1X thermo pol buffer

0.25 mM dNTP,

10 ng di amplicone 1,

10 ng di amplicone 2,

1,75 U di Deep Vent (New England Biolabs),

H₂O (Sigma) a 50 µl.

94°C per 5'

94°C per 1' }
56°C per 50'' } 5 cicli
72°C per 2' }

72°C per 2'

Si aggiungono poi 1 µl di ciascun primer (20 µM) esterno non mutagenico.

94°C per 4'

94°C per 45'' }
56°C per 45'' } 25 cicli
72°C per 1' }

72°C per 5'

4°C for ever

3.2.2 Costruzione dei mutanti *SNR52 Ddown* e *SNR52 D less*

In entrambi i casi è stata assemblata una reazione di PCR utilizzando un primer reverse mutagenico, rispettivamente snR52 Ddown_rev e snR52_rev_mini (per le sequenze vedi tabella 3.1).

1X Pfu buffer (Promega),

0.25 mM dNTP,

0.5 mM primer snR52_fw,

0.5 mM primer reverse,

1 U di Pfu (Promega),

50 ng di pNEB-SNR52 wt,

H₂O (Sigma) a 50 µl

94°C per 5'

94°C per 45''	}	35 cicli
55°C per 45''		
72°C per 2'		

72°C per 5'

4°C for ever

NOME PRIMER	SEQUENZA
snR52_fw	CTTTGAAAAGATAATGTATGATTATGC
snR52_rev	GCGTTCCATACTGTCAGAGGT
snR52_Cdown_rev	ATGTCATGGAGGATAGTAGATCATTATC
snR52_Cdown_fw	GATCTACTATCCTCCATGACATTAGC
snR52_Ddown_rev	GACAAAAAATAAATTGGAGAAGGAAGGC
snR52_rev_mini	AAAAAATAAATCCTAAAATCTTTCG

Tabella 3.1: primer utilizzati per ottenere le versioni mutate di *SNR52*.

3.2.3 Costruzione dei mutanti per lo studio della regione a monte dei geni per snoRNA

In generale è stato seguito il seguente protocollo, ad eccezione dei mutanti *SNR31 RRPE down*, *SNR32 -127*, *SNR65 -59*. Tutte le sequenze dei primer utilizzati sono riportate nella tabella 3.2.

1X Thermo Pol buffer

0.3 mM dNTP

0.5 μ M primer forward

0.5 μ M primer reverse

1 U di Deep Vent (New England Biolabs)

50 ng di DNA genomico

H₂O (Sigma) a 50 μ l

Per amplificare i restanti tre mutanti è stato utilizzato l'enzima GoTaq (Promega), con il seguente mix di reazione.

1X GoTaq buffer

0.2 mM dNTP

0.4 μ M primer forward

0.4 μ M primer reverse

1 U di GoTaq (Promega)

50 ng di DNA genomico

H₂O (Sigma) a 50 μ l

Nel caso dei costrutti con delezioni o mutazioni in 5' rispetto ai geni *SNR31* e *SNR32* i templati utilizzati sono pFL45S- *SNR31* e pFL45S- *SNR32*, rispettivamente.

94°C per 4'

94°C per 45''
X°C per 45''
72°C per Y' } 35 cicli

72°C per 7'

4°C for ever

Per le temperature di “annealing” (X) e i tempi di allungamento (Y) utilizzati nelle varie reazioni di PCR vedi tabella 3.2.

Per le PCR “touchdown” e “touchup” si sono fatti 2 cicli per ogni temperatura di “annealing”, che viene variata di 2°C fino al raggiungimento dell’ultima temperatura, a cui si fanno 20 cicli.

COSTRUTTO	PRIMERS	X	Y
<i>SNR31 wt</i>	snR31_fw TGAAGTGAAGATATGTACACC	55°C	1' 30''
	snR31_rev GCCAATTCTTGGTTTCAACGC		
<i>SNR31 -220</i>	snR31_-220 GGTTTTCTTATCCTGAAGGATCC	54°C	1' 30''
	snR31_rev		
<i>SNR31 -124</i>	snR31_-124 TGAAAATTTTTCGCTTTATCTAGG	48°C	2' 15''
	snR31_rev		
<i>SNR31 RRPE down</i>	snR31_RRPEdown TGAGCGTACCTCGCTTTATCTAGG	touch	2' 15''
	snR31_rev	42/54	
<i>SNR31 -107</i>	snR31_-107 ATCTCTTAGGGCTCCTACTGTCCG	54°C	1' 30''
	snR31_rev		
<i>SNR31 -9</i>	snR31_-9 CAGAAAAGTGAAGCAAAATTACACCA	54°C	1' 30''
	snR31_rev		
<i>SNR32 wt</i>	snR32_fw CCTATCCATACCCGTTTCTAGACG	50°C	1' 30''
	snR32_rev TTAACTTTTTGTAATTTTGC		
<i>SNR32 -224</i>	snR32_-220 TAAGCTCTTTGTATGTAAAGTGAGC	54°C	1' 30''
	snR32_rev		
<i>SNR32 -148</i>	snR32_-148 GAACCCAAAAAATCAATCACC	48°C	1' 30''
	snR32_rev		
<i>SNR32 -127</i>	snR32_-127 AGGGTAAGCGGGAAAAAAAAA	touch	1' 30''
	snR32_rev	47/41	
<i>SNR32 Reb down</i>	snR32_reb_down AATGTCAGCGGGAAAAAAAAA	50°C	2' 15''
	snR32_rev		

COSTRUTTO	PRIMERS	X	Y
<i>SNR32 poliA down</i>	snR32_poliAdown AGGGTAAGCGGGAACAGTAACAAGTGTAGGAACC GCACTATAAA	55°C	2' 15''
	snR32_rev		
<i>SNR32 Reb down-poliA down</i>	snR32_Rebdown_poliAdown AATGTCAGCGGGAACAGTAACAAGTGTAGGAACC GCACTATAA	55°C	2' 15''
	snR32_rev		
<i>SNR32 -100</i>	snR32_-100 AGGAACCGCACTATAAAAAGTCG	54°C	1' 30''
	snR32_rev		
<i>SNR32 -7</i>	snR32_-7 AAAGGCCAACATCATAATTATATCG	54°C	1' 30''
	snR32_rev		
<i>SNR64 wt</i>	snR64_fw TTCAATGATTACTACAAGAGGC	48°C	2' 15''
	snR64_rev TTGTTTGCTTATATCCAGCG		
<i>SNR64 -231</i>	snR64_-276 GCGCAAATACTACTTTGTGG	56°C	2'
	snR64_rev		
<i>SNR64 -89</i>	snR64_-134 CGAAAAATACTCCAGCGGAGG	56°C	2'
	snR64_rev		
<i>SNR64 -66</i>	snR64_-111 GTATAAAAGGGAGGAAAATAC	53°C	2'
	snR64_rev		
<i>SNR64 +27</i>	snR64_-18 GCAAAAACAGCGACTGGCGCG	56°C	2'
	snR64_rev		
<i>SNR65 wt</i>	snR65_fw ATAATCATATACAAATATCACC	48°C	2' 15''
	snR65_rev TGTATTGTACTIONTGTATTGCCG		
<i>SNR65 -219</i>	snR65_-240 GCTGAAACATGATAGCCGCC	58°C	2'
	snR65_rev		
<i>SNR65 -98</i>	snR65_-119 GATTATTTTATGCGCGCCTCC	58°C	2'
	snR65_rev		
<i>SNR65 -59</i>	snR65_-176 TATATAACATAGGTGGAAAACTT	touch	2'
	snR65_rev	48/38	
<i>SNR65 +7</i>	snR65_-15 CACGAAAAGGCGTAGACATGG	58°C	2'
	snR65_rev		

Tabella 3.2: primer utilizzati per ottenere gli ampliconi utilizzati per lo studio dei promotori dei geni per gli snoRNA.

3.3 Clonaggi

Per l'espressione in lievito di tutti i costrutti è stato utilizzato il vettore "shuttle" pFL45S (*Amp^r*, *TRP1*). Per tutti i clonaggi riguardanti lo studio di *SNR52* è stato utilizzato il metodo descritto in Bolchi *et al.*, 2005; stessa cosa per lo studio dei promotori, ad eccezione dei costrutti *SNR31 RRPEdown*, *SNR31 -124* e *SNR65 -59* che sono stati dapprima clonati in pGEM[®]-T Easy Vector (Promega) e successivamente subclonati in pFL45S.

Nel primo caso la reazione (10 µl finali) comprende

50 ng di pFL45S,

1X buffer 4 (New England Biolabs),

0,1 mg/ml BSA,

1 mM ATP,

2,5 U di ligasi (GE Healthcare),

2,5 U di SmaI (GE Healthcare),

inserto in eccesso di almeno 20 volte (in moli) rispetto al vettore.

La reazione è stata incubata per 3 ore a 25°C e quindi precipitata. Si sono trasformate cellule di *E. coli* (ceppo XL1b), poi il plasmide è stato estratto con il metodo della lisi alcalina.

Nel secondo caso la reazione comprende

50 ng di pGEM-T Easy Vector,

1X rapid ligation buffer,

3 U di ligasi,

inserto in rapporto 1:1 rispetto al vettore (in moli).

La reazione viene incubata o/n a 4°C e poi precipitata. Dopo la trasformazione di batteri elettrocompetenti e l'estrazione del plasmide, quest'ultimo è stato digerito con opportuni enzimi di restrizione, è stato recuperato l'inserto (mediante estrazione da gel d'agarosio) e quindi questo è stato subclonato in pFL45S, secondo il seguente protocollo (20 µl finali):

50 ng di pFL45S (tagliato con gli stessi enzimi di restrizione dell'inserto e defosforilato),

1X ligation buffer,

1 mM ATP,

8 U di ligasi (GE Healthcare).

È stata fatta una nuova trasformazione e infine si è recuperato il plasmide.

3.4 Trasformazione di lievito (“quick yeast transformation”)

Il ceppo di lievito da trasformare viene fatto crescere in 50 ml di YPD (1% yeast extract, 2% peptone, 2% glucosio) fino ad una OD di 0.4-0.5; le cellule vengono raccolte e lavate con 20 ml di H₂O sterile e poi con 10 ml di TE/0.1 M LiOAc pH 7.5, quindi risospese in un volume di TE/0.1 M LiOAc pH 7.5 proporzionale all'OD ottenuta (500 µl per una OD pari a 1).

A 50 µl di cellule si aggiungono 15 µg di “salmon sperm DNA” a singolo filamento e 2 µg di DNA plasmidico, si mettono poi a 30°C per 30 minuti, agitando delicatamente ogni 10 minuti. Si aggiungono 300 µl di 40% PEG 8000/TE/0.1 M LiOAc pH 7.5 a ciascuna trasformazione e si mescola a lungo, poi si rimettono a 30°C per 30 minuti. Si fa poi uno shock termico, lasciando le trasformazioni a 42°C per 15 minuti (questo passaggio non è stato fatto nel caso del ceppo ΔSNR52). Le cellule sono infine piastrate su terreno minimo (0.67% yeast nitrogen base, 2% glucosio, 2% agar) supplementato con opportuni amminoacidi.

3.5 Estrazione dell'RNA

Le cellule vengono fatte crescere in 100 ml di terreno liquido fino al raggiungimento della fase logaritmica; viene raccolta una quantità di cellule pari a 10 OD da cui poi viene estratto l'RNA.

Il campione viene risospeso in 400 µl di AE buffer (50 mM sodio acetato pH 5, 10 mM EDTA), si aggiunge SDS 1% finale e si passa sul vortex; si aggiungono 500 µl di fenolo acido/cloroformio 5:1 e si passa a lungo sul vortex, poi si mette il campione a 65°C per 4 minuti e successivamente in ghiaccio secco, finché non è completamente congelato. Si centrifuga per almeno 5 minuti e si recupera il surnatante, a cui si aggiunge 1 volume di fenolo acido/cloroformio 1:1, si passa sul vortex e si centrifuga a 4°C per almeno 8 minuti. Al surnatante recuperato si aggiungono 1/10 di volume di sodio acetato 3 M, pH 5.2 e 2.5 volumi totali di etanolo 100% freddo per precipitare l'RNA; si lava il pellet con 1 ml di etanolo 70% freddo e infine l'RNA è risospeso in 50 µl di H₂O DEPC (o 30 µl se la resa è bassa). Gli RNA vengono quantificati allo

spettrofotometro e 1 o 2 μg sono caricati su gel di acrilammide 6%, 7 M urea come controllo.

3.6 Northern blot

Su gel di acrilammide 6%, 7 M urea si caricano 10 μg di RNA totale, risospesi in soluzione denaturante colorata; dopo la corsa elettroforetica l'RNA è trasferito sulla membrana di nitrocellulosa GeneScreen Plus (PerkinElmer), mediante "semi-dry blotting".

La preibridazione viene effettuata per almeno 1 ora e 30 minuti in una soluzione 5X SSC, 5X Denhardt solution, 0.1 mg/ml salmon sperm DNA, 0.5% SDS (per le temperature utilizzate nei vari Northern blot vedi tabella 3.3).

Per i piccoli RNA, la sonda è costituita da un oligonucleotide marcato con ^{32}P in 5':

1X buffer

10 pmol di oligonucleotide

10 U di T_4 polinucleotide chinasi (Promega)

5 μl di $\gamma^{32}\text{P}$ -ATP (10 mCi/ml) (GE Healthcare)

H_2O (Sigma) a 20 μl

La reazione viene lasciata a 37°C per almeno 30 minuti, poi l'enzima è inattivato a 90°C per 7 minuti, quindi la sonda viene purificata con colonnine Micro Bio-Spin 6 (Bio-Rad), dopo aver portato il volume a 40 μl , e la radioattività viene misurata al β counter. Dopo 5 minuti di denaturazione a 90°C la sonda viene aggiunta alla soluzione di preibridazione e lasciata o/n, alla stessa temperatura della preibridazione.

Lavaggi:

1°: 2X SSC, 0.1% SDS; 10 minuti; temperatura ambiente.

2°: 1X SSC, 0.1% SDS; 7 minuti; temperatura ambiente o 42°C (se l'ibridazione è stata fatta a temperature maggiori di 30°C).

3° (facoltativo): 0.2X SSC; 5 minuti ; stessa temperatura dell'ibridazione.

Il filtro è esposto al Phosphorimager e la lastra successivamente scannerizzata.

Per le sonde per gli mRNA di dimensioni maggiori è stato utilizzato il kit Megaprime DNA labeling system (GE Healthcare): è stato amplificato un frammento del gene di interesse di circa 500 bp (la regione amplificata non deve avere omologia con altri geni). A 50 ng di amplicone vengono aggiunti 2.5 μl di primers random e H_2O (Sigma),

come da protocollo, si mette per 5 minuti a 95°C, quindi in ghiaccio e successivamente si aggiungono il buffer, $\alpha^{32}\text{P}$ -dCTP ed enzima, si mette a 37°C per 25-30 minuti, poi si inattiva con 2.5 μl di EDTA 0.2 M e si porta a 40 μl . Si purifica con colonnine Micro Bio-Spin 6 (Bio-Rad), poi si quantifica la radioattività. Dopo 5 minuti di denaturazione la sonda viene aggiunta alla soluzione di preibridazione, che è stata effettuata a 65°C. Per questi Northern blot l'RNA è stato caricato su gel d'agarosio 1.2%, 1.9% formaldeide, 1X MOPS, quindi trasferito con il metodo "semi-dry blotting".

SONDA	SEQUENZA	T°
snr52_mini_north	GTATCAGAGATTGTTTCACGCTAATG	42°C
U3	CTTCCAAAGGAAAAAACTACAAATGCAACGGCAAAG	42°C
snR33-r	CAATTCCAATAAGAAAACGAAGGA	28°C
Ala_AGC_probe	GGAGACCTCTCCCATGCTAAGGGAGCGCGC	42°C
SNU13_probe	GGTTCACAGTCAGCAGCCATAATGATGAATTCGGA	28°C
snR31_probe	GGTGTTTAAAAAGGCGAGGTGA	28°C
snR32_probe	GTGGGCTTTTAGTCATTCCCTG	28°C
snR64_probe	GAGAATCTGTTGTCCCTATCTGG	28°C
snR65_probe	GATACTATCTAGCATAAATTGTG	28°C
snR24_probe	GGAACTCAAAGTTCCATCTG	28°C
NOP1	NOP1_fw TGGTGCCCGTGGTGGTGCCAAGG	65°C
	NOP1_rev TGTGGGTGTGTAGCATCTTCAATGATTGG	
NOP58	NOP58_full_fw GCTTACGTTTTAACTGAACTTCAGC	65°C
	NOP58_full_rev CTTGTCCTTCTTTTCCTTCTTAGC	

Tabella 3.3: sonde utilizzate negli esperimenti di Northern blot

3.7 Deplezione di *Nop1p*, *Nop58p*, *Snu13p*

Per studiare gli effetti sulla biosintesi di snR52 della mancanza delle proteine "core" della snoRNP, sono stati utilizzati ceppi in cui i geni *NOP1*, *NOP58* e *SNU13* sono posti sotto il controllo di un promotore inducibile con galattosio (Tollervey *et al.*, 1991; Lafontaine and Tollervey, 1999; Dobbyn and O'Keefe, 2004), in modo da poterli spegnere, dal momento che tutti i geni sono essenziali.

Le cellule sono cresciute in YPGal (1% yeast extract, 2% peptone, 2% galattosio) fino a OD pari a 0.4, quindi raccolte, lavate due volte con H₂O sterile e risospese in un uguale volume di YPD; si sono poi effettuate raccolte a vari tempi. I lieviti sono stati mantenuti in fase di crescita esponenziale diluendo la coltura con YPD preriscaldato. Per il ceppo Gal::NOP58 e il relativo wt il terreno “permissivo” è costituito da 1% yeast extract, 2% peptone, 2% galattosio, 2% raffinosisio, 2% saccarosio.

3.8 Chromatin immunoprecipitation (ChIP)

1° giorno

I ceppi “TAP-tagged” d’interesse e il ceppo “parental” vengono inoculati in 200 ml di YPD e lasciati crescere fino a OD pari a 0.4-0.5; per “cross-linkare” DNA e proteine, tratto con formaldeide 1% finale per 20 minuti a temperatura ambiente, agitando occasionalmente, quindi aggiungo glicina 240 mM finale, per bloccare la formaldeide, e lascio per altri 5 minuti. Si raccolgono le cellule e si lavano due volte con 5 ml di TBS (20 mM Tris-HCl, pH 7.5, 150 mM NaCl) freddo, e una volta con 5 ml di FA-lysis buffer (50 mM HEPES-KOH, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodio deossicolato, 0.1% SDS, 1 mM PMSF).

2° giorno

Le cellule vengono risospese in 1 ml di FA-lysis buffer freddo contenente 0.5% SDS, viene aggiunto lo stesso volume di “glass beads” (diam. 0.5 mm) e quindi si passa sul vortex per 15 minuti a 4°C, per rompere le cellule; si aggiungono 8 ml di FA e si buca il fondo della provette con uno spillo incandescente: il lisato viene raccolto in tubi da centrifuga tenuti in ghiaccio. La cromatina viene sedimentata centrifugando per 20 minuti a 20000 g, viene poi lavata per 1 ora a 4°C con 1.5 ml di FA, centrifugata nuovamente e risospesa in 1.5 ml di FA, quindi lasciata per 1 ora a 4°C.

Si sonica in modo da avere frammenti di 100-300 bp: per verificare la taglia effettiva del DNA, si trattano 80 µl di ciascun campione con Pronase (Calbiochem) 1.2 mg/ml finali, a 42°C per 1 ora e poi a 65°C per 3 ore; estraggo con fenolo/cloroformio, quindi carico 2 µl su gel d’agarosio 1%.

Il volume dei campioni viene portato a 4 ml con FA, si centrifuga per 20 minuti a 20000 g e si recupera il surnatante; si preparano aliquote da 800 µl e due da 400 µl che possono essere conservate a -80°C. A una delle aliquote da 800 µl aggiungo 10 µl di

“beads” IgG-Sepharose (GE Healthcare), precedentemente lavate due volte con 100 µl di TE (10 mM Tris, pH 8, 1 mM EDTA); si lascia o/n a 4°C in agitazione.

3° giorno

Si lavano le “beads” con 1.4 ml di FA+275 mM NaCl per 5 minuti, quindi si ripete con 1.4 ml di FA+500 mM NaCl, poi con 1.4 ml di “wash buffer” (10 mM Tris, pH 8, 0.25 M LiCl, 1 mM EDTA, 0.5% NP-40, 0.5 % sodio deossicolato) e infine con 1.4 ml di TE. La cromatina viene fluita scaldando le “beads” per 10 minuti a 65°C in 200 µl di buffer di eluizione (50 mM Tris, pH 7.5, 10 mM EDTA, 1% SDS), il surnatante viene recuperato e le “beads” vengono lavate con 200 µl di TE, che viene recuperato e unito al surnatante raccolto in precedenza. Si aggiunge Pronase 0.8 mg/ml e si lascia 1 ora a 42°C e poi 5 ore a 65°C: lo stesso trattamento viene fatto su una delle aliquote da 400 µl di cromatina, che diventerà così il controllo “input”.

Si estrae il DNA con fenolo/cloroformio/isoamyl alcool e poi con cloroformio; si precipita con etanolo e sodio acetato, in presenza di 20 µg di glicogeno.

4° giorno

Il DNA viene risospeso in 50 µl di Tris 10 mM, pH 8; diluisco ulteriormente il controllo “input” 1/100.

PCR (10 µl finali)

0.25 µM di ciascun primer specifico (vedi tabella 3.4 per le sequenze)

0.25 µM di ciascun primer di controllo

0.1 mM di ciascun dNTP

1X PCR buffer

0.5 U di GoTaq (Promega)

0.06 mCi/ml di $\alpha^{32}\text{P}$ -dCTP (PerkinElmer)

templato (in genere 1/100 del DNA immunoprecipitato; 1/20000 o 1/50000 del DNA “input”).

94°C per 5'

94°C per 45''
53°C per 45''
72°C per 1'20'' } 25 cicli

72°C per 7'

4°C for ever

I prodotti di PCR sono caricati su gel di acrilammide 6%-TBE standard 1X; il gel seccato è stato esposto al Phosphorimager e quindi la lastra è stata scannerizzata. Come controllo si amplifica anche una regione a cui non si dovrebbero legare le proteine d'interesse (la regione scelta è compresa tra i geni *PST2* e *MRH1*). Il segnale ottenuto nei campioni immunoprecipitati viene normalizzato rispetto all'efficienza di amplificazione (determinata dal segnale dei campioni "input"), quindi si effettua il rapporto tra il segnale normalizzato ottenuto per la regione d'interesse e il segnale normalizzato ottenuto per la regione di controllo: si ottiene così un valore che esprime quante volte il segnale della regione di interesse è maggiore rispetto al "background".

NOME PRIMER	SEQUENZA
snR31_-220	GGTTTTCTTATCCTGAAGGATCC
snR31_ChIP_rev	CCTCATGCTTTAAATATACTCAAC
snR32_-224	TAAGCTCTTTGTATGTAAAGTGAGC
snR32_ChIP_rev	GGAAAATTAATACAAACCTTCG
snR42_ChIP_fw	CGTGTTCTACAGAGCACACGG
snR42_ChIP_rev	GCAAATGCTATAGTTCACGGC
snR45_ChIP_fw	CATCTCATCTCATCAAAACCG
snR45_ChIP_rev	TACCTGTAAATGAAAATCG
PST2-MRH_IntergReg_fw	GTCGTCTAAGGATGGTAACTAGC
PST2-MRH_IntergReg_rev	AGAACA AAAATTACTTAAGGGTCG

Tabella 3.4: primer utilizzati nelle PCR della ChIP.

3.9 Deplezione d'azoto

Le cellule (BY4741) vengono cresciute in terreno SC fino al raggiungimento della fase logaritmica, quindi viene raccolto il campione al tempo 0, mentre il resto delle cellule viene lavato con H₂O sterile e inoculato in un uguale volume di terreno minimo con

concentrazione limitante di azoto e addizionato con opportuni amminoacidi; il terreno minimo è così composto: 0.67% “yeast nitrogen base w/o ammonium, copper, iron” (Q-BIO gene), 0.25 μ M (40 μ g/l) solfato di rame, 200 μ g/l cloruro di ferro, 25 mg/l solfato d’ammonio. Si effettuano poi raccolte di cellule a vari tempi.

3.10 Induzione da glucosio

Le cellule vengono cresciute in YP contenente glicerolo al 3% come fonte di carbonio (non fermentabile), fino al raggiungimento della fase logaritmica precoce (OD pari a 0.4-0.5). Una parte della coltura viene raccolta (tempo 0), quindi al resto viene aggiunto glucosio al 2% e le cellule vengono lasciate crescere per altri 20 minuti, tempo necessario per avere la massima induzione dei geni legati alla biogenesi del ribosoma (Wang *et al.*, 2004).

3.11 Primer extension

A 5 μ g di RNA si aggiungono 2 pmol di primer marcato con γ^{32} P-ATP (vedi paragrafo 3.6), 1 μ l di dNTP 10 mM e si porta a 12 μ l con H₂O (Sigma); si mette a 65°C per 5 minuti e poi in ghiaccio per 2 minuti. Si aggiungono 4 μ l di “first strand buffer” 5X, 1 μ l di DDT 0.1 M, 40 U di Superase (Ambion) e 200 U di SuperScript III (Invitrogen); si mette a 55°C per 1 ora, quindi si inattiva a 70°C per 15 minuti. Si precipita il cDNA con ½ V di acetato d’ammonio 7.5 M e etanolo 100% e alla fine il pellet viene risospeso in 10 μ l di soluzione denaturante colorata.

Si caricano 5 μ l di ciascuna reazione (dopo aver denaturato per 5 minuti) su gel di acrilammide 7%, 7 M urea, 1X TBE standard; si caricano anche 3 μ l delle reazioni di sequenziamento (vedi paragrafo 3.11.1) denaturate. La corsa si effettua a 30 W per circa 2 ore e 30 minuti; il gel viene quindi seccato ed esposto o/n al Phosphorimager; la lastra viene quindi scannerizzata.

3.11.1 Sequenziamento

Si utilizza il kit Thermo-Sequenase (GE Healthcare): a 500 ng di templatato plasmidico si aggiungono 1 pmol di primer marcato, 2 μ l di “Thermo-Sequenase rxn buffer concentrated”, 2 μ l di DNA polimerasi e H₂O (Sigma) fino a 17.5 μ l. Di questo mix si

prelevano 3.5 µl e si aggiungono a ciascun ddNTP, quindi si aggiungono 5 µl di olio per PCR.

94°C per 45'' }
55°C per 45'' } 40 cicli
72°C per 2' }

4°C for ever

Si aggiungono infine 4 µl di soluzione denaturante colorata.

3.12 Isolamento dell'RNA poli A+

È stato utilizzato il kit *PolyAtract*® *mRNA Isolation Systems IV* (Promega): 1 mg di RNA totale, contenuto in 500 µl di H₂O DEPC, viene denaturato per 10 minuti a 65°C, vengono poi aggiunti 3 µl di *Biotinylated Oligo(dT) Probe* e 13 µl di SSC 20X "RNase-free", e la soluzione viene quindi incubata a temperatura ambiente, fino al raffreddamento.

Nel frattempo le particelle paramagnetiche a cui è associata la streptavidina vengono risospese nel loro buffer e poi separate dal surnatante grazie a un magnete, il surnatante viene rimosso e le particelle sono lavate tre volte con 300 µl di SSC 0.5X, quindi risospese in 100 µl di SSC 0.5X.

La reazione di "annealing" viene aggiunta alle particelle e si tiene in incubazione a temperatura ambiente per 10 minuti, miscelando per inversione; il surnatante viene rimosso e le particelle sono lavate quattro volte con 300 µl di SSC 0.1X, infine le particelle vengono risospese in 100 µl di H₂O DEPC, per eluire l'RNA. L'eluato viene trasferito in un nuovo tubo e le particelle vengono lavate con 150 µl di H₂O DEPC, che vengono poi recuperati e aggiunti al precedente eluato. La concentrazione dell'RNA poli A+ viene determinata allo spettrofotometro (tipicamente si ottiene l'1-5% dell'RNA totale di partenza).

3.13 Trascrizione *in vitro*

La reazione avviene in un volume finale di 50 μ l, con le seguenti condizioni: 10 mM Tris/HCl, pH 8, 5 mM MgCl₂, 100 mM KCl, 1 mM DDT, glicerolo 8%, 500 μ M ATP, CTP, GTP, 25 μ M UTP, 2 μ Ci di α^{32} P-UTP, 200 ng di templatato, 1 μ l di estratto nucleare (preparato come descritto in Dieci *et al.*, 1993), 10 U di Superase (Ambion). Si lascia per 20 minuti a temperatura ambiente, poi si ferma la reazione mettendo i tubi in ghiaccio secco. Si estrae aggiungendo 50 μ l di fenolo-cloroformio 1:1, pH 8 e 60 μ l di H₂O (Sigma); si recuperano 88 μ l di surnatante e si passano in tubi contenenti 15 μ g di RNA “carrier”, quindi si precipita con 50 μ l di acetato d’ammonio 7.5 M ed etanolo 100%. Il pellet ottenuto è risospeso in 15 μ l di soluzione denaturante colorata. Il campione viene denaturato e caricato su gel d’acrilammide 6%, 7 M urea; il gel è successivamente seccato ed esposto al Phosphorimager e quindi la lastra viene scannerizzata.

3.14 Analisi computazionale

Le sequenze codificanti per gli snoRNA di *S. cerevisiae* sono quelle presenti nel database SGD (www.yeastgenome.org); le sequenze codificanti per gli snoRNA nelle specie *S. bayanus* (o la variante *S. uvarum*), *S. mikatae*, *S. paradoxus*, *S. kudriavzevii* sono state ricavate mediante ricerca di omologia, utilizzando il programma BLASTN (Altschul *et al.*, 1997) (le sequenze con valore E>0.1 sono state considerate poco significative). Le sequenze a monte dei geni per snoRNA nelle cinque specie considerate sono state ricavate mediante ricerca nel database NCBI (www.ncbi.nlm.nih.gov), avendo a disposizione le sequenze codificanti per gli snoRNA: si sono considerate 400 bp a monte della sequenza codificante se lo snoRNA non viene processato in 5’, 600 bp se lo snoRNA è processato (o se non c’è la certezza del contrario).

Per l’allineamento delle sequenze è stato utilizzato il programma ClustalX v 1.81 e gli allineamenti ottenuti sono stati visualizzati con GeneDoc v 2.5.

Le ORF adiacenti ai geni per snoRNA in *S. cerevisiae* sono state classificate in base alla funzione utilizzando “The MIPS Functional Catalogue Database” (<http://mips.gsf.de/genre/proj/yeast/Search/Catalogs/catalog.jsp>).

4. RISULTATI E DISCUSSIONE

4.1 Analisi delle regioni a monte dei geni per snoRNA

Poco si sa dei promotori dei geni per snoRNA indipendenti in *S. cerevisiae*; più studiato è stato il promotore di *SNR52*, trascritto dalla RNA polimerasi III e di cui si parlerà nel paragrafo 4.1.1.

Un altro promotore già analizzato è quello del “cluster” costituito dai geni *SNR78*, *SNR77*, *SNR76*, *SNR75*, *SNR74*, *SNR73*, *SNR72* (Qu *et al.*, 1999): a monte sono stati individuati, oltre alla TATA box una sequenza “A/T rich” e due siti di legame per la proteina Rap1p, tutti elementi richiesti per una efficiente trascrizione del “cluster”. Queste sequenze sono state individuate anche a monte di alcuni altri geni per snoRNA, così come il sito di legame per Abf1p (Qu *et al.*, 1999).

Nello studio qui presentato si sono scelti quattro geni indipendenti monocistronici per l’analisi sperimentale delle regioni a monte; in parallelo è stata effettuata un’analisi computazionale delle regioni a monte di tutti i geni indipendenti, per identificare regioni conservate, e quindi potenzialmente coinvolte nella trascrizione, nei Saccaromiceti

4.1.1 Analisi computazionale

A partire dalle sequenze codificanti per snoRNA in *S. cerevisiae* è stata fatta una ricerca di omologia all’interno del regno dei funghi per identificare le specie in cui i geni per snoRNA sono maggiormente conservate: la conservazione è evidente in numerose specie di funghi e in particolare all’interno di cinque specie del genere *Saccharomyces*. Inoltre si può notare come nessun gene per snoRNA sia stato duplicato nelle specie considerate. Da questa prima ricerca sono stati esclusi i geni intronici, in quanto non saranno presi in considerazione per le successive analisi. I risultati ottenuti sono riportati nella Tabella 1, presente in Appendice; per altri dettagli sulla ricerca effettuata si veda il paragrafo 3.14.

Nel caso delle cinque specie di Saccaromiceti che presentano maggiore conservazione, *S. cerevisiae*, *S. bayanus* (o la variante *S. uvarum*), *S. mikatae*, *S. paradoxus*, *S. kudriavzevii*, sono state considerate poi le sequenze “upstream” (vedi paragrafo 3.14);

le altre specie di funghi presentano una notevole divergenza nel caso delle regioni a monte, per cui non sono state considerate.

Dall'allineamento delle sequenze a monte di ogni gene di *S. cerevisiae* con le sequenze a monte dei geni ortologhi nelle altre specie è emersa la conservazione di alcuni elementi ricorrenti. La TATA box è molto ben conservata e quasi sempre presente (vedi Tabella 4.1); la sequenza consenso, che si ricava confrontando le TATA box in *S. cerevisiae*, è in accordo con quella presente in Basehoar *et al.*, 2004 (vedi allineamento presente in Appendice). A monte della TATA box sono spesso presenti altri due elementi di sequenza conservati: una è molto simile (e a volte uguale) alla sequenza consenso per Reb1p, l'altra (ARCCCTaa) è riconducibile alla sequenza che recentemente è stata indicata come la sequenza consenso per Nrg1p (Stanley *et al.*, 2006), ma che può anche essere ricondotta alla sequenza che viene legata da Tbf1p (figura 4.1).



Figura 4.1: allineamento di alcune regioni a monte dei geni per snoRNA in *S. cerevisiae* che mostra la conservazione dell'elemento che può essere ricondotto alla sequenza consenso per Nrg1p o Tbf1p. A destra è riportata la distanza dell'elemento dalla TATA box, per ogni gene.

Reb1p è una proteina legante il DNA (consenso: CGGGTAA o TTACCCG), essenziale per la vitalità della cellula; questa proteina è implicata nella trascrizione della RNA polimerasi I e della RNA polimerasi II, nella terminazione pol I, nell'organizzazione della cromatina e, inoltre, può esercitare una regolazione sia positiva che negativa sul

proprio gene (Wang and Warner, 1998). È stato anche mostrato un ruolo di Reb1p come fattore che impedisce il propagarsi del “silencing” oltre i telomeri (Fourel *et al.*, 2001). Nel caso del consenso è permessa una certa flessibilità e l’affinità di legame può essere influenzata dai nucleotidi adiacenti.

Nrg1p è noto come repressore richiesto per la repressione da glucosio (Park *et al.*, 1999) ed è coinvolto in altri processi quali la crescita invasiva e la risposta a pH alcalini. La sequenza consenso fornita in Park *et al.*, 1999 è CCCT, in Harbison *et al.*, 2004 è GGACCCT, mentre in Stanley *et al.*, 2006 è GCCCTAA: quest’ultima sequenza è quella che viene ritrovata a monte dei geni per snoRNA nei cinque Saccaromiceti analizzati.

Tbf1p è in grado di legare le ripetizioni TTAGGG o TTAGGG-like presenti nelle regioni subtelomeriche, dove sono associate ai siti di legame per Reb1p; inoltre è stato osservato come Tbf1p possa legarsi anche all’interno dei cromosomi, in particolare nelle regioni intergeniche, cosa che permette di ipotizzare un ruolo di questa proteina nel delimitare regioni di espressione sul genoma. Stesso ruolo può essere ipotizzato anche per Reb1p (Koering *et al.*, 2000).

Un altro elemento la cui presenza è stata evidenziata in questa analisi è RRPE (ribosomal RNA processing element), frequente a monte dei geni coinvolti nella biogenesi del ribosoma, e il cui consenso è TGAAAA(A/T)TTT(T/C) (Hughes *et al.*, 2000); Stb3p è stata recentemente individuata essere una proteina legante questo motivo (Liko *et al.*, 2007). La mancanza di questa proteina ha effetti sull’induzione da glucosio di molti geni che presentano RRPE a monte, ma anche su geni che sembrano non avere RRPE: da questo si può ipotizzare che Stb3p faccia parte di un processo che coinvolge più componenti ridondanti. Il motivo RRPE è molto meno frequente rispetto agli altri due individuati a monte dei geni per snoRNA, e meno conservato; inoltre la sua posizione è estremamente variabile.

Quando è presente a monte di geni codificanti per proteine, RRPE è spesso seguito, ad una distanza di 6-60 bp, da un altro elemento di sequenza, chiamato PAC (“polymerase A and C”), per la ricorrenza nei promotori dei geni codificanti per le subunità delle RNA polimerasi I e III; il consenso è GCGATGAG o CTCATCGC. La combinazione RRPE-PAC viene identificata a monte di *SNR45* e *SNR190* ed effettivamente in questi casi è presente in 5’ un gene coinvolto nella biosintesi del ribosoma, trascritto in

direzione opposta (rispettivamente *NOCA* e *RPA34*). È stato notato come molti geni coinvolti nella biogenesi del ribosoma (e dei quali molti presentano RRPE e/o PAC nel promotore) siano adiacenti sul genoma, con orientamento sia divergente, sia convergente, sia in tandem (Wade *et al.*, 2006).

Entrambi gli elementi sono presenti anche a monte del gene *SNR84* ma in questo caso il gene presente a monte (*HLR1*) è piuttosto distante (1560 bp tra la base +1 di *HLR1* e la prima base codificante per *snR84* maturo) e non sembra essere coinvolto nella biogenesi del ribosoma.

In numerosi casi è stata messa in evidenza la presenza di una sequenza poli(A/T) poco a monte rispetto alla TATA box (si è assunto che poli(A/T) sia una sequenza continua di almeno sette A o sette T; nella Tabella 4.1 il poli(A/T) è stato considerato solo se la prima base è distante dalla TATA box non più di 55 bp); sequenze di questo tipo sono state descritte a monte della TATA box di svariati geni ed è stato dimostrato come esse siano necessarie per una efficiente trascrizione (Struhl, 1985). Sembra che la particolare struttura di queste sequenze di DNA (rigide, con un solco minore schiacciato e con 10 bp per giro, invece di 10.5), chiamate “T-tracts”, interferisca con la formazione e la disposizione dei nucleosomi, anche se sono noti alcuni esempi di “T-tracts” presenti in regioni organizzate in nucleosomi (Suter *et al.*, 2000). Probabilmente, per impedire la formazione dei nucleosomi *in vivo* sono richieste altre attività oltre alla presenza di un elemento “T-rich”, come osservato nel caso del gene *RPS28A* in cui è necessaria la presenza di un “T-tract” e di un sito di legame per Abf1p per creare una regione priva di nucleosomi (Lascaris *et al.*, 2000).

Il sito per Rap1p, già noto essere presente a monte di alcuni geni per snoRNA in *S. cerevisiae* (Qu *et al.*, 1999) viene ritrovato discretamente conservato tra i Saccaromiceti, mentre il sito per Abf1p, anche lui identificato a monte di geni per snoRNA nello stesso studio, non risulta essere conservato.

Alcune regioni a monte risultano essere estremamente poco conservate nei cinque genomi confrontati, per esempio nel caso di *SNR87* o di *SNR52*: quest’ultimo gene è trascritto dalla polimerasi III e le A e B box sono contenute all’interno della regione trascritta (vedi paragrafo 4.2.1), tuttavia risulta non essere conservata neanche la TATA box presente a monte del gene in *S. cerevisiae* (vedi allineamento presente nell’Appendice).

I geni paraloghi *SNR17A* e *SNR17B*, codificanti per U3, presentano sequenze “upstream” notevolmente differenti: nella regione a monte di *SNR17A* è possibile riconoscere una TATA box, un motivo RRPE, un motivo poli(T) (anche se distante 80 bp dalla TATA box), una sequenza AACCCCTGA e un possibile sito di legame per Rap1p, mentre nella regione a monte di *SNR17B* si può identificare solo la TATA box e un possibile sito di legame per Rap1p; la maggiore abbondanza di U3A rispetto a U3B precedentemente osservata (Hughes *et al.*, 1987) potrebbe essere spiegata ammettendo che uno degli elementi presenti a monte di *SNR17A* stimoli fortemente la trascrizione (gli allineamenti sono presenti nell’Appendice).

Molto ben conservata è invece la regione, lunga 163 bp, tra *NME1*, codificante per l’RNA MRP, e *SNR66* (l’allineamento è presente in Appendice). Lo snoRNA snR66 è sintetizzato come precursore, quindi il sito d’inizio della trascrizione di *SNR66* dovrebbe essere molto vicino all’estremità 3’ di *NME1*; in letteratura non è noto che i due geni facciano parte di una unità dicistronica, anche se, da un esperimento di “primer extension” presente in Lee *et al.*, 2003, sembra che la prima base trascritta di *SNR66* si trovi all’interno di *NME1*. La Tabella 4.1 riporta gli elementi identificati a monte delle 56 unità trascrizionali indipendenti in *S. cerevisiae* e la loro frequenza; le posizioni sono invece riassunte nella figura 4.2.

La Tabella 2 presente in Appendice riporta in dettaglio tutti gli elementi di sequenza presenti a monte di ciascun gene e conservati nelle cinque specie di Saccaromiceti analizzati.

	C/D (29)	H/ACA (27)	TOTALE (56)
TATA box	23	22	45
Reb1p binding site	8	14	22
ARCCCTaa	20	17	37
RRPE	9	5	14
Poli (A/T)	12	11	23

Tabella 4.1: frequenza degli elementi di sequenza identificati a monte dei geni indipendenti per snoRNA in *S. cerevisiae*. Se RRPE è presente due volte a monte dello stesso gene, in questa tabella è conteggiato una volta sola.

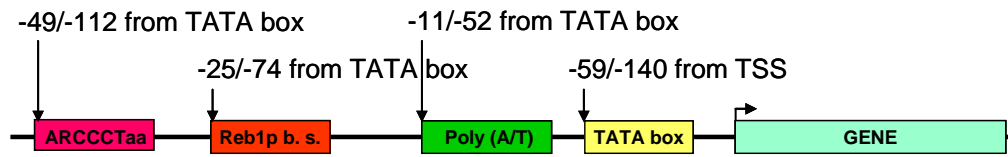


Figura 4.2: schema riassuntivo delle posizioni degli elementi identificati. Quando sono presenti più motivi nella regione a monte del gene per snoRNA, l'ordine è quello mostrato in figura; RRPE non è indicato in quanto la sua posizione è estremamente variabile rispetto alla TATA box e agli altri elementi. TSS: “transcription start site”.

Dall'allineamento mostrato in Appendice. si può vedere come, in *S. cerevisiae*, sia conservata discretamente anche la sequenza attorno al sito di inizio; tale sequenza è in accordo con il consenso $A(A\text{-rich})_5NYA(A/T)NN(A\text{-rich})_6$, riportato in precedenza (Zhang and Dietrich, 2005), ad eccezione della sequenza $(A\text{-rich})_6$.

È infine interessante sottolineare che una porzione non trascurabile dei geni che fiancheggiano i geni per snoRNA è coinvolta nella sintesi proteica, ed in particolare nella biogenesi del ribosoma (vedi Tabella 3 presente in Appendice).

4.1.2 Analisi sperimentale: scelta dei geni

Parallelamente all'analisi computazionale è stata condotta anche una analisi *in vivo* delle regioni a monte di alcuni geni per snoRNA in *S. cerevisiae*. Sono stati scelti due geni per snoRNA H/ACA box (*SNR31* e *SNR32*) e due geni per snoRNA C/D box (*SNR64* e *SNR65*): questi geni sono non essenziali e costituiscono unità indipendenti monocistrioniche.

Studiando il processamento dell'estremità 5' degli snoRNA trascritti in modo indipendente, è stato possibile notare come gli snoRNA H/ACA box siano prevalentemente non processati (16 su 27; per 6 non sono stati trovati riferimenti in letteratura), mentre gli snoRNA C/D box sono nelle maggior parte dei casi processati (24 su 29) (Ghazal *et al.*, 2005; Hiley *et al.*, 2005; Davis and Ares, 2006): i geni scelti seguono questo criterio.

4.1.3 Analisi della regione a monte di *SNR31*

L'allineamento della regione a monte del gene *SNR31* fra i genomi dei cinque Saccaromiceti è mostrato nella figura 4.3.

Per studiare *in vivo* la regione “upstream” di *SNR31* si sono inizialmente preparati quattro costrutti (*SNR31 wt*, *SNR31 -220*, *SNR31 -107*, *SNR31 -9*) in cui tale regione è progressivamente deleta. I costrutti sono stati trasformati nel ceppo YS669, deleto per il gene *SNR31* (Balakin et al., 1993), quindi l'espressione di snR31 è stata analizzata mediante Northern blot (figura 4.4 B): si osserva la trascrizione dei costrutti *SNR31 wt* e *SNR31 -220* mentre i mutanti che presentano una delezione maggiore a monte del gene non vengono trascritti (nonostante la presenza della TATA box e di un elemento conservato con sequenza uguale al sito di legame di Tbf1p, nel costrutto *SNR31 -107*). Per quanto riguarda i mutanti trascritti si può riscontrare un aumento della trascrizione per *SNR31 -220* rispetto al wt: nella regione deleta è presente il gene *SNR5* che viene trascritto nel senso opposto rispetto a *SNR31* (figura 4.4 A); si può ipotizzare che tra i due geni ci sia “competizione” per la trascrizione nella situazione wt e che l'eliminazione della sola sequenza codificante per snR5 provochi una maggiore efficienza di trascrizione di snR31.

Dall'analisi computazionale risulta che nella regione compresa tra 107 e 220 bp a monte del gene è situato, oltre alla TATA box di *SNR5*, un elemento RRPE, per cui si sono preparati due ulteriori costrutti in cui la regione “upstream” comprende l'elemento RRPE e ciò che si trova a valle: in un caso RRPE è integro (costrutto *SNR31 -124*), nell'altro si sono introdotte mutazioni per distruggere RRPE (costrutto *SNR31 RRPE down*). Mediante Northern blot è possibile osservare che la trascrizione avviene nel caso di *SNR31 -124* mentre è drasticamente ridotta nel caso di *SNR31 RRPE down*, cosa che permette di affermare che la presenza di RRPE è fondamentale per una efficiente trascrizione di *SNR31*, la quale invece non è influenzata dalla TATA box di *SNR5*.

Si può rilevare come il motivo RRPE sia richiesto per l'espressione di *SNR31* nonostante tra RRPE e il gene sia presente un possibile sito di legame per Tbf1p, per cui è noto un ruolo nella formazione di regioni genomiche con funzione di “insulator” (Fourel et al., 2001). È possibile che RRPE sia coinvolto anche nella trascrizione di *SNR5* (la sequenza può essere considerata palindromica) e l'eliminazione di

quest'ultimo gene potrebbe rendere RRPE disponibile esclusivamente per la trascrizione di *SNR31*, spiegando così l'aumento di trascrizione che si osserva nel caso del mutante *SNR31 -220*.

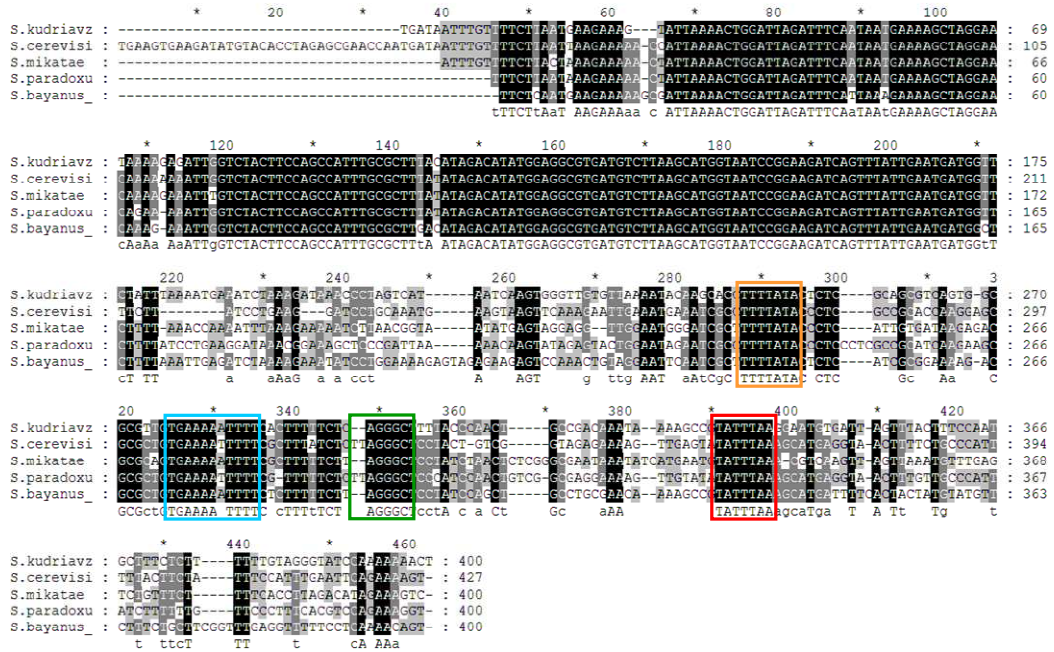


Figura 4.3: allineamento delle regioni a monte di *SNR31* in cinque specie di Saccaromiceti; gli elementi conservati sono evidenziati con colori differenti: in rosso la TATA box, in verde l'ipotetico sito di legame per Tbf1p, in azzurro RRPE e in arancione la TATA box di *SNR5*.

A

wt
TGAAGTGAAGATATGTACACCTAGAGCGAACCAATGATAATTTGTTTTCTTAATTAAG
AAAAACCATTAAACTGGATTAGATTTCAATAATGAAAAGCTAGGAACAAAAAATTT
GGTCTACTTCCAGCCATTTGCGCTTTATATAGACATATGGAGGCGTGATGTCTTAAGC
ATGGTAATCCGGAAGATCAGTTTATTGAATGAT⁻²²⁰GGTTTTCTTATCCTGAAGGATCCTG
CAAATGAAGTAAGTTCAAAGAATTGAAATGAAATCGCCTTTTATACCCTCGCCGCACC
AAGGAGCGCGCTGTGAAAATTTTTCGCTTTATCTCTTAGGGCTCCTACTGTCGGTAGA
GAAAAGTTGAGTATATTTAAAGCATGAGGTAACTTTTCTGCCCATTTTTACTTCTATT
TCCATTTGAATTCAGAAAAGT⁻⁹

-124/RRPEd -107
gcg acc

B

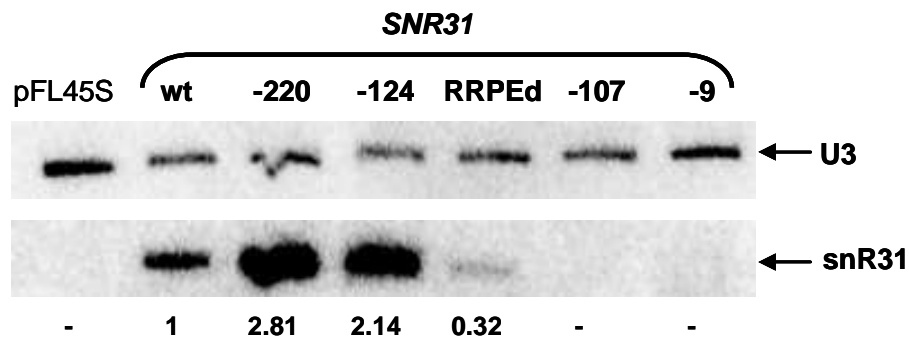


Figura 4.4: A: la regione a monte di *SNR31* in *S. cerevisiae*. Sono riportati i costrutti analizzati (la prima base di ciascuno è in arancione); la sequenza sottolineata corrisponde a *SNR5*; la TATA box di *SNR31*, il motivo RRPE, la sequenza uguale al sito di legame per Tbf1p e la TATA box di *SNR5* sono evidenziati (stessi colori utilizzati nella figura 4.3); le mutazioni introdotte per ottenere *SNR31 RRPE down* sono riportate al di sotto di RRPE. B: analisi della trascrizione di *SNR31* nei diversi costrutti mediante Northern blot; i nomi dei costrutti sono riportati al di sopra delle corsie (pFL45S: plasmide vuoto). U3 è stato utilizzato come controllo. Le quantificazioni sono riportate al di sotto di ciascuna corsia; il valore ottenuto per il wt è stato fissato come pari a 1.

4.1.4 Analisi della regione a monte di *SNR32*

L'allineamento delle 400 bp a monte di *SNR32* nelle cinque specie di Saccaromiceti è mostrato nella figura 4.5.

Anche nel caso di *SNR32* si sono inizialmente preparati quattro costrutti (*SNR32 wt*, *SNR32 -224*, *SNR32 -100*, *SNR32 -7*), successivamente trasformati nel ceppo YS670 delecto per *SNR32* (Balakin et al., 1993). Dall'analisi Northern blot (figura 4.6 B) risulta un'efficiente trascrizione dei primi due costrutti, mentre non si osserva trascrizione

negli altri due casi (malgrado la presenza della TATA box nel mutante *SNR32 -100*). Nella regione tra 100 e 224 bp a monte del gene sono presenti un elemento molto simile al sito di legame per Reb1p (-38 bp dalla TATA box) e un elemento AGCCCTAC (-71 bp dalla TATA box) (figura 4.6 A); con un ulteriore costrutto viene eliminato il sito più distante dal gene (costrutto *SNR32 -148*): la trascrizione è ancora presente a livelli paragonabili agli altri casi, quindi l'ipotetico sito di legame per Nrg1p (GCCCTAC) non sembra influire sulla trascrizione. A questo punto si preparano altri due mutanti che escludono tutta la porzione a monte rispetto al sito per Reb1p: questo elemento è mantenuto in un caso (*SNR32 -127*) e mutato nell'altro (*SNR32 Reb down*): in entrambi i casi è ancora presente la trascrizione del gene. L'attenzione viene quindi posta a una sequenza poliA presente subito a valle del sito per Reb1p e conservata tra i Saccaromiceti (figure 4.6 A e 4.5). Per verificare un eventuale coinvolgimento della sequenza poliA nella trascrizione di *SNR32* vengono preparati due ulteriori mutanti, i quali presentano sempre 127 bp a monte del gene ma sono entrambi privi della sequenza poliA, in combinazione con la sequenza per Reb1p wt (costrutto *SNR32 poliA down*) o mutata (costrutto *SNR32 Reb down-poliA down*): in entrambi i casi la trascrizione è abolita, per cui si può concludere che la sequenza poliA ha un ruolo fondamentale nella trascrizione di *SNR32*, mentre il sito di legame per Reb1p non è richiesto per la trascrizione (figura 4.6 B). Forse Reb1p è coinvolto nell'organizzazione del genoma in domini indipendenti di espressione genica: è noto come questa proteina, così come Tbf1p, sia responsabile della funzione di "insulator" osservata per alcuni promotori (Fourel *et al.*, 2001).

Figura 4.6: A: la regione a monte di *SNR32* in *S. cerevisiae*. Sono riportati i costrutti analizzati (la prima base di ciascuno è in arancione); la TATA box, l'elemento GCCCTAC, il sito di legame per Reb1 (AGGGTAA) e la sequenza poli(A) sono evidenziati (stessi colori utilizzati nella figura 4.5); le mutazioni introdotte per ottenere *SNR32 Reb down*, *SNR32 poliA down* e *SNR32 Reb down-poliA down* sono riportate al di sotto della sequenza; per questi mutanti la prima base è la stessa di *SNR32 -127*. B: analisi della trascrizione di *SNR32* nei diversi costrutti mediante Northern blot; i nomi dei costrutti sono riportati al di sopra delle corsie (pFL45S: plasmide vuoto). U3 è stato utilizzato come controllo. Le quantificazioni sono riportate al di sotto di ciascuna corsia; il valore ottenuto per il wt è stato fissato come pari a 1.

4.1.5 Analisi della regione a monte di *SNR64*

L'allineamento delle regioni a monte di *SNR64* nelle cinque specie di Saccaromiceti è mostrato nella figura 4.8.

L'analisi della regione "upstream" di *SNR64* è complicata dal fatto che lo snoRNA è sintetizzato in forma di precursore e il sito d'inizio della trascrizione non è stato definito in modo preciso. Inizialmente si è deciso di considerare 516 bp a monte rispetto alla prima base del trascritto maturo per creare il costrutto *SNR64 wt*; in più sono stati preparati altri tre costrutti progressivamente deleti, i quali sono stati trasformati nel ceppo YM4585, deleto per *SNR64* (Lowe and Eddy, 1999).

Successivamente è stata effettuata una "primer extension" (figura 4.7) per cercare di identificare il sito d'inizio della trascrizione *in vivo*: dopo tale analisi tuttavia rimane ancora un'incertezza tra tre basi (A -164, A -163, G -162 rispetto alla prima base dello snoRNA maturo). Nella denominazione dei mutanti e nella definizione della posizione degli elementi conservati, per convenzione assumerò come base +1 la A -164; lo stesso criterio è stato utilizzato nell'allineamento mostrato nella figura 4.8. Dall'analisi di omologia risulta che il sito di inizio è situato in una regione, presente solo in *S. cerevisiae*, in cui si ripete la sequenza AAG. L'RNA su cui è stata eseguita la "primer extension" è stato estratto dal ceppo *rnt1Δ xrn1Δ rat1-1* (Danin-Kreiselman *et al.*, 2003), in cui i precursori degli snoRNA sono più stabili, a causa del sistema di processamento difettivo.

I costrutti quindi sono *SNR64 -231*, *SNR64 -89*, *SNR64 +27*, mentre il wt comprende 352 bp a monte del gene. Dal Northern blot (figura 4.9 B) risulta che il mutante *SNR64 -231* è trascritto in modo meno efficiente rispetto al wt mentre *SNR64 -89* è trascritto

più efficientemente; per *SNR64* +27 c'è un crollo della trascrizione, dato non sorprendente dal momento che in tale costrutto è stato eliminato anche il sito d'inizio; negli altri casi è sempre presente una TATA box a -65 bp, mentre in *SNR64* -231 è presente anche un motivo AGCCCTAA a -80 bp dalla TATA box, la cui eliminazione ha un debole effetto positivo sulla trascrizione. Per verificare che effettivamente non ci sono sequenze a monte della TATA box che agiscono come attivatori trascrizionali, viene preparato un altro mutante, *SNR64* -66, in cui è presente solo la TATA box e la regione a valle: in effetti la trascrizione del gene avviene ad alti livelli anche in questo caso. In questo mutante è stato eliminato un sito poli(A), precedentemente identificato (Qu *et al.*, 1999) per il quale quindi può essere escluso un ruolo rilevante nella trascrizione. Il lieve aumento della trascrizione per *SNR64* -89 potrebbe essere spiegato dall'eliminazione di siti di legame per repressori, forse dall'eliminazione proprio del motivo AGCCCTAA; il calo della trascrizione osservato per *SNR64* -231 potrebbe essere spiegato dalla eliminazione di uno o più siti di legame per attivatori, la cui presenza non emerge dall'analisi di omologia. Un sito imperfetto per il repressore Sko1p è presente a monte del gene ma è escluso in tutti i costrutti preparati.

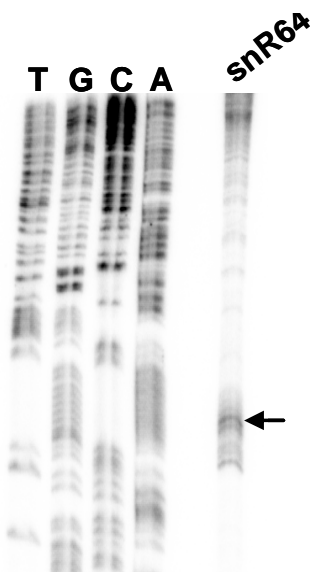


Figura 4.7: determinazione della prima base trascritta per *SNR64*, mediante primer extension.

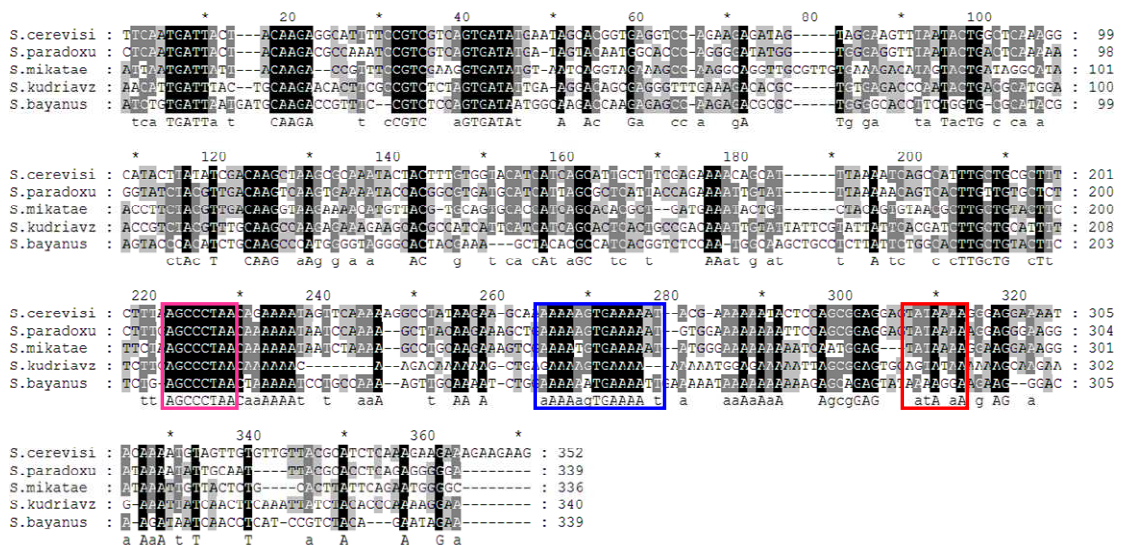


Figura 4.8: allineamento delle regioni a monte di *SNR64* in cinque specie di Saccaromiceti; gli elementi conservati sono evidenziati con colori differenti: in rosso la TATA box, in blu la sequenza poli(A) e in rosa l'ipotetico sito di legame per Nrg1p.

A

wt
TTCAATGATTACTACAAGAGGCATTTTCCGTCGTCAGTGATATGAATAGCACGGTGAGG
TCCAGAAGAGATAGTAGGAAGTTTAATACTGGCTCAAAGGCATACTTATATCGACAAGC
-231
TAAG**G**CGCAAATACTACTTTTGTGGTACATCATCAGCATTGCTTTCGAGAAAACAGCATTT
AAAATCAGCCATTTGCTGCGCTTTCTTTAAG**CCCTAAC**CAGAAAATAGTTCAAAAAGGCC
-89 **-66**
TATAAGAA**GCAAAAAA**AGTGAAAAATACGAAAAATACTCCAGCGGAGGA**GTATAAAA**G
GGAGGAAAATACAAAATGTAGTTGTGTGTTACGCATCTCAAAGAAGAAAGAAG

B

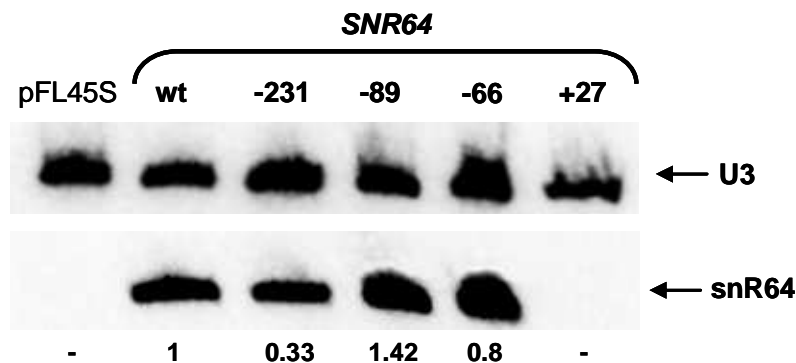


Figura 4.9: A: la regione a monte di *SNR64*. Sono riportati i costrutti analizzati (la prima base di ciascuno è in arancione), ad eccezione di *SNR64 +27*; la TATA box, il putativo sito di legame per Nrg1p e la sequenza poli(A) sono evidenziati (stessi colori utilizzati nella figura 4.8). B: analisi della trascrizione di *SNR64* nei diversi costrutti mediante Northern blot; i nomi dei costrutti sono riportati al di sopra delle corsie (pFL45S: plasmide vuoto). U3 è stato utilizzato come controllo. Le quantificazioni sono riportate al di sotto di ciascuna corsia; il valore ottenuto per il wt è stato fissato come pari a 1.

4.1.6 Analisi della regione a monte di *SNR65*

L'allineamento delle regioni a monte di *SNR65* nelle cinque specie di Saccaromiceti è mostrato nella figura 4.11 (il sito d'inizio della trascrizione è stato definito mediante "primer extension", come indicato di seguito).

Anche *snR65* è sintetizzato come precursore e anche in questo caso non è stato definito il sito d'inizio della trascrizione, quindi si sono considerate, per il costrutto *SNR65 wt*, 529 bp a monte della sequenza codificante e successivamente si sono preparati tre costrutti progressivamente deleti, i quali sono stati trasformati nel ceppo YM4585, deleto per *SNR65* (Lowe and Eddy, 1999).

In seguito si è fatta una "primer extension" per cercare di stabilire il sito d'inizio di *SNR65* (figura 4.10): come per *SNR64* resta un'incertezza tra tre basi (A -117, A -116, T -115 rispetto al trascritto maturo). Per convenzione assumo A -117 come base +1 e in base a questo i costrutti risultano essere definiti come *SNR65 -219*, *SNR65 -98*, *SNR65 +7*, mentre *SNR65 wt* comprende 412 bp a monte del gene. Dal Northern blot (figura 4.11 B) risulta che la trascrizione nel costrutto *SNR65 -219* è più efficiente che nel costrutto wt: forse viene rimosso un sito di legame per un repressore, anche se non è nota la presenza di tali siti a monte di *SNR65*. Procedendo con le delezioni si osserva, per il costrutto *SNR65 -98*, un livello di trascrizione simile al wt; in tale costrutto è stato eliminato il motivo AACCCCTAA (-67 bp rispetto alla TATA box), che quindi, in questo caso, sembrerebbe avere un ruolo nell'attivazione della trascrizione. Per verificare che non ci siano elementi richiesti per la trascrizione a monte della TATA box, viene preparato un costrutto aggiuntivo, *SNR65 -59*, in cui è deleta tutta la regione a monte della TATA box: la trascrizione è ancora paragonabile al costrutto *SNR65 -98* e al wt, quindi sembra che la presenza della sola TATA box sia sufficiente per garantire una apprezzabile trascrizione di *SNR65*.

Un sito di legame per Abf1p è stato precedentemente identificato a monte della TATA box (Qu *et al.*, 1999) ma questo sito non è stato evidenziato dalla nostra analisi di omologia (inoltre non viene riportato tra i siti di legame per proteine a monte di *SNR65* nel database SGD); nello stesso studio viene identificata anche una sequenza poli(A) che è chiaramente riconoscibile a monte della TATA box (figura 4.5 B) in *S. cerevisiae*, ma che risulta conservata solo in alcuni degli altri Saccaromiceti considerati e inoltre non sembra essere coinvolta nella trascrizione (la sua eliminazione nel mutante *SNR65 -59* non provoca cambiamenti nei livelli di espressione).

Dall'analisi dei due geni per snoRNA C/D box sembra emergere che solo la TATA box è richiesta per la trascrizione; non si può escludere la presenza di motivi di sequenza importanti a valle della TATA box, ma è anche possibile ipotizzare, per questi snoRNA, un legame tra il sistema di trascrizione e l'apparato di processamento dello snoRNA, e quindi un controllo a livello post-trascrizionale piuttosto che un controllo della trascrizione.

Non si può escludere che gli elementi di sequenza identificati possano essere importanti per la trascrizione in condizioni di crescita differenti da quelle saggiate negli esperimenti precedentemente riportati.

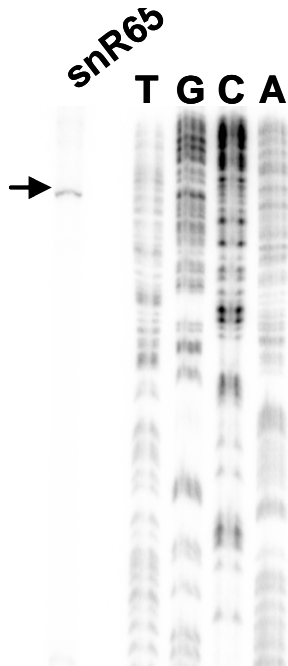


Figura 4.10: determinazione della prima base trascritta per *SNR65*, mediante primer extension.

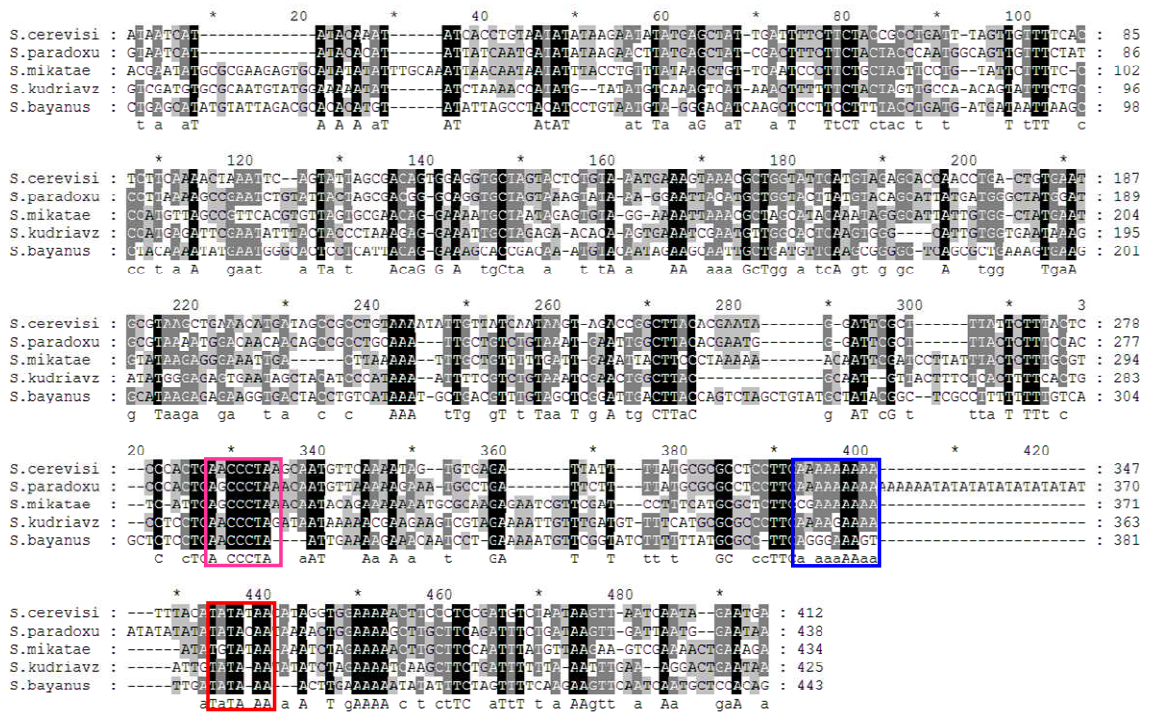


Figura 4.11: allineamento delle regioni a monte di *SNR65* in cinque specie di Saccaromiceti; gli elementi conservati sono evidenziati con colori differenti: in rosso la TATA box, in blu la sequenza poli(A) e in rosa l'ipotetico sito di legame per Nrg1p.

A

wt
ATAATCATATACAAATATCACCTGTAATATATAAGAATATATGAGCTATTGATTTTCTTC
TACCGCCTGATTTAGTTGTTTTCACTCTTCAAACTAAATTCAGTATTAGCGACAGTGGA
GGTGCTAGTACTCTGTAAATGAAAGTAAACGCTGGTATTCATGTAGAGCACCAACCTGAC
TGTGAATGCGTAA**G**CTGAAACATGATAGCCGCCTGTAAAATATTGTTATCAATAAGTAGA
CCGGCTTACACGAATAGGATTCGCTTTATTCTTTACTCCCCACTCA**ACCCTAA**GCAATGT
TCAAAATAGTGTGA**G**ATTATTTTATGCGCGCCTCCTTC**AAAAAAAA**TTTACAT**TATATAA**
CATAGGTGGAAAACTTCCCTCCGATGTCTAATAAGTTAATCAATAGAATGA

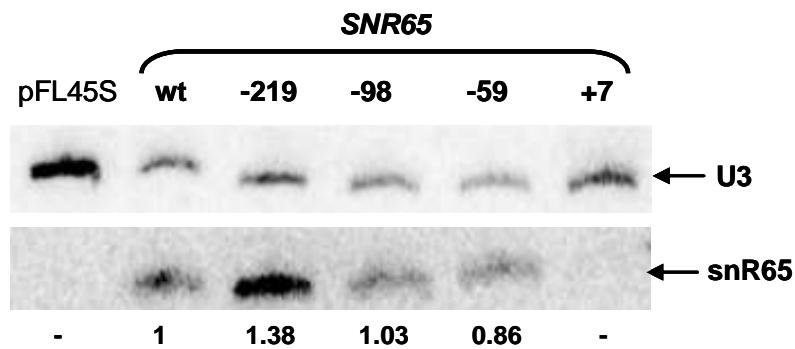
B

Figura 4.12: A: la regione a monte di *SNR65*. Sono riportati i costrutti analizzati (la prima base di ciascuno è in arancione), ad eccezione di *SNR65* +7; la TATA box, il putativo sito di legame per Nrg1p e la sequenza poli(A) sono evidenziati (stessi colori utilizzati nella figura 4.11). B: analisi della trascrizione di *SNR65* nei diversi costrutti mediante Northern blot; i nomi dei costrutti sono riportati al di sopra delle corsie (pFL45S: plasmide vuoto). U3 è stato utilizzato come controllo. Le quantificazioni sono riportate al di sotto di ciascuna corsia; il valore ottenuto per il wt è stato fissato come pari a 1.

4.1.7 Legame di proteine a monte dei geni per snoRNA

Mediante ChIP si è verificato il legame di alcune proteine nella regione “upstream” dei geni per snoRNA. Le proteine inizialmente prese in esame sono Reb1p, Rap1p e Hmo1p: quest’ultima è stata presa in considerazione in quanto in precedenti studi è stata trovata legata a monte di geni per proteine ribosomiali e lungo il gene per l’rRNA (Hall *et al.*, 2006); il legame ai promotori per proteine ribosomiali dipende dalla presenza di Rap1p e in misura minore, di Fhl1p. Le regioni “upstream” prese in considerazione sono quelle di *SNR32*, *SNR42* e *SNR45*: nelle ultime due il consenso per Reb1p è perfetto mentre nel caso di *SNR32* è presente un “mismatch” (AGGGTAA).

Gli allineamenti per le regioni a monte di *SNR42* e *SNR45* sono riportati nell'Appendice.

Nelle condizioni dell'esperimento, per Reb1p si ottiene un arricchimento di 5-7 volte rispetto alla regione di controllo (in cui è possibile comunque ritrovare una sequenza AGGGTAA), mentre l'arricchimento per Rap1p è solo di 2 volte nel caso della regione a monte di *SNR42* (nella quale non è stato però individuato il sito di legame per Rap1p); il segnale nel caso di Hmo1p non si discosta dal "background" in tutti i casi (figura 4.13).

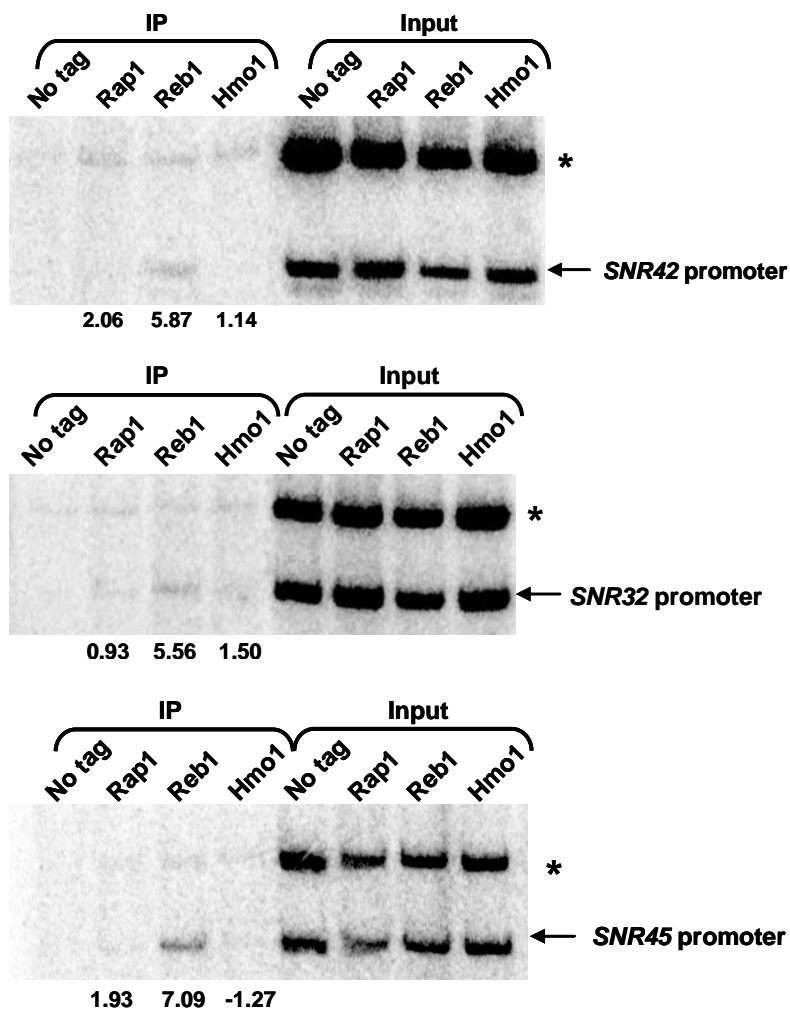


Figura 4.13: saggi di ChIP per la verifica del legame delle proteine Rap1p, Reb1p e Hmo1p alle regioni a monte dei geni *SNR42*, *SNR32* e *SNR45*. Come controllo intergenico interno si è scelta una regione compresa tra i geni *PST2* e *MRH1* (cromosoma IV); la banda corrispondente all'amplicone della regione di controllo è indicata con un asterisco. La quantificazione è espressa come arricchimento rispetto al "background" ed è riportato al di sotto delle corsie. IP: cromatina immunoprecipitata, Input: cromatina non immunoprecipitata.

Successivamente si sono considerate Nrg1p e Tbf1p: la prima potrebbe legare il motivo GCCCTAA mentre la seconda lega il complementare invertito TTAGGG: nelle condizioni in cui è stata condotta la ChIP sembra che Nrg1p non leghi la regione a monte dei geni analizzati in modo specifico, in quanto non si ottiene un arricchimento rispetto alla regione di controllo, mentre si ottiene un arricchimento di poco superiore a 2 per Tbf1p. Nella figura 4.14 è riportato uno dei risultati di questi saggi di ChIP.

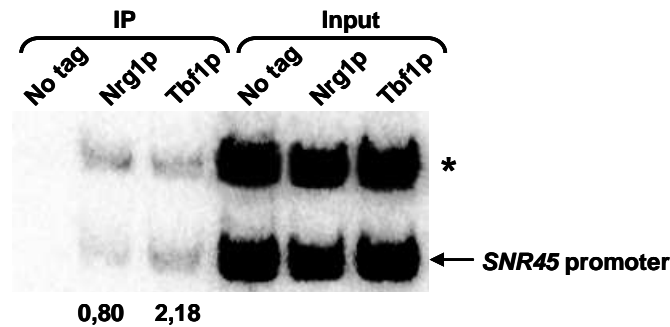


Figura 4.14: saggio di ChIP per la verifica del legame delle proteine Nrg1p e Tbf1p alla regione a monte del gene *SNR45*. La banda ottenuta per il controllo intergenico interno è indicata con un asterisco. La quantificazione è stata determinata come nella figura 4.13. IP: cromatina immunoprecipitata, Input: cromatina non immunoprecipitata.

Anche il legame di Stb3p ai promotori contenenti RRPE è stato saggiato in varie condizioni: nel caso di crescita delle cellule in terreno contenente glucosio, in terreno contenente glicerolo e in terreno contenente glicerolo addizionato successivamente con glucosio. È stato osservato infatti di recente che la mancanza di Stb3p provoca la mancata induzione da glucosio di una larga parte dei geni che presentano il motivo RRPE (Liko *et al.*, 2007) quindi è ipotizzabile che la proteina si leghi solo quando il glucosio viene aggiunto ad una coltura contenente altre fonti di carbonio. La raccolta dopo l'aggiunta di glucosio viene effettuata dopo 20 minuti in quanto è stato osservato che a questo tempo si ha la massima induzione dei geni coinvolti nella biogenesi del ribosoma (Wang *et al.*, 2004). In nessuna condizione si nota un arricchimento (figura 4.15), sia a monte di *SNR31* sia a monte di *SNR45*.

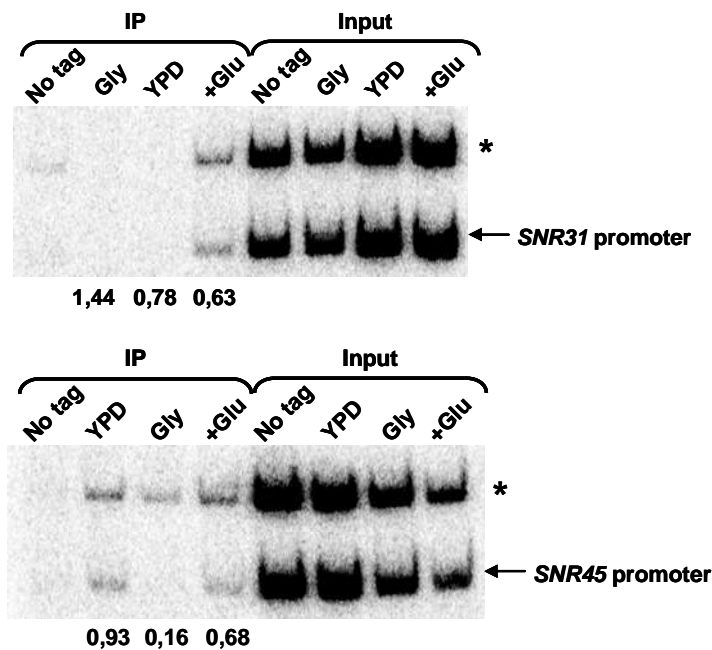


Figura 4.15: saggio di ChIP per la verifica del legame della proteina Stb3p alle regioni a monte dei geni *SNR45* e *SNR31*. La banda ottenuta per il controllo intergenico interno è indicata con un asterisco. La quantificazione è stata determinata come nella figura 4.13. IP: cromatina immunoprecipitata; Input: cromatina non immunoprecipitata; Gly: cellule cresciute in YP+3% glicerolo; YPD: cellule cresciute in YPD; +Glu: cellule cresciute in YP+3% glicerolo addizionato con 2% glucosio e raccolte dopo 20 minuti.

4.1.8 Espressione degli snoRNA in particolari condizioni di crescita

Non sono attualmente conosciute delle condizioni di crescita in cui la trascrizione dei geni per snoRNA risulti attivata oppure repressa; tuttavia si possono individuare delle condizioni in cui si osserva la repressione della trascrizione dei geni ospiti degli snoRNA intronici. Spesso negli stessi casi risultano repressi anche i geni per le proteine delle snoRNP (Gasch *et al.*, 2000).

Una marcata repressione dei geni citati sopra si osserva durante la fase stazionaria, per cui l'espressione di uno snoRNA, snR65, viene studiata mediante Northern blot, al fine di verificare se anche gli snoRNA subiscono una repressione nella stessa condizione di crescita: i livelli di snoRNA maturo sembrano aumentare leggermente al raggiungimento della fase stazionaria (figura 4.16). L'aumento dell'intensità della banda di snR65 che si osserva per gli ultimi tre tempi dell'esperimento potrebbe essere dovuto al caricamento della stessa quantità di RNA totali su gel: poiché in fase stazionaria si ha un forte calo degli rRNA (Dammann *et al.*, 1993), i quali

rappresentano l'80% dell'RNA presente in una cellula (Warner, 1999), si tende a caricare una maggiore quantità degli altri RNA, per raggiungere la stessa quantità di RNA totali.

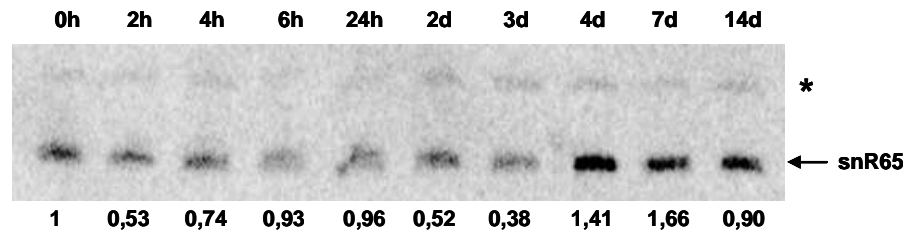


Figura 4.16: analisi dell'espressione dello snoRNA snR65 in fase stazionaria; i tempi a cui è stata effettuata ciascuna raccolta sono riportati al di sopra di ciascuna corsia. L'asterisco indica l'rRNA 5S. Le quantificazioni sono riportate al di sotto di ciascuna corsia; il valore ottenuto al tempo 0 è stato fissato come pari a 1.

Anche in carenza d'azoto si ha repressione dei geni citati in precedenza: analizzando la trascrizione di *SNR65* si osserva un aumento dell'intensità della banda di interesse per gli ultimi due tempi dell'esperimento: probabilmente in queste condizioni si ha un aumento della trascrizione di snR65. In precedenti lavori (Basehoar *et al.*, 2004) è stato notato che i geni di lievito che presentano una TATA box nella regione del promotore tendono a essere indotti in condizioni di stress. Se si analizza un gene intronico (*SNR24*), invece, sembra che non si abbiano variazioni nella quantità di snoRNA, neanche dopo 3 o 4 giorni in carenza d'azoto (figura 4.17); *SNR24* è contenuto nell'introne di *ASCI*, la cui trascrizione è repressa in carenza d'azoto (Gasch *et al.*, 2000). È necessario comunque ricordare che, deprivando le cellule di azoto, i livelli di trascrizione da parte delle RNA polimerasi I e III diminuiscono (Clarke *et al.*, 1996; Oliver and McLaughlin, 1977), anche se, a differenza del precedente esperimento, non ci sono state grosse variazioni nella quantità di RNA totale estratto dai campioni raccolti ai differenti tempi.

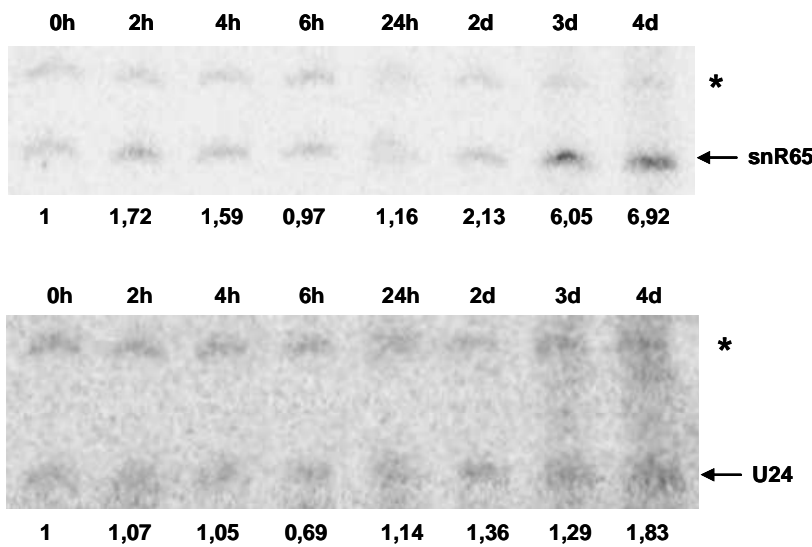


Figura 4.17: analisi dell'espressione dello snoRNA snR65 (pannello in alto) e dello snoRNA U24 (pannello in basso) in carenza di azoto; i tempi a cui è stata effettuata ciascuna raccolta sono riportati al di sopra di ciascuna corsia. L'asterisco indica l'rRNA 5S. Le quantificazioni sono riportate al di sotto di ogni corsia; il valore ottenuto al tempo 0 è stato fissato come pari a 1.

È stata anche presa in considerazione l'aggiunta di glucosio ad una coltura contenente glicerolo come fonte di carbonio (non fermentabile): in seguito all'aggiunta di glucosio è stata precedentemente osservata l'induzione di geni che presentano il motivo RRPE. Tuttavia la trascrizione del gene *SNR31* non sembra variare dopo l'aggiunta di glucosio al terreno (figura 4.18).

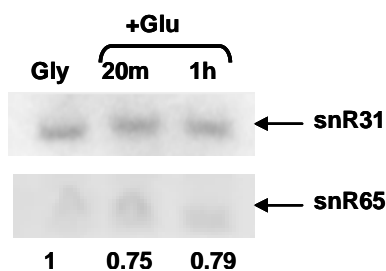


Figura 4.18: analisi dell'espressione dello snoRNA snR31 in seguito all'aggiunta di glucosio alla coltura; snR65 è stato utilizzato come controllo. Le quantificazioni sono riportate al di sotto di ogni corsia; il valore ottenuto al tempo 0 è stato fissato come pari a 1.

4.2 Maturazione dello snoRNA snR52

4.2.1 Il gene SNR52

Il gene *SNR52* è l'unico gene per snoRNA trascritto dalla RNA polimerasi III finora identificato in *S. cerevisiae* (Harismendy *et al.*, 2003; Moqtaderi and Struhl, 2003). Il promotore è costituito da una A e una B box, posizionate nella sequenza "leader" dello snoRNA, e da una TATA box localizzata a monte del sito di inizio della trascrizione (Harismendy *et al.*, 2003). La distanza tra le due box (86 bp) è notevolmente più ampia di quella osservata per i tRNA di lievito e tra esse è presente una sequenza costituita da sei T: solitamente una ripetizione di questo tipo funziona come un forte terminatore per la RNA polimerasi III in lievito, mentre nel contesto di *SNR52* agisce come un terminatore debole; tuttavia questo elemento è presente anche in altri Saccaromiceti (*S. paradoxus* e *S. mikatae*), cosa che permette di ipotizzare un ruolo nell'espressione di *SNR52* (Braglia *et al.*, 2005).

Sia la A che la B box sono richieste per la trascrizione del gene *in vivo* mentre sembra che la B box non sia indispensabile per la trascrizione *in vitro* (Guffanti *et al.*, 2006). Mediante analisi "DNase footprinting" si è osservata la protezione dalla digestione della B box, ma non della A box, da parte di TFIIC: questo è probabilmente dovuto alla considerevole distanza tra le due box; mutazioni nella B box provocano la perdita della protezione, mentre mutazioni nella A box causano una riduzione della protezione della B box: questo permette di ipotizzare che l'inattivazione della trascrizione causata da mutazioni nella A box non sia dovuta da una riduzione del reclutamento di TFIIC, mentre mutazioni nella B box permettono la formazione del complesso TFIIC-TFIIB e il successivo reclutamento della RNA polimerasi III *in vitro*, nonostante la stabilità del complesso sia molto ridotta rispetto al wt (Guffanti *et al.*, 2006).

Mutazioni nella TATA box e in una sequenza "T rich" presente poche basi a monte della TATA box non hanno ripercussioni sulla trascrizione, mentre si osserva una riduzione se viene rimossa la regione a monte del gene, sia in un contesto in cui il gene *SNR52* è "wild type", sia nel caso in cui la B box è mutata o rimossa: questo indica che la A box costituisce un elemento del promotore che può dirigere la trascrizione della RNA polimerasi III in modo autonomo; la riduzione della trascrizione nel caso della rimozione della regione a monte di *SNR52* può essere dovuta alla sequenza plasmidica

che viene a trovarsi a fianco del gene, oppure nella regione rimossa sono presenti elementi attivatori differenti dalla TATA box e dalla sequenza “T rich”(Guffanti *et al.*, 2006).

L’organizzazione del gene è rappresentata nella figura 4.19.

```

CTTTGAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGAACTTGATGTTT
TCTTTCGAGTATATACAAGGTGATTACATGTACGTTTGAAGTACAACTAGATTTTGTA
GTGCCCTCTTGGGCTAGCGGTAAAGGTGCGCATTTTTTCACACCCTACAATGTTCTGTTC
AAAAGATTTTGGTCAAACGCTGTAGAAGTGAAAGTTGGTGCGCATGTTTCGGCGTTCGAA
ACTTCTCCGCAGTGAAAGATAAATGATCTACTATGATGAATGACATTAGCGTGAACAATC
TCTGATACAAAATCGAAAGATTTTAGGATTAGAAAACTTATGTTGCCCTTCCTTCTGAA
TTTATTTTTTTTGTCACTATTGTTATGTAAAATGCCACCTCTGACAGTATGGAACGC
  
```

TTTCTTT: T rich
TATATA: TATA box
A: sito d’inizio
TTGGGCTAGCGGTAAAGG: A box
TTTTTT: T box
CGTTCGAAAC: B box
TTTTTT: terminatore

Figura 4.19: il gene *SNR52* e la sua regione a monte. Sono evidenziati gli elementi legati alla trascrizione (A box, B box, sito d’inizio, terminatore) o comunque analizzati in Guffanti *et al.*, 2006 (T rich, TATA box, T box). La sequenza codificante per lo snoRNA maturo è sottolineata e in grassetto.

4.2.2 Ruolo della C box e della D box nella biosintesi di snR52

Lo snoRNA C/D box snR52 viene trascritto come precursore lungo 250 nt da cui viene poi rimossa la sequenza in 5’ per dare origine al trascritto maturo di 92 nt.

Come già ricordato nel paragrafo 1.3, le box C e D sono coinvolte nella maturazione degli snoRNA trascritti dalla RNA polimerasi II: una parte del mio lavoro di tesi è stata dedicata all’analisi, mediante mutagenesi sito-specifica, del ruolo delle box C e D nella maturazione di snR52, la cui trascrizione avviene ad opera di un sistema differente, quello della RNA polimerasi III. Inoltre, si ha una precisa terminazione dei trascritti e

di precursore (94% della quantità totale), ma in questo caso non si osserva una riduzione globale della quantità dello snoRNA (figura 4.21). L'accumulo del precursore di *SNR52* nel caso del mutante *SNR52 D-less* è stato osservato anche mediante "primer extension", grazie alla quale viene anche confermato che l'estremità 5' di *SNR52 D-less* è la stessa prevista per il trascritto wt (figura 4.23), quindi la rimozione della D box non altera il sito di inizio della trascrizione.

Per spiegare questi risultati si può ipotizzare che la presenza di una delle due box mutate favorisca la degradazione dello snoRNA, quando è ancora in forma di precursore, da parte delle nucleasi, degradazione non favorita invece, quando è rimossa una sequenza ed è presente la sola box C. Forse questa rimozione interferisce con il processamento in 5' da parte dell'endonucleasi Rnt1p, la cui interazione con Nop1p è stata precedentemente dimostrata (Giorgi *et al.*, 2001).

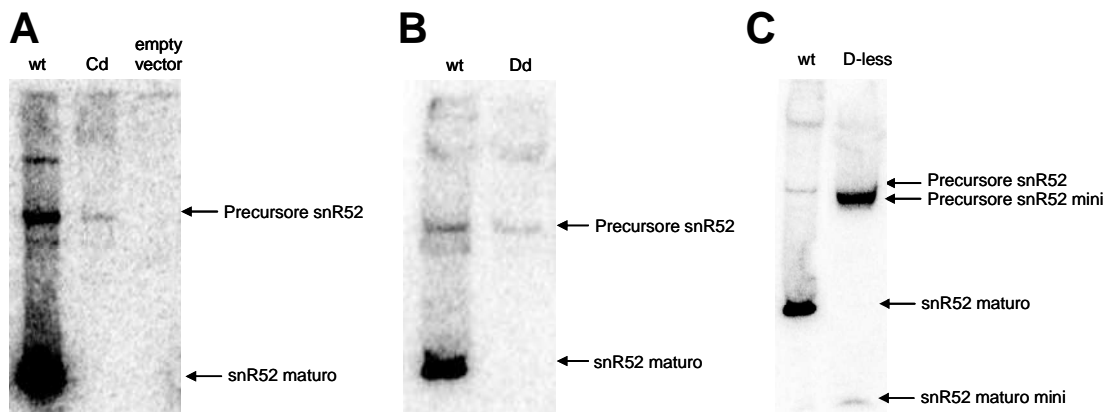


Figura 4.21: analisi della trascrizione di snR52 e delle sue versioni mutate, mediante Northern blot. A: analisi della trascrizione del mutante Cdown (Cd) rispetto a snR52 wt; B: analisi della trascrizione del mutante Ddown (Dd); C: analisi della trascrizione del mutante D-less.

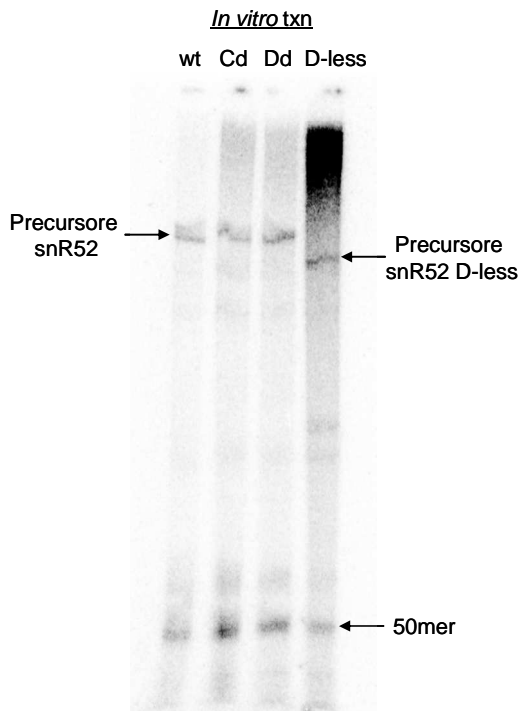


Figura 4.22: saggio di trascrizione *in vitro* per i templati *SNR52 wt* e le versioni mutate (i nomi sono indicati sopra ciascuna corsia).

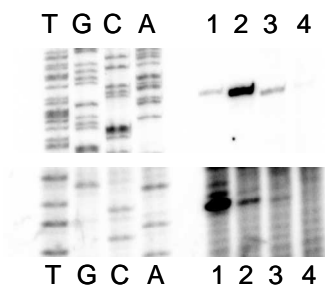


Figura 4.23: saggio di “primer extension” per i trascritti codificati da *SNR52 wt* e dalle versioni mutate.

1: *snR52 wt*; 2: *snR52 D-less*; 3: *snR52 Ddown*; 4: *snR52 Cdown*

In alto sono riportate le bande relative ai precursori, in basso le bande relative ai trascritti maturi.

4.2.3 Ruolo delle proteine della snoRNP nella maturazione di snR52

La forte instabilità di snR52 nei mutanti *SNR52 Cdown* e *SNR52 Ddown* potrebbe essere dovuto a problemi nella formazione della snoRNP: quindi anche la mancanza delle proteine del “core” potrebbe alterare il processamento di snR52. Per verificare tale ipotesi sono stati utilizzati ceppi di lievito in cui i geni *NOPI*, *NOP58* o *SNU13* sono sotto il controllo di un promotore che permette la trascrizione solo in presenza di galattosio nel terreno di coltura (Tollervey *et al.*, 1991; Lafontaine and Tollervey, 1996; Dobbyn and O’Keefe, 2004): in questo contesto è possibile spegnere in modo controllato i geni d’interesse, che sono tutti essenziali.

Mediante Northern blot si può osservare come la deplezione di Nop1p o di Nop58p provochi l'accumulo del precursore di snR52 e, nel solo caso di Nop58p, la diminuzione della forma matura (figure 4.24 e 4.25). L'aumento dello snoRNA H/ACA box snR33, visibile nel caso della deplezione di Nop58p, è stato precedentemente osservato (Lafontaine and Tollervey, 1996) ed è forse dovuto a un meccanismo di "crosstalk" tra i sistemi di biosintesi dei due tipi di snoRNP (vari fattori richiesti per l'assemblaggio sono comuni per le due classi: forse, non formandosi correttamente le snoRNP C/D box, c'è una maggiore disponibilità di fattori per l'assemblaggio delle snoRNP H/ACA, per cui una maggiore quantità di snoRNA H/ACA box risulta protetto dalla degradazione). Come già osservato in precedenza (Tollervey *et al.*, 1991), la deplezione di Nop1p non altera la quantità di uno snoRNA C/D box, U3, che è stato quindi utilizzato come controllo in questo esperimento.

La deplezione di Snu13p non altera il processamento in 5' dello snoRNA (figura 4.26), al contrario delle proteine viste in precedenza. Questo dato risulta essere inatteso in quanto è stato dimostrato che Snu13p è la prima proteina che si lega allo snoRNA ed è richiesta per il legame delle altre proteine del "core" (Watkins *et al.*, 2002); si potrebbe ipotizzare la presenza di una proteina che lega snR52, successivamente rimpiazzata da Snu13p e che, in assenza di quest'ultima, resta legata al trascritto e lo protegge dalla degradazione, allo stesso tempo permettendo una normale maturazione del trascritto primario.

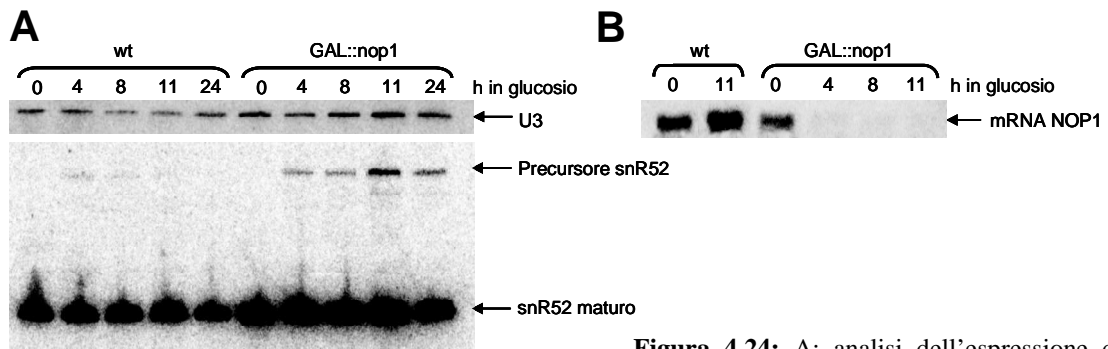


Figura 4.24: A: analisi dell'espressione di snR52 in risposta alla deplezione di Nop1p, mediante Northern blot (U3 è utilizzato come controllo); B: controllo dell'effettivo spegnimento del gene *NOP1* in risposta alla crescita in terreno contenente glucosio; C: incremento della quantità del precursore di snR52 rispetto alla forma matura nel ceppo wt e nel ceppo GAL::nop1.

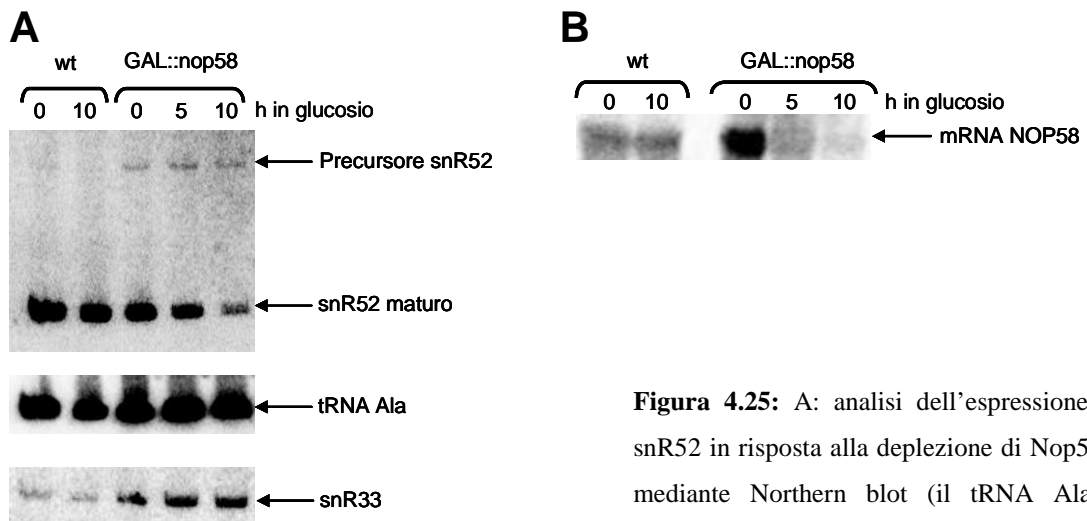
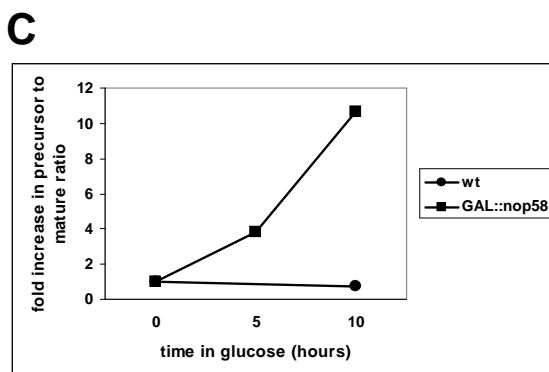
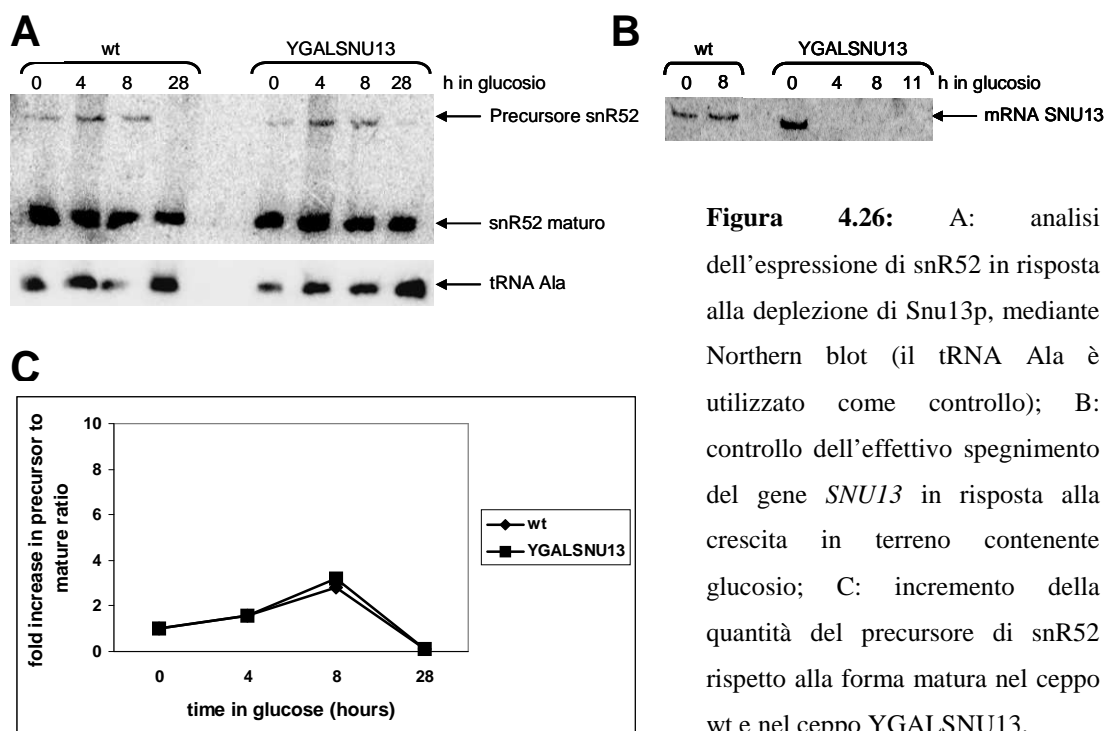


Figura 4.25: A: analisi dell'espressione di snR52 in risposta alla deplezione di Nop58p, mediante Northern blot (il tRNA Ala è utilizzato come controllo). È mostrato anche l'incremento della quantità dello snoRNA snR33 nel ceppo GAL::nop58; B: controllo dell'effettivo spegnimento del gene *NOP58* in risposta alla crescita in terreno contenente glucosio; C: incremento della quantità del precursore di snR52 rispetto alla forma matura nel ceppo wt e nel ceppo GAL::nop58.





Probabilmente Nop1p e Nop58p sono direttamente coinvolte nel processamento dell'estremità 5' di snR52, forse facilitando il taglio effettuato dall'endonucleasi Rnt1p, per la quale è già stata dimostrata l'interazione con Nop1p (Giorgi *et al.*, 2001).

I risultati ottenuti depongono a favore dell'accoppiamento tra processamento dello snoRNA e assemblaggio della snoRNP, come già osservato per gli altri snoRNA; visto il risultato ottenuto in assenza di Snu13p, non è però possibile escludere qualche variazione nella sequenza di reclutamento delle proteine del "core".

Dagli esperimenti si può notare come il rapporto precursore/maturo osservato nei ceppi wt sia variabile: questo può essere dovuto alla ridotta quantità del precursore di snR52, quando il gene è presente in singola copia, cosa che rende difficile avere delle quantificazioni accurate.

4.2.4 Maturazione dell'estremità 3' di snR52

Dagli esperimenti precedenti non sono emerse alterazioni nella maturazione dell'estremità 3' di snR52 quando le box C o D vengono mutate. La terminazione dei trascritti pol III avviene in modo preciso ad una sequenza di 5 o più T (Braglia *et al.*, 2005) e, in seguito, la coda poliU presente nell'RNA viene rimossa; l'apparato necessario per il processamento delle estremità 3' dei trascritti pol III è tuttora

sconosciuto. Nei tRNA e in altri trascritti pol III, il tratto ricco in U all'estremità 3' viene legata dalla proteina Lhp1p ("La homologous protein") che influenza la stabilità e la maturazione degli RNA stessi (Wolin and Cedervall, 2002); inoltre è stato osservato come Lhp1p sia associato anche agli snoRNA (Inada and Guthrie, 2004).

La maturazione dell'estremità 3' degli snoRNA richiede la forma nucleare del complesso dell'esosoma, e un ruolo rilevante in questo processo è svolto da Rrp6p (vedi paragrafo 1.13).

Dall'analisi Northern blot effettuata su RNA estratti da ceppi Δ lhp1 e Δ rrp6 emerge che l'assenza di Lhp1p non influenza la maturazione di snR52, mentre l'assenza di Rrp6p causa l'accumulo di una forma estesa in 3' di qualche nucleotide (figura 4.27); quest'ultimo dato era stato precedentemente osservato per altri snoRNA (van Hoof *et al.*, 2000). Anche la presenza di forme poliadenilate di snR52, dovuta al mancato processament in 3', è stata rilevata (figura 4.28), così come osservato in precedenza per altri snoRNA (van Hoof *et al.*, 2000). La poliadenilazione degli snoRNA è dovuta alle poli(A)-polimerasi Trf4p e Trf5p e Rrp6p è responsabile della rimozione della coda poli(A) durante il normale processamento dell'estremità 3' degli snoRNA.

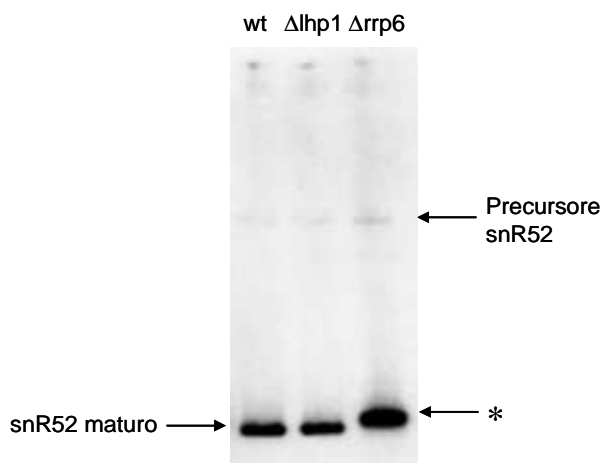


Figura 4.27: analisi di snR52 nei ceppi BY4742 (wt), Δ lhp1 e Δ rrp6 mediante Northern blot. L'asterisco indica snR52 maturo nel ceppo Δ rrp6, in cui le dimensioni dello snoRNA sono maggiori.

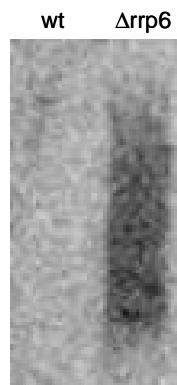


Figura 4.28: rilevazione della forma poliadenilata di snR52 nei ceppi BY4742 (wt) e Δ rrp6 mediante Northern blot. Gli RNA che presentano poliadenilazione in 3' sono stati purificati (vedi Materiali e Metodi, paragrafo 3.12) e quindi caricati su gel.

È noto che l'esosoma abbia un ruolo nella degradazione dei trascritti pol III, probabilmente nel caso in cui questi siano ripiegati in modo scorretto e abbiano subito poliadenilazione da parte di Trf4p (Kadaba *et al.*, 2006); nel caso di snR52, invece, l'esosoma fa parte del "pathway" di processamento, come nel caso degli altri snoRNA trascritti dalla RNA polimerasi II.

La mancanza di Rrp6p non interferisce con il processamento dell'estremità 5', come verificato mediante primer extension (figura 4.29).

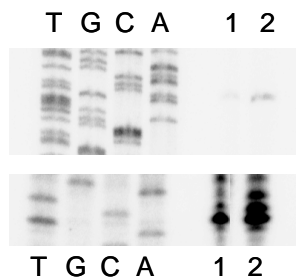


Figura 4.29: saggio di "primer extension" per snR52 nei ceppi wt e Δ rrp6.

1:BY4742 (wt); 2: Δ rrp6. In alto sono riportate le bande relative ai precursori, in basso le bande relative ai trascritti maturi.

5. BIBLIOGRAFIA

Allmang C., Kufel J., Chanfreau G., Mitchell P., Petfalsky E. and Tollervey D. (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.*, **18**, 5399-5410.

Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W. and Lipman D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Res.*, **25**, 3389-3402.

Antal M., Mougín A., Kis M., Boros E., Steger G., Jakab G., Solymosy F. and Branlant C. (2000). Molecular characterization at the RNA and gene level of U3 snoRNA from a unicellular green alga, *Chlamydomonas reinhardtii*. *Nucleic Acid Res.*, **28**, 2959-2968.

Aspegren A., Hinas A., Larsson P., Larsson A. and Söderbom F. (2004). Novel non-coding RNAs in *Dictyostelium discoideum* and their expression during development. *Nucleic Acid Res.*, **32**, 4646-4656.

Bachelier J.P., Cavaille J. and Huttenhofer A. (2002) The expanding snoRNA world. *Biochimie*, **84**, 775-790.

Balakin A.G., Schneider G.S., Corbett M.S., Ni J. and Fournier M.J. (1993). SnR31, snR32, and snR33: three novel, non-essential snRNAs from *Saccharomyces cerevisiae*. *Nucleic Acid Res.*, **21**, 5391-5397.

Ballarino M., Morlando M., Pagano F., Fatica A. and Bozzoni I. (2005). The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **25**, 5396-5403.

Beltrame M., Henry Y. and Tollervey D. (1994). Mutational analysis of an essential binding site for the U3 snoRNA in the 5' external transcribed spacer of yeast pre-rRNA. *Nucleic Acids Res.*, **22**, 5139-5147.

Beltrame M., and Tollervey D. (1995). Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis. *EMBO J.*, **14**, 4350-4356.

- Basehoar A.D., Zanton S.J. and Pugh B.F. (2004). Identification and distinct regulation of yeast TATA box-containing genes. *Cell*, **116**, 699-709.
- Bernstein K.A., Gallagher J.E.J., Mitchell B.M., Granneman S. and Baserga S.J. (2004). The small-subunit processome is a ribosome assembly intermediate. *Eukaryot Cell*, **3**, 1619-1626.
- Bertrand E., Houser-Scott F., Kendall A., Singer R.H. and Engelke D.R. (1998). Nucleolar localization of early tRNA processing. *Genes Dev.*, **12**, 2463-2468.
- Bolchi A., Ottonello S. and Petrucco S. (2005). A general one-step method for the cloning of PCR products. *Biotechnol Appl Biochem.*, **42**, 205-209.
- Bonnerot C., Pintard L. and Lutfalla G. (2003). Functional redundancy of Spb1p and a snR52-dependent mechanism for the 2'-O-ribose methylation of a conserved rRNA position in yeast. *Mol. Cell*, **12**, 1309-1315.
- Borovjagin A.V. and Gerbi S.A. (1999). U3 small nucleolar RNA is essential for cleavage at sites 1, 2 and 3 in pre-rRNA and determines which rRNA processing pathway is taken in *Xenopus* oocytes. *J. Mol. Biol.*, **286**, 1347-1363.
- Braglia P., Percudani R. and Dieci G. (2005). Sequence context effects on oligo(dT) termination signal recognition by *Saccharomyces cerevisiae* RNA polymerase III. *J. Biol. Chem.*, **280**, 19551-19562.
- Brouwer R., Pruijn G.J.M. and van Venrooij W.J. (2001). The human exosome: an autoantigenic complex of exoribonucleases in myositis and scleroderma. *Arthritis Res.*, **3**, 102-106.
- Brown J.W.S., Echeverria M. and Qu L.H. (2003). Plant snoRNAs: functional evolution and new modes of gene expression. *Trends Plant Sci.*, **8**, 42-49.
- Cahill N.M., Friend K., Speckmann W., Li Z.H., Terns R.M., Terns M.P. and Steitz J.A. (2002). Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snoRNP. *EMBO J.*, **21**, 3816-3828.

5. Bibliografia

- Carneiro T., Carvalho C., Braga J. Rino J., Milligan L., Tollervey D. and Carmo-Fonseca M. (2007). Depletion of the yeast exosome subunit Rrp6 results in accumulation of polyadenylated RNAs in a discrete domain within the nucleolus. *Mol. Cell. Biol.*, **27**, 4157-4165.
- Carroll K.L., Ghirlando R., Ames J.M. and Corden J.L. (2007). Interaction of yeast RNA-binding protein Nrd1 and Nab3 with RNA polymerase II terminator elements. *RNA*, **13**, 361-373.
- Cavaillé J., Buiting K., Kiefmann M., Lalande M., Brannan C.I., Horsthemke B., Bachellerie J.P., Brosius J. and Huttenhofer A. (2000). Identification of brain-specific and imprinted. *Proc Natl Acad Sci U S A.*, **19**, 14311-14316.
- Chanfreau G., Legrain P. and Jacquier A. (1998). Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. *J. Mol. Biol.*, **284**, 975-988.
- Chanfreau G., Rotondo G, Legrain P. and Jacquier A. (1998). Processing of a dicistronic small nucleolar RNA precursor *EMBO J.*, **17**, 3726-3737.
- Chen C.L., Liang D., Zhou H., Zhuo M., Chen Y.Q. and Qu L.H. (2003). The high diversity of snoRNAs in plants: identification and comparative study of 120 snoRNA genes from *Oryza sativa*. *Nucleic Acids Res.*, **31**, 2601-2613.
- Chen C.L., Perasso R., Qu L.H. and Amar L. (2007). Exploration of pairing constraints identifies a 9 base-pair core within box C/D snoRNA-rRNA duplexes. *J. Mol. Biol.*, **369**, 771-783.
- Clarke E.M., Peterson C.L., Brainard A.V. and Riggs D.L. (1996). Regulation of the RNA polymerase I and III transcription systems in response to growth conditions. *J. Biol. Chem.*, **271**, 22189-22195.
- Dammann R., Lucchini R., Koller T. and Sogo J.M. (1993). Chromatin structures and transcription of rDNA in yeast *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **21**, 2331-2338.
- Danin-Kreiselman M., Lee C.Y. and Chanfreau G. (2003). RNase III-mediated degradation of unspliced pre-mRNAs and lariat introns. *Mol. Cell*, **11**, 1279-1289.

5. Bibliografia

- Davis C.A. and Ares M.Jr. (2006). Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*, **103**, 3262-3267.
- Decatur W.A. and Fournier M.J. (2002). rRNA modifications and ribosome function. *Trends Biochem. Sci.*, **27**, 344-351.
- Deng W., Zhu X., Skogerbø G., Zhao Y., Fu Z., Wang Y., He H., Cai L., Sun H., Liu C., Li B., Bai B., Wang J., Jia D., Sun S., He H., Cui Y., Wang Y., Bu D. and Chen R. (2006). Organization of the *Caenorhabditis elegans* small non-coding transcriptome: Genomic features, biogenesis, and expression. *Genome Res.*, **16**, 20-29.
- De Turrís V., Di Leva G., Caldarola S., Loreni F., Amaldi F. and Bozzoni I. (2004). TOP promoter elements control the relative ratio of intron-encoded snoRNA versus spliced mRNA biosynthesis. *J. Mol. Biol.*, **344**, 383-394.
- Dheur S., Vo L.T.A., Voisinet-Hakil F., Minet M., Schmitter J.M., Lacroute F., Wyers F., Minvielle-Sebastia L. (2003). Pti1p and Ref2p found in association with the mRNA 3' end formation complex direct snoRNA maturation. *EMBO J.*, **22**, 2831-2840.
- Dieci G., Duimio L., Coda-Zabetta F., Sprague K.U. and Ottonello S. (1993). A novel RNA polymerase III transcription factor fraction that is not required for template commitment. *J. Biol. Chem.*, **268**, 11199-11207.
- Dobbyn H.C. and O'Keefe (2004). Analysis of Snu13p mutations reveals differential interactions with the U4 snRNA and U3 snoRNA. *RNA*, **10**, 308-320.
- Dragon F., Gallagher J.E., Compagnone-Post P.A., Mitchell B.M., Porwancher K.A., Wehner K.A., Wormsley S., Settlage R.E., Shabanowitz J., Osheim Y., Beyer A.L., Hunt D.F., Baserga S.J.(2002). A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature*, **27**, 967-970.
- Dunbar D.A. and Baserga S.J. (1998). The U14 snoRNA is required for 2'-O-methylation of the pre-18S rRNA in *Xenopus* oocytes. *RNA*, **4**, 195-204.

5. Bibliografia

- Fatica A., Morlando M. and Bozzoni I. (2000). Yeast snoRNA accumulation relies on a cleavage-dependent/polyadenylation-independent 3'-processing apparatus. *EMBO J.*, **19**, 6218-6229.
- Fourel G., Boscheron C., Revardel E., Lebrun E., Hu Y.F., Simmen K.C., Müller K., Li R., Mermod N. and Gilson E. (2001). An activation-independent role of transcription factors in insulator function. *EMBO Rep.*, **2**, 124-132.
- Fragapane P., Prislei S., Michienzi A., Caffarelli E. and Bozzoni I. (1993). A novel small nucleolar (U16) is encoded inside a ribosomal protein intron and originates by processing of the pre-mRNA. *EMBO J.*, **12**, 2921-2928.
- Galardi S., Fatica A., Bachi A., Scaloni A. Presutti C. and Bozzoni I. (2002). Purified box C/D snoRNPs are able to reproduce site-specific 2'-O-methylation of target RNA *in vitro*. *Mol. Cell. Biol.*, **22**, 6663-6668.
- Ganot P., Caizergues-Ferrer M., and Kiss T. (1997). The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev* **11**, 941-956.
- Gasch A.P., Spellman P.T., Kao C.M., Carmel-Harel O., Eisen M.B., Storz G., Botstein D. and Brown P.O. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell*, **11**, 4241-4257.
- Gaspin C., Cavallé J., Erauso G. and Bachellerie J.P. (2000). Archaeal homologs of eukaryotic methylation guide small nucleolar RNAs: lessons from the Pyrococcus genomes. *J. Mol. Biol.*, **297**, 895-906.
- Gaudin C., Ghazal G., Yoshizawa S., Abou Elela S. and Fourmy D. (2006). Structure of an AAGU tetraloop and its contribution to substrata selection by yeast RNase III. *J. Mol. Biol.*, **363**, 322-331.
- Gautier T., Berges T., Tollervey D. and Hurt E. (1997). Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. *Mol. Cell. Biol.*, **17**, 7088-7098.

5. Bibliografia

- Ghazal G. and Abou Elela S. (2006). Characterization of the reactivity determinants of a novel hairpin substrate of yeast RNase III. *J. Mol. Biol.*, **363**, 332-344.
- Ghazal G., Ge D., Gervais-Bird J., Gagnon J. and Abou Elela S. (2005). Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. *Mol. Cell. Biol.*, **25**, 2981-2994.
- Giorgi C., Fatica A., Nagel R. and Bozzoni I. (2001). Release of U18 snoRNA from its host intron requires interaction of Nop1p with the Rnt1p endonuclease. *EMBO J.*, **20**, 6856-6865.
- Gonzales B., Henning D., So R.B., Dixon J., Dixon M.J. and Valdez B.C. (2005). The Treacher Collins syndrome (*TCOF1*) gene product is involved in pre-rRNA methylation. *Hum. Mol. Genet.*, **14**, 2035-2043.
- Gonzales F.A., Zanchin N.I.T., Luz J.S. and Oliveira C.C. (2005). Characterization of *Saccharomyces cerevisiae* Nop17p, a novel Nop58p-interacting protein that is involved in pre-rRNA processing. *J. Mol. Biol.*, **346**, 437-455.
- Guffanti E., Ferrari R., Preti M., Forloni M., Harismendy O., Lefebvre O. and Dieci G. (2006). A minimal promoter for TFIIC-dependent *in vitro* transcription of snoRNA and tRNA genes by RNA polymerase III. *J. Biol. Chem.*, **281**, 23945-23957.
- Hall D.B., Wade J.T. and Struhl K. (2006). An HMG protein, Hmo1, associates with promoters of many ribosomal protein genes and throughout the rRNA gene locus in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **26**, 3672-3679.
- Harbison C.T., Gordon D.B., Lee T.I., Rinaldi N.J., Macisaac K.D., Danford T.W., Hannett N.M., Tagne J.B., Reynolds D.B., Yoo J., Jennings E.G., Zeitlinger J. Pokholok D.K., Kellis M., Rolfe P.A., Takusagawa K.T., Lander E.S., Gifford D.K., Fraenkel E. and Young R.A. (2004). Transcriptional regulatory code of a eukaryotic genome. *Nature*, **431**, 99-104.
- Harismendy O., Gendrel C.G., Soularue P., Gidrol X., Sentenac A., Werner M. and Lefebvre O. (2003). Genome-wide location of yeast RNA polymerase III transcription machinery. *EMBO J.*, **22**, 4738-4747.

5. Bibliografia

- Hamma T., Reichow S.L., Varani G and Ferré-D'Amaré A.R. (2005). The Cbf5–Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat Struct Mol Biol*, **12**, 1101-1107.
- He H., Cai L., Skogerbø G., Deng W., Liu T., Zhu X., Wang Y., Jia D., Zhang Z., Tao Y., Zeng H., Aftab M.N., Cui Y., Liu G. and Chen R. (2006). Profiling *Caenorhabditis elegans* non-coding RNA expression with a combined microarray. *Nucleic Acids Res.*, **34**, 2976-2983.
- Henras A., Capeyrou R., Henry Y. and Caizergues-Ferrer M. (2004). Cbf5p, the putative pseudouridine synthase of H/ACA-type snoRNPs, can form a complex with Gar1p and Nop10p in absence of Nhp2p and box H/ACA snoRNAs. *RNA*, **10**, 1704-1712.
- Henras A., Dez C. and Henry Y. (2004). RNA structure and function in C/D and H/ACA s(no)RNPs. *Curr. Opin. Cell. Biol.*, **14**, 335-343.
- Hiley S.L., Babak T. and Hughes T.R. (2005). Global analysis of yeast RNA processing identifies new targets of RNase III and uncovers a link between tRNA 5' end processing and tRNA splicing. *Nucleic Acids Res.*, **33**, 3048-3056.
- Hirose T., Ideue T., Nagai M., Hagiwara M., Shu M.D. and Steitz J.A. (2006). A spliceosomal intron binding protein, IBP160, links position-dependent assembly of intron-encoded box C/D snoRNP to pre-mRNA splicing. *Mol. Cell*, **23**, 673-684.
- Hoang T., Peng W.T., Vanrobays E., Krogan N., Hiley S., Beyer A.L., Osheim Y.N., Greenblatt J., Hughes T.R. and Lafontaine D.L.J. (2005). Esf2p, a U3-associated factor required for small-subunit processome assembly and compaction. *Mol. Cell. Biol.*, **25**, 5523-5534.
- Huang Z.P., Zhou H., He H.L., Chen C.L., Liang D. and Qu L.H. (2005). Genome-wide analyses of two families of snoRNA genes from *Drosophila melanogaster*, demonstrating the extensive utilization of introns for coding of snoRNAs. *RNA*, **11**, 1303-1316.
- Hughes, J.D., Estep, P.W., Tavazoie, S., and Church, G.M. (2000). Computational identification of cis-regulatory elements associated with groups of functionally related genes in *Saccharomyces cerevisiae*. *J. Mol. Biol.*, **296**, 1205-1214.

5. Bibliografia

Hughes, J.M., Konings D.A. and Cesareni G. (1987). The yeast homologue of U3 snRNA. *EMBO J.*, **6**, 2145-2155.

Huttenhofer A., Kiefmann M., Meier-Ewert S., O'Brien J, Lehrach H., Bachellerie J.P. and Brosius J. (2001). RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *EMBO J.*, **20**, 2943-2953.

Inada M. and Guthrie C. (2004). Identification of Lhp1p-associated RNAs by microarray analysis in *Saccharomyces cerevisiae* reveals association with coding and noncoding RNAs. *Proc. Natl. Acad. Sci. USA*, **101**, 434-439.

Isogai Y., Takada S., Tjian R. and Keleş S. (2007). Novel TRF1/BRF target genes revealed by genome-wide analysis of *Drosophila* Pol III transcription. *EMBO J.*, **26**, 76-89.

Jady B.E. and Kiss T.(2001). A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO J.*, **20**, 541-551.

Jones K.W., Gorzynski K., Hales C.M., Fischer U., Badbanchi F., Terns R.M. and Terns M.P. (2001). Direct interaction of the spinal muscular atrophy disease protein SMN with the small nucleolar RNA-associated protein fibrillarin. *J. Biol. Chem.*, **276**, 38645-38651.

Kadaba S., Wang X. and Anderson J.T. (2006). Nuclear RNA surveillance in *Saccharomyces cerevisiae*: Trf4p-dependent polyadenylation of nascent hypomethylated tRNA and an aberrant form of 5S rRNA. *RNA*, **12**, 508-521.

Kickhoefer V.A., Searles R.P., Kedersha N.L., Garber M.E., Johnson D.L. and Rome L.H. (1993). Vault ribonucleoprotein particles from rat and bullfrog contain a related small RNA that is transcribed by RNA polymerase III. *J. Biol. Chem.* **268**, 7868-7873.

Kim M., Vasiljeva L., Rando O.J., Zhelkoysky A., Moore C. and Buratowski S. (2006). Distinct pathways for snoRNA and mRNA termination. *Mol. Cell*, **24**, 723-734.

King T.H., Decatur W.A., Bertrand E., Maxwell E.S. and Fournier M.J. (2001). A well-connected and conserved nucleoplasmic helicase is required for production of box C/D and H/ACA snoRNAs and localization of snoRNP proteins. *Mol. Cell. Biol.*, **21**, 7731-7746.

5. Bibliografia

- Kishore S. and Stamm S. (2006). The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science*, **311**, 230-232.
- Kiss-Laszlo Z., Henry Y., and Kiss T. (1998). Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *EMBO J*, **17**, 797-807.
- Kiss T., Marshallsay C. and Filipowicz W. (1991). Alteration of the RNA polymerase specificity of U3 snRNA genes during evolution and *in vitro*. *Cell*, **65**, 517-526.
- Kiss T. (2002). Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell*, **109**, 145-148.
- Koering C.E., Fourel G., Binet-Brasselet E., Laroche T., Klein F. and Gilson E. (2000). Identification of high affinity Tbf1p-binding sites within the budding yeast genome. *Nucleic Acid Res.*, **28**, 2519-2526.
- Kos M. and Tollervey D. (2005). The putative RNA helicase Dbp4p is required for release of the U14 snoRNA from preribosomes in *Saccharomyces cerevisiae*. *Mol. Cell*, **20**, 53-64.
- Kruszka K., Barneche F., Guyot R., Aihias J., Meneau I., Schiffer S., Marchfelder A. and Echeverria M. (2003). Plant dicistronic tRNA-snoRNA genes: a new mode of expression of the small nucleolar RNAs processed by RNase Z. *EMBO J*, **22**, 621-632.
- Kufel J., Allmang C., Chanfreau G., Petfalsky E., Lafontaine D.L.J. and Tollervey D. (2000). Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. *Mol. Cell. Biol.*, **20**, 5415-5424.
- LaCava J., Houseley J., Saveanu C., Petfalski E., Thompson E., Jacquier A. and Tollervey D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell*, **121**, 713-724.
- Lafontaine D.L.J., Preiss T. and Tollervey D. (1998). Yeast 18S rRNA dimethylase Dim1p: a quality control mechanism in ribosome synthesis? *Mol. Cell. Biol.*, **18**, 2360-2370.

5. Bibliografia

- Lafontaine D.L.J. and Tollervey D. (1996). One-step PCR mediated strategy for the construction of conditionally expressed and epitope tagged yeast proteins. *Nucleic Acids Res.*, **24**, 3469-3472.
- Lafontaine D.L.J. and Tollervey D. (1999). Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. *RNA* **5**, 455-467.
- Lafontaine D.L.J. and Tollervey D. (2000). Synthesis and assembly of the box C+D small nucleolar RNPs. *Mol. Cell. Biol.*, **20**, 2650-2659.
- Lascaris R.F., Groot E., Hoen P.B., Mager W.H. and Planta R.J. (2000). Different roles for abf1p and a T-rich promoter element in nucleosome organization of the yeast RPS28A gene. *Nucleic Acids Res.*, **28**, 1390-1396..
- Lapeyre B. and Purushothaman S.K. (2004). Spb1p-directed formation of Gm₂₉₂₂ in the ribosome catalytic center occurs at a late processing stage. *Mol. Cell*, **16**, 663-669.
- Lee C.Y., Lee A. and Chanfreau G. (2003). The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5'-end processing of *S. cerevisiae* box C/D snoRNAs. *RNA*, **9**, 1362-1370.
- Leung A.K. and Lamond A.I. (2002) In vivo analysis of NHPX reveals a novel nucleolar localization pathway involving a transient accumulation in splicing speckles. *J. Cell. Biol.*, **157**, 615-629.
- Li H.D., Zagorski J. and Fournier M.J. (1990). Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, **10**, 1145-1152.
- Liang D., Zhou H., Zhang P., Chen Y.Q., Chen X., Chen C.L. and Qu L.H. (2002). A novel organization: intronic snoRNA gene clusters from *Oryza sativa*. *Nucleic Acids Res.*, **30**, 3262-3272.
- Liang X.H. and Fournier M.J. (2006). The helicase Has1p is required for snoRNA release from pre-rRNA. *Mol. Cell. Biol.*, **26**, 7437-7450.

5. Bibliografia

- Liang X.H., Liu Q. and Michaeli S. (2003). Small nucleolar RNA interference induced by antisense or double-stranded RNA in trypanosomatids. *Proc. Natl. Acad. Sci. USA.*, **100**, 7521-7526.
- Liang X.H., Ochaion A., Xu Y.X., Liu Q. and Michaeli S. (2004). Small nucleolar RNA clusters in trypanosomatid *Leptomonas collosoma*. Genome organization, expression studies, and the potential role of sequences present upstream from the first repeated cluster. *J. Biol. Chem.*, **279**, 5100-5109.
- Liko D., Slattery M.G. and Heideman W. (2007). Stb3 binds to ribosomal RNA processing element motifs that control transcriptional responses to growth in *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **282**, 26623-2662.
- Lowe T.M. and Eddy S. (1999). A computational screen for methylation guide snoRNAs in yeast. *Science*, **283**, 1168-1171.
- Luo Y. and Li S. (2007). Genome-wide analyses of retrogenes derived from the human box H/ACA snoRNAs. *Nucleic Acids Res.*, **35**, 559-571.
- Marakova J.A. and Kramerov D.A. (2005). Noncoding RNA of U87 host gene is associated with ribosomes and is relatively resistant to nonsense-mediated decay. *Gene*, **363**, 51-60.
- Matera A.G., Terns R.M. and Terns M.P. (2007). Non-coding RNAs. Lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.*, **8**, 209-220.
- Mattick J.S. (2004). RNA regulation: a new genetics? *Nat. Rev. Genet.*, **5**, 316-323.
- Mattick J.S. and Makunin I.V. (2006). Non-coding RNA. *Hum. Mol. Genet.*, **15**, R17-29.
- Ma X., Yang C, Alexandrov A., Grayhack E.J., Behm-Ansmant I. and Yu Y.T. (2005). Pseudouridylation of yeast U2 snRNA is catalyzed by either an RNA-guided or RNA-independent mechanism. *EMBO J.*, **24**, 2403-2413.

5. Bibliografia

- Moqtaderi Z. and Struhl K. (2003). Genome-wide occupancy profile of the RNA polymerase III machinery in *Saccharomyces cerevisiae* reveals loci with incomplete transcription complexes. *Mol. Cell. Biol.*, **24**, 4118-4127.
- Morey C. and Avner P. (2004). Employment opportunities for non-coding RNAs. *FEBS Lett.*, **567**, 27-34.
- Morlando M., Ballarino M., Greco P., Caffarelli E., Dichtl B. and Bozzoni I (2004). Coupling between snoRNP assembly and 3' processing controls box C/D snoRNA biosynthesis in yeast. *EMBO J.*, **23**, 2392-2401.
- Morlando M., Greco P., Dichtl B., Fatica A., Keller V. and Bozzoni I. (2002). Functional analysis of yeast snoRNA and snRNA 3'-end formation mediated by uncoupling of cleavage and polyadenylation. *Mol. Cell Biol.*, **22**, 1379-1389.
- Morrissey J.P. and Tollervey D. (1995). Birth of the snoRNPs: the evolution of RNase MRP and the eukaryotic pre-rRNA-processing system. *Trends Biochem Sci.*, **20**, 78-82.
- Mouaikel J., Verheggen C. and Bertrand E. (2002). Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol Cell*, **9**, 891-901.
- Narayanan A., Lukowiak A., Jady B.E., Dragon F, Kiss T., Terns R.M. and Terns M.P. (1999). Nucleolar localization signals of box H/ACA small nucleolar RNAs. *EMBO J.*, **18**, 5120-5130
- Ochs R.L., Lischwe M.A. and Spohn W.H.(1985). Fibrillarin: A new protein of the nucleolus identified by autoimmune sera. *Biol Cell*, **54**, 123-133.
- Ofengand J. and Bakin A. (1997). Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaebacteria, mitochondria and chloroplasts. *J. Mol. Biol.*, **266**, 246-268.
- Oliver S.G. and McLaughlin C.S. (1977). The regulation of RNA synthesis in yeast. I: Starvation experiments. *Mol. Gen. Genet.*, **154**, 145-153.

5. Bibliografia

- Omer A.D., Lowe T.M., Russell A.G., Ebhardt H., Eddy S.R. and Dennis P.P. (2000). Homologs of small nucleolar RNAs in Archaea. *Science*, **288**, 517-522.
- Park S.H., Koh S.S., Chun J.H., Hwang H.J. and Kang H.S. (1999). Nrg1 is a transcriptional repressor for glucose repression of *STAI* gene expression in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **19**, 2044-2050.
- Pelczar P. and Filipowicz W. (1998). The host gene for intronic U17 small nucleolar RNAs in mammals has no protein-coding potential and is a member of the 5'-terminal oligopyrimidine gene family. *Mol. Cell. Biol.*, **18**, 4509-4518.
- Pellizzoni L., Baccon J., Charroux B. and Dreyfuss G. (2001). The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1. *Curr. Biol.*, **11**, 1079-1088.
- Pellizzoni L., Kataoka N., Charroux B. and Dreyfuss G. (1998). A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell*, **95**, 615-624.
- Peng W.T., Krogan N.J., Richards D.P., Greenblatt J.F. and Hughes T.R. (2004). ESF1 is required for 18S rRNA synthesis in *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **32**, 1993-1999.
- Peng W.T., Robinson M.D., Mnaimneh S., Krogan N.J., Cagney G., Morris Q., Davierwala A.P., Grigull J., Yang X., Zhang W., Mitsakakis N., Ryan O.W., Datta N., Jojic V., Pal C., Canadien V., Richards D., Beattie B., Wu L.F., Altschuler S.J., Roweis S., Frey B.J., Emili A., Greenblatt J.F. and Hughes T.R. (2003). A panoramic view of yeast noncoding RNA processing. *Cell*, **113**, 919-933.
- Petfalski E., Dandekar T., Henry Y. and Tollervey D. (1998). Processing of precursors to small nucleolar RNAs and ribosomal RNAs requires common components. *Mol. Cell. Biol.*, **18**, 1181-1189.
- Piekna-Przybylska D., Decatur W.A. and Fournier M.J. (2007). New bioinformatic tools for analysis of nucleotide modifications in eukaryotic rRNA. *RNA*, **13**, 305-312.

5. Bibliografia

- Preti M., Guffanti E., Valitutto E. and Dieci G. (2006). Assembly into snoRNP controls 5'-end maturation of a box C/D snoRNA in *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.*, **351**, 468-473.
- Qu L.H., Henras A., Lu Y.J., Zhou H., Zhou W.X., Zhu Y.Q., Zhao J., Henry Y., Caizergues-Ferrer M. and Bachellerie J.P. (1999). Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. *Mol. Cell. Biol.*, **19**, 1144-1158.
- Richard P., Kiss M.A., Darzacq X. and Kiss T. (2006). Cotranscriptional recognition of human intronico box H/ACA snoRNAs occurs in a splicing-independent manner. *Mol. Cell. Biol.*, **26**, 2540-2549.
- Rozhdestvensky T., Tang T., Tchirkova I., Brosius J., Bachellerie J.P. and Huttenhofer A. (2003). Binding of L7Ae protein to the K-turn of archaeal snoRNAs: A shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea. *Nucleic Acids Res.*, **31**, 869-877.
- Runte M., Varon R., Horn D., Horsthemke B. and Buiting K. (2005). Exclusion of the C/D box snoRNA gene cluster HBII-52 from a major role in Prader-Willi syndrome. *Hum. Genet.*, **116**, 228-230.
- Russell A.G., Schnare M.N. and Gray M.W. (2004). Pseudouridine-guide RNAs and other Cbf5p-associated RNAs in *Euglena gracilis*. *RNA*, **10**, 1034-1046.
- Samarsky, D.A., and Fournier M.J. (1998). Functional mapping of the U3 small nucleolar RNA from the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **18**, 3431-3444.
- Schattner P., Decatur W.A., Davis C.A., Ares M.Jr., Fournier M.J. and Lowe T.M. (2004). Genome-wide searching for pseudouridylation guide snoRNAs: analysis of the *Saccharomyces cerevisiae* genome. *Nucleic Acids Res.*, **32**, 4281-4296.
- Schmitt M.E. and Clayton D.A. (1993). Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **13**, 7935-7941.

5. Bibliografia

- Shimba S., Buckley B., Reddy R., Kiss T. and Filipowicz W. (1992). Cap structure of U3 small nucleolar RNA in animal and plant cells is different. *J. Biol. Chem.*, **19**, 13772-13777.
- Singh S.K., Gurha P., Tran E.J., Maxwell E.S. and Gupta R. (2004) sequential 2'-O-methylation of archaeal pre-tRNA^{Trp} nucleotides is guided by the intron-encoded but *trans*-acting box C/D ribonucleoprotein of pre-tRNA. *J. Biol. Chem.*, **279**, 47661-47671.
- Speckmann W.A., Terns R.M. and Terns M.P. (2000). The box C/D motif directs snoRNA 5'-cap hypermethylation. *Nucleic Acid Res.*, **28**, 4467-4473.
- Stanley S.M., Bailey T.L. and Mattick J.S. (2006). GONOME: measuring correlations between GO terms and genomic position. *BMC Bioinformatics*, **7**.
- Starostina N.G., Marshburn S., Johnson L.S., Eddy S.R., Terns R.M. and Terns M.P. (2004). Circular box C/D RNAs in *Pyrococcus furiosus*. *Proc Natl Acad Sci USA.*, **101**, 14097-14101.
- Steinmetz E.J., Conrad N.K., Brow D.A., Corden J.L. (2001). RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature*, **413**, 327-331.
- Storz G. (2002). An expanding universe of noncoding RNAs. *Science*, **296**, 1260-1263.
- Struhl K. (1985). Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc Natl Acad Sci USA.*, **82**, 8419-8423.
- Suter B., Schnappauf G. and Thoma F. (2000). Poly(dA-dT) sequences exist as rigid DNA structures in nucleosome-free yeast promoters *in vivo*. *Nucleic Acids Res.*, **28**, 4083-4089.
- Tanaka-Fujita R., Soeno Y., Satoh H., Nakamura Y. and Mori S. (2007). Human and mouse protein-noncoding snoRNA host genes with dissimilar nucleotide sequences show chromosomal synteny. *RNA*, **13**, 811-816.
- Tollervey D., Lehtonen H., Carmo-Fonseca M. and Hurt E.C. (1991). The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre-rRNA processing in yeast. *EMBO J.*, **10**, 573-583.

5. Bibliografia

- Tollervey D., Lehtonen H. Jansen R., Kern H. and Hurt E.C. (1993). Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell*, **72**, 443-457.
- Torchet C., Badis G., Devaux F., Costanzo G., Werner M. and Jacquier A. (2005). The complete set of H/ACA snoRNAs that guide rRNA pseudouridylations in *Saccharomyces cerevisiae*. *RNA*, **11**, 928-938.
- Tremblay A., Lamontagne B., Catala M., Yam Y., Larose S., Good L. and Abou Elela S. (2002). A physical interaction between Gar1p and Rnt1p is required for the nuclear import of H/ACA small nucleolar RNA-associated proteins. *Mol. Cell. Biol.*, **22**, 4792-4802.
- Tycowski K.T., Shu M.D. and Steitz J.A. (1993). A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. *Genes & Dev.*, **7**, 1176-1190.
- Tycowski K.T., Shu M.D. and Steitz J.A. (1994). Requirement for intron-encoded U22 small nucleolar RNA in 18S ribosomal RNA maturation. *Science*, **266**, 1558-1561.
- Tycowski K.T., You Z.H., Graham P.J and Steitz J.A. (1998). Modification of U6 spliceosomal RNA is guided by other small RNAs. *Mol. Cell*, **2**, 629-638.
- van Hoof A., Lennertz P. and Parker R. (2000). Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol. Cell. Biol.*, **20**, 441-452.
- Vasiljeva L. and Buratowski S. (2006). Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. *Mol. Cell*, **21**, 239-248.
- Venema J. and Tollervey D. (1999). Ribosome synthesis in *Saccharomyces cerevisiae*. *Annu. Rev. Genet.*, **33**, 261-311.
- Verheggen C., Lafontaine D.L., Samarsky D., Mouaikel J., Blanchard J.M., Bordonné R. and Bertrand E. (2002). Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. *EMBO J.*, **21**, 2736-2745.

5. Bibliografia

- Villa T., Ceradini F. and Bozzoni I. (2000). Identification of a novel element required for processing of intron-encoded box C/D small nucleolar RNAs. *Mol. Cell. Biol.*, **20**, 1311-1320.
- Vincenti S., De Chiara V., Bozzoni I. and Presutti C. (2007). The position of yeast snoRNA-coding regions within host introns is essential for their biosynthesis and for efficient splicing of the host pre-mRNA. *RNA*, **13**, 138-150.
- Vitali P., Royo H., Seitz H., Bachellerie J.P., Hüttenhofer A. and Cavallé J. (2003). Identification of 13 novel human modification guide RNAs. *Nucleic Acids Res.*, **31**, 6543-6551.
- Wade C.H., Umbarger M.A. and McAlear M.A. (2006). The budding yeast rRNA and ribosome biosynthesis (RRB) regulon contains over 200 genes. *Yeast*, **23**, 293-306.
- Waibel F. and Filipowicz W. (1990). RNA-polymerase specificity of transcription of *Arabidopsis* U snRNA genes determined by promoter element spacing. *Nature*, **346**, 199-202.
- Wang K.L.C. and Warner J.R. (1998). Positive and negative autoregulation of *REB1* transcription in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **18**, 4368-4376.
- Wang Y., Pierce M., Schnepfer L., Güldal C.G., Zhang X., Tavazoie S. and Broach J. (2004). Ras and Gpa2 mediate one branch of a redundant glucose signalling pathway in yeast. *PLoS Biol.*, **2**, 610-622.
- Warner J.R. (1999). The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.*, **24**, 437-440.
- Watkins N., Dickmanns A. and Lührmann R. (2002). Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Mol. Cell. Biol.*, **22**, 8342-8352.
- Watkins N.J., Segault V., Charpentier B., Nottrott S., Fabrizio P., Bachi A., Wilm M., Rosbash M., Branlant C. and Lührmann R. (2000). A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP. *Cell*, **103**, 457-466.

5. Bibliografia

- Weber M.J. (2006). Mammalian small nucleolar RNAs are mobile genetic elements. *PLoS Genet.*, **2**, e205.
- Wolin S.L. and Cedervall T. (2002). The Ia protein. *Annu. Rev. Biochem.*, **71**, 375-403.
- Wyers F., Rougemaille M., Badis G., Rousselle J.C., Dufour M.E., Boulay J., Régnault B., Devaux F., Namane A., Séraphin B., Libri D. and Jacquier A. (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell*, **121**, 725-737.
- Xiao S., Scott F., Fierke C.A. and Engelke D.R. (2002). Eukaryotic ribonuclease P: a plurality of ribonucleoprotein enzymes. *Annu. Rev. Biochem.*, **71**, 165-189.
- Xiao Y., Zhou H. and Qu L.H. (2006). Characterization of three novel imprinted snoRNAs from mouse *Irm* gene. *Biochem. Biophys. Res. Commun.*, **340**, 1217-1223.
- Yang P.K., Hoareau C., Froment C. Monsarrat B., Henry Y. and Chanfreau G. (2005). Cotranscriptional recruitment of the pseudouridyltransferase Cbf5p and of the RNA binding protein Naf1p during H/ACA snoRNP assembly. *Mol. Cell. Biol.*, **25**, 3295-3304.
- Yang P.K., Rotondo G., Porras T., Legrain P. and Chanfreau G. (2002). The Shq1p:Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis. *J. Biol. Chem.*, **277**, 45235-45242.
- Zebarjadian Y., King T., Fournier M.J., Clarke L. and Carbon J. (1999). Point mutations in yeast CBF5 can abolish in vivo pseudouridylation of rRNA. *Mol. Cell Biol.*, **19**, 7461-7472.
- Zemann A., op de Bekke A., Kiefmann M., Brosius J. and Schmitz J. Evolution of small nucleolar RNAs in nematodes. *Nucleic Acids Res.*, **34**, 2676-2685.
- Zhang Z. and Dietrich F.S. (2005). Mapping of transcription start sites in *Saccharomyces cerevisiae* using 5' SAGE. *Nucleic Acids Res.*, **33**, 2838-2851.

5. Bibliografia

Zhao J., Hyman L. and Moore C. (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.*, **63**, 405-445.

Zhao Y. and Srivastava D. (2007). A developmental view of microRNA function. *Trends Biochem. Sci.*, **32**, 189-197.

Zhou H., Chen Y.Q. and Du Y.P. and Qu L.H. (2002). The *Schizosaccharomyces pombe* mgU6-47 gene is required for 2'-O-methylation of U6 snRNA at A41. *Nucleic Acids Res.*, **30**, 894-902.

6. Appendice

In questo capitolo sono riportate le tabelle e gli allineamenti a cui si fa riferimento nel testo.

Tabella 1: conservazione dei geni per snoRNA in differenti specie di Funghi

snoRNA	<i>S.cerevisiae</i>	<i>S.paradoxus</i>	<i>S.kudriavzevii</i>	<i>S.bayanus</i>	<i>S.castellii</i>	<i>S.mikatae</i>	<i>K.lactis</i>	<i>C.glabrata</i>	<i>S.kluyveri</i>	<i>K.waltii</i>	<i>E.gossypii</i>
U3a	3.1e-68	3.8e-62	1.5e-57	2.0e-55	2.4e-28	9.6e-58	3.4e-36	5.1e-21	7.3e-38	5.2e-27	1.2e-27
U3b	2.6e-66	6.7e-58	3.6e-55	1.2e-54	4.5e-27	5.3e-57	2.1e-34	6.1e-25	2.3e-38	1.1e-25	4.6e-27
U14	2.3e-22	2.2e-20	1.4e-19	1.1e-20	1.9e-15	3.2e-22	1.1e-14	1.2e-14	4.5e-15	9.7e-15	3.9e-15
snR3	1.1e-37	1.0e-35	8.3e-34	9.8e-31	0.014	1.1e-32	0.00041	0.0017			
snR4	3.4e-35	1.1e-32	2.8e-30		3.3e-13	8.2e-28					
snR5	1.2e-37	1.6e-33	3.6e-32	1.4e-30	6.2e-13	2.4e-34	0.0036		0.0002		
snR8	3.8e-36	3.3e-34	6.7e-33	4.5e-30	2.1e-12	3.0e-32	3.8e-07	1.0e-09	5.2e-05	3.1e-07	7.0e-06
snR9	6.1e-36	1.1e-29	2.1e-27	2.8e-27		1.6e-26				0.023	
snR10	2.4e-49	1.5e-44	3.2e-44	5.1e-43	4.4e-21	1.2e-42	7.5e-11	5.1e-19	2.7e-24	1.0e-22	1.8e-19
snR11	5.2e-52	5.4e-47	4.9e-44	6.7e-24	1.4e-10	3.6e-43		1.4e-11			
snR13	7.4e-22	2.4e-20	1.2e-17	6.2e-19	0.00018	2.0e-20	0.00075	1.1e-07			0.00024
snR30	3.2e-130	9.4e-117	1.7e-106	1.0e-93	5.5e-22	2.0e-105	8.0e-18	2.8e-27	3.6e-12	1.4e-09	6.2e-08
snR31	5.3e-45	3.8e-36	3.9e-30	2.3e-33	1.1e-10	5.1e-35	1.1e-09		0.00098		
snR32	3.6e-36	1.5e-29	1.1e-26	1.4e-26	1.5e-07	5.4e-25	2.3e-05	3.5e-08	0.00017	0.00016	0.017
snR33	1.5e-34	5.3e-29	1.9e-25	6.2e-26		6.2e-25					
snR34	7.4e-112	3.5e-63	1.7e-42	1.1e-40	3.3e-11	5.2e-57	6.1e-11	3.0e-11	2.6e-17	7.1e-14	9.6e-05
snR35	1.8e-110	3.3e-26	9.8e-60	2.8e-35	1.3e-13	1.2e-45	2.0e-17	2.7e-07	0.045	8.1e-09	9.6e-09
snR36	7.4e-68	3.0e-37	8.3e-24	5.3e-23	4.4e-11	1.3e-27		0.077	1.0e-05		
snR37	1.9e-81	2.4e-72	3.4e-68	4.3e-63	1.1e-12	4.0e-63	2.4e-17	1.1e-17	2.7e-24	1.8e-18	0.00055
snR39b	1.4e-14	3.4e-14	3.2e-11	3.2e-11	0.019	3.2e-12		0.023	0.093		
snR40	8.8e-15	7.6e-14		3.2e-12	1.5e-08	3.1e-13	3.3e-05	0.00018	1.6e-05	7.6e-05	8.5e-05
snR41	2.1e-14	4.1e-13	3.6e-12	2.2e-12	7.3e-09	4.2e-10	0.00018	0.0021	0.0012	0.0082	0.0024
snR42	2.0e-63	4.1e-48	5.3e-39	1.2e-40	9.9e-09	1.1e-37	2.8e-07	1.2e-06	8.4e-06	7.0e-06	1.9e-08
snR43	3.0e-41	3.5e-35	1.2e-32	6.0e-29	2.2e-05	3.5e-35	2.2e-05		9.8e-06	0.00088	
snR45	9.1e-32	9.0e-31	3.7e-29	1.7e-29	5.0e-19	1.3e-29	4.7e-13	1.6e-12	4.8e-12	1.7e-17	1.2e-16
snR46	7.5e-38	6.2e-29	4.0e-24	4.7e-28	0.00044	4.9e-26	0.002	1.5e-05	0.0026	2.9e-16	
snR47	3.9e-15	1.3e-10	8.4e-09	1.6e-08		3.2e-09					

snoRNA	<i>S.cerevisiae</i>	<i>S.paradoxus</i>	<i>S.kudriavzevii</i>	<i>S.bayanus</i>	<i>S.castelii</i>	<i>S.mikatae</i>	<i>K.lactis</i>	<i>C.glabrata</i>	<i>S.kluyveri</i>	<i>K.waltii</i>	<i>E.gossypii</i>
snR48	1.7e-18	8.7e-16	5.0e-14	2.1e-15	1.6e-07	1.8e-15	0.0014	0.0032	5.8e-07	1.5e-07	9.1e-10
snR49	2.8e-26	1.9e-23	2.5e-19	1.2e-20	1.8e-05	7.4e-20			0.045		0.0067
snR50	2.9e-13	1.2e-12	2.2e-10	4.0e-10		3.4e-09		0.045			
snR51	4.1e-17	3.1e-15	1.8e-14	1.8e-13		2.8e-14	0.0015				0.00084
snR52	9.4e-14	4.1e-13	2.5e-12	3.8e-13	5.7e-07	6.8e-13	0.0038		0.00081		0.0047
snR53	2.0e-13	1.7e-12	2.0e-10	9.5e-11		4.5e-12					
snR55	1.6e-12	1.6e-11	9.9e-11	1.4e-11	5.2e-08	1.6e-10	0.00027	0.00019	0.038	0.0003	0.0065
snR56	3.5e-13	3.3e-12	3.2e-12	2.0e-12		2.7e-12	0.0044		1.8e-06	0.0016	
snR57	9.8e-13	6.1e-12	5.3e-11	4.5e-07	7.1e-08	3.0e-09	1.5e-05	0.0002	2.4e-05	0.071	0.00011
snR58	1.2e-14	2.2e-13	5.1e-13	4.9e-13		3.5e-12					
snR60	1.9e-16	3.1e-16	3.0e-16	5.0e-16		5.8e-16	4.2e-06	1.1e-05	1.5e-05	0.0039	0.019
snR61	3.4e-13	5.0e-12	1.4e-11	1.4e-11	8.4e-05	1.3e-09	0.0021	9.8e-08	0.038	0.00016	0.00072
snR62	1.7e-15	6.4e-15	7.5e-13	5.1e-12	7.0e-06	9.0e-14					
snR63	2.9e-51	7.5e-46	3.2e-42	2.6e-40		1.6e-41	0.00023	0.0029	0.00013		
snR64	9.9e-16	3.8e-15	9.6e-13	2.2e-12	2.1e-08	1.3e-14	9.7e-08		2.6e-07	2.5e-08	1.1e-09
snR65	1.7e-15	5.0e-13	4.0e-12	1.2e-10	0.00026	2.6e-11			4.1e-05		
snR66	1.0e-12	4.4e-11	1.1e-10	4.3e-11	0.00092	3.7e-11			0.00014		
snR67	2.4e-11	2.9e-11	1.6e-09	9.4e-10		2.2e-11			0.035	0.078	
snR68	3.7e-24	1.5e-20	9.2e-20	2.7e-22	0.038	6.3e-20	0.0043				
snR69	9.9e-16	1.5e-15	7.9e-15	1.0e-14	1.2e-08	6.5e-15	7.8e-06		2.8e-05	1.3e-05	
snR70	3.6e-30	4.4e-26	2.0e-24	6.6e-21	1.1e-12	2.3e-23		4.5e-12			
snR71	3.4e-13	1.2e-12	1.2e-10	5.0e-08	0.0003	1.7e-11	8.5e-05	2.5e-05	2.9e-06	9.4e-06	0.0001
snR72	4.9e-15	1.7e-14	3.3e-13	1.6e-11		1.6e-12					
snR73	7.0e-17	2.7e-16	7.1e-16	2.1e-17	1.7e-07	2.2e-16	9.9e-08	3.6e-10	8.3e-07		1.8e-07
snR74	9.8e-13	3.5e-12	2.6e-11	8.5e-12	2.7e-06	6.8e-11					
snR75	5.8e-13	4.8e-12	4.9e-12	1.3e-10	8.5e-05	4.0e-12	5.1e-06				
snR76	1.4e-17	1.1e-13	3.2e-11	8.5e-12	6.3e-06	1.4e-11		1.3e-05	0.00053	0.0019	0.031

snoRNA	<i>S.cerevisiae</i>	<i>S.paradoxus</i>	<i>S.kudriavzevii</i>	<i>S.bayanus</i>	<i>S.castellii</i>	<i>S.mikatae</i>	<i>K.lactis</i>	<i>C.glabrata</i>	<i>S.kluyveri</i>	<i>K.waltii</i>	<i>E.gossypii</i>
snR77	9.8e-13	5.3e-12	4.4e-11	1.1e-09	4.7e-08	1.7e-12	1.2e-05	1.8e-05	1.5e-05	5.8e-07	5.1e-06
snR78	2.8e-12	2.3e-11	2.4e-11	5.5e-11	0.0032	1.9e-11			0.041		
snR79	6.7e-12	9.5e-12	9.0e-12	2.5e-11	6.2e-05		0.019		0.0003	0.0006	0.011
snR189	2.2e-36	2.6e-32	8.1e-27	1.3e-26		1.2e-27	2.4e-05		2.3e-06	3.0e-06	0.064
snR190	5.8e-34	2.0e-14	5.7e-24	2.9e-22	8.2e-07	1.2e-25	0.00075	7.6e-05	0.0058	0.00019	
snR87	1.0e-50	2.0e-43	2.0e-37	1.0e-35		5.0e-39					
snR82	9.3e-54	1.6e-47	4.3e-38	8.4e-36	1.7e-06	1.0e-41	6.9e-08	1.4e-06	2.6e-08	1.1e-08	0.037
snR83	1.9e-62	1.4e-50	6.1e-43	7.3e-35	2.9e-06	6.0e-48					
snR84	1.1e-117	7.6e-75	4.2e-37	3.1e-55		1.3e-45					
snR80	1.4e-31	1.9e-27	1.3e-25	6.1e-24	2.5e-10	1.7e-27	3.0e-05			0.084	0.0048
snR85	1.0e-91	1.0e-36	6.0e-29	6.0e-20		4.0e-27		0.003			
snR86	1.2e-219	4.9e-142	3.3e-64	7.5e-88		3.1e-104					
snR161	3.5e-29	1.2e-25	2.3e-17	3.5e-15		1.1e-20	0.00067		0.0014	0.014	0.094
snR81	2.0e-109		2.0e-51			7.0e-54					
MRP RNA	5.4e-71	1.3e-61	1.2e-58	3.0e-53	8.8e-31	1.1e-57	7.6e-15	9.7e-22	1.4e-26	1.5e-28	4.7e-09

Tabella 2: elementi conservati presenti a monte dei geni per snoRNA

snoRNA H/ACA box	5' processing	TATA b. (respect to mature snoRNA cds)	TATA box^	Poly (A/T)*	Reb1p binding site*	ARCCCTaa*	RRPE*	Other
snR3	N	-90	-90	10 bp (-11)	AGGGTAA (-67)	TGCCCTAA (-90)		
snR5	N	-70	-70			AGCCCTAA (-56)	GAAAAATTTT (-28)	
snR8	N	-89	-89					
snR9	N	-99	-99			AGCCCTTA (-91)		
snR10	N	-99	-99	9 bp (-15)	TTACCCG (-41)	AGCCCTAC (-68)		
snR11	N	-103	-103	8 bp (-79)	TTACCCG (-55)	AACCCTAC (-89)		
snR30	N	-101	-101		TTACCCT (-56)	AACCCTAA (-90)	GAAAAATTTT (-28)	PAC
snR31	N	-66	-66				TGAAAAATTTT (-58)	TTAGGGC (-36)
snR32	N	-89	-89	11 bp (-26)	AGGGTAA (-38)	AGCCCTAC (-71)		
snR33	N			10 bp [-28]			TGAAAAATTT [-99]	
snR34	N	-140	-140					
snR35	N	-92	-92	9 bp (-23)	CGGGTAA (-49)	ACCCTAA (-80)		
snR36	Y			12 bp [-244]				
snR37	N	-91	-91		CGGGTCA (-56)	AACCCTAC (-90)		
snR42	N	-95	-95		TTACCCG (-54)	TGCCCTGA (-78)		
snR43	Y	-170		13 bp (-30)	TCACCCG (-57)	AGCCCTAA (-89)		
snR46	Y					AACCCTAA [-382]		
snR49	N	-112	-112	8 bp (-52)		AGCCCTAA (-62)		
snR80		-199						CTAGGGTT (-89)
snR81		-146		7 bp (-42)		AACCCTAG (-112)		
snR82					TTACCCG [-123]	agccctaa [-150]		
snR83	Y	-193			CGGGTAA (-59)	AGCCCTCA (-91)		
snR84	Y				CGGGTAT [-205]		TGAAAAATTTT [-181]	PAC
snR85		-148		7 bp (-40)				Rap1p b.s.
snR86					TCACCCG [-121]	AGCCCTAT [-148]		
snR161		-79		7 bp (-17)				
snR189	N	-73	-73	9 bp (-23)	CGGGTAA (-46)			
MRP RNA	N	-104	-104	14 bp (-18)		AACCCTAA (-72)		

snoRNA C/D box	5' processing	TATA b. (respect to mature snoRNA cds)	TATA box^	Poly (A/T)*	Reb1p binding site*	ARCCCTaa*	RRPE*	Other
snR4	N	-104	-104	9 bp (-29)				
snR13	N	-83	-83	10 bp (-24)		AGCCCTAA (-77)		
snR17a	N	-86	-86	8 bp (-80)		AACCCTGA (-88)	TGAAAATTTT (-35)	Rap1p b.s.
snR17b	N	-96	-96					Rap1p b.s.
snR39b	Y	-259		18 bp (-68)				
snR40	Y	-216			CGGGTGA (-54)	AGCCCTAA (-88)		
snR41	Y	-214			CGGGTGA (-25)	AGCCCTAA (-49)		
snR45	N	-80	-80	8 bp (-26)	TTACCCG (-65)	AGCCCTAA (-90)	GAAAATTTT (-118); GAAAATTT (-106)	PAC
snR47	Y	-178		10 bp (-52)		AACCCTAG (-100)		Abf1p b.s.
snR48	Y	-255		11 bp (-14)	CGGGTAA (-74)	AACCCTAA (-109)		
snR50	Y			13 bp [-272]		AGCCCTAC [-281]		
snR52	Y							A and B box
snR56	Y	-237				AACCCTAA (-85)		
snR57	Y	-258				AACCCTAA (-89)	GAAAATTTT (-18)	
snR58	Y	-164	-59 #				GAAAATTT (-18); TGAAAATTT (-45)	
snR60	Y	-221	-93 #	7 bp (-30)		AGCCCTAA (-70)		
snR62	Y	-216	-75 #	9 bp (-38)			TGAAAATTTT (+18)	Rap1p b.s.; PAC
snR63	Y				CGGGTGA [-337]	AGCCCTAA [-369]	TGAAAATTTT [-314]	
snR64	Y	-229	-65 ##	7 bp (-41)		AGCCCTAA (-81)		
snR65	Y	-176	-59 ##	9 bp (-15)		AACCCTAA (-68)		
snR66	Y							
snR67	Y	-318				AGCCCTAA (-56)	GAAAATTTT (-16)	
snR68	Y	-336	-167 #					
snR69	Y	tataaaa (-215)		12 bp [-284]				
snR71	Y	-277	-77 #	7 bp (-38)	CGGGTTT (-46)	AGCCCTAA (-77)		
snR78	Y	-247	-75 §		TTACCCG (-73)	AGCCCTCA (-97)	TGAAAATTTT (-34)	2 Rap1p b.s.
snR79	Y	-201		15 bp (-39)	TGGGTAA (-55)	AGCCCTAA (-80)		
snR87	Y			7 bp [-443]				
snR190	Y	-227	-77			TGCCCTAA (-79)	GAAAATTTT (-47)	PAC; Rap1p

Gli elementi indicati in tabelle sono conservati in almeno tre specie di Saccaromiceti. Le distanze riportate sono riferite a *S. cerevisiae*, così come le sequenze per il sito di legame per Reb1p e i motivi ARCCCTaa e RRPE e le lunghezze degli elementi poli(A/T). In marrone chiaro sono indicati gli snoRNA H/ACA box, in verde l'RNA della RNasi MRP e in turchese gli snoRNA C/D box.

I siti riportati all'interno di caselle gialle e scritti in rosso non presentano consensi perfetti; l'elemento RRPE indicato in rosso è presente solo in *S. cerevisiae* (negli altri Saccaromiceti è presente un poli(T)).

Se la sequenza riportata è conservata tra i Saccaromiceti ma è mutata in *S. cerevisiae*, è indicata in lettere minuscole (le distanze sono comunque riferite a *S. cerevisiae*).

Sono state indicate anche sequenze poli(A/T) molto distanti dalla TATA box (in blu).

^: la distanza riportata è riferita al sito d'inizio della trascrizione.

*: la distanza riportata è rispetto alla TATA box; se questa non è stata individuata, la distanza indicata è riferita alla prima base del trascritto maturo ed è riportata in parentesi quadre.

#: il sito d'inizio della trascrizione è stato ricavato da dati presenti in letteratura.

##: il sito d'inizio è stato individuato mediante primer extension nel corso di questo lavoro di tesi e confrontato con dati presenti in letteratura.

§: in Qu *et al.*, 1999 sono riportati quattro probabili siti di inizio della trascrizione; si fa qui riferimento a quello che dà origine al trascritto più lungo.

Tabella 3: classi funzionali maggiormente rappresentate tra le ORF adiacenti ai geni per snoRNA

FUNCTIONAL CATEGORY	n° SET	% SET	n° GENOME	% GENOME	Fold (%set/%gen)	P-VALUE	genes SET
12.01 ribosome biogenesis	19	19,3	310	5,05	3,82	3,09E-07	YBL072c YCR031c YDL075w YDR471w YEL050c YEL054c YER131w YGL076c YGL099w YHR141c YKL006w YLR106c YLR333c YLR388w YMR194w YOR234c YOR272w YPL143w YPR144c
12.01.01 ribosomal proteins	16	16,3	246	4,01	4,06	1,38E-06	YBL072c YCR031c YDL075w YDR471w YEL050c YEL054c YER131w YGL076c YHR141c YKL006w YLR106c YLR333c YLR388w YMR194w YOR234c YPL143w
12 PROTEIN SYNTHESIS	21	21,4	480	7,83	2,73	1,62E-05	YBL072c YCR031c YDL075w YDR471w YEL050c YEL054c YER131w YGL076c YGL099w YHR141c YKL006w YLR106c YLR333c YLR388w YMR194w YOL097c YOR234c YOR272w YOR276w YPL143w YPR144c
14.07.05 modification by ubiquitination, deubiquitination	5	5,1	79	1,28	3,98	0,008356889	YDL074c YGL087c YJL047c YKL210w YNR051c
43.01.03.09 development of asco-basidio- or zygosporo	7	7,14	167	2,72	2,63	0,017045522	YER132c YGL099w YGR059w YLR028c YMR013c YOL051w YOL052c
01.07.01 biosynthesis of vitamins, cofactors, and prosthetic groups	5	5,1	110	1,79	2,85	0,030888429	YBR035c YOL052c YOR184w YOR236w YOR278w
14.10 assembly of protein complexes	7	7,14	199	3,24	2,20	0,039421068	YBL074c YDL074c YGL087c YJL047c YKL210w YLR106c YML102w
01.01 amino acid metabolism	8	8,16	243	3,96	2,06	0,039858156	YKL211c YLR027c YNR050c YOL052c YOR040w YOR184w YOR236w YPR145w
42.10 nucleus	5	5,1	149	2,43	2,10	0,089525176	YDL074c YJL148w YML103c YOR185c YOR272w

La tabella rappresenta i risultati ottenuti mediante ricerca in “The MIPS Functional Catalogue Database” (http://mips.gsf.de/proj/funecatDB/search_main_frame.html), a partire dalla lista delle ORF adiacenti ai geni per snoRNA; è stata qui riportata solo la parte della tabella che comprende i risultati più significativi. Le categorie funzionali sono riportate in ordine decrescente di P-value (settima colonna); va sottolineato che “ribosomal proteins” è una sottocategoria di “ribosomal biogenesis” che, a sua volta, è una sottocategoria di “protein synthesis”. Nella sesta colonna (Fold) è riportato il rapporto tra la percentuale dei geni della categoria adiacenti ai geni per snoRNA (terza colonna) e la percentuale dei geni della stessa categoria presenti nell’intero genoma (quinta colonna).

Allineamento della regione attorno al sito d'inizio in *S. cerevisiae*

```

          *           20           *           40           *           60
snR42 : ACCTTTGATGCTTTAAAAAACATTTCTCTTAAAAGAAAGCTAACAATTTGTGATGCTTT : 60
snR189 : GTGACCGTTGTTTTTCAACTGTGAACTTATAAGAAATTACAAAAATAATTTACATTCC : 60
snR8 : CATTFTTGTAGTTATCTTTGATTTTCGGTTCGGAAGAAAGGCTTGATACATATTAACATGC : 60
snR35 : GTATATTTACTCTTTTATTTCAAATTTAAAAAAAACCCTACAAAATTAATCGTGCGG : 60
snR4 : GTCATTGATCTTTTCATTTTTTATTTCAAATCCCCATCTACTATTACAGTCGATGAGG : 60
snR30 : CCATTCAAGTTTTCTCTTTGCTTATTTGTAAAAAAAAGTAACCATAGTCTCGTGCTAGT : 60
snR45 : CGCGTTAGATTTTCTTTGGAGGTTTCATTTAAAGAAACCAATGACCTTCCAAGTTTTTAA : 60
snR32 : TTATAATCAAATTCCTTTTGGTTTTTGTTTTAAAAAGGCCACATCATAAATTATATCGCG : 60
snR37 : TAAAATTCAAATTTTCTTTTGTATTTTTTCAAAGGGCTACATCAGTAGTGGTTAATA : 60
snR10 : TTACTTTCCATTTCTTATTTTAGTGTATTTTATAACAGTACGCAAATTTAACAGCCAT : 60
snR33 : CTCTCTTTCACATTTTTTTTTTCATAAGAATTAAAAAAATAGCTTAATGCCCTCTTTGT : 60
snR13 : CGTCTCTATATCTTTTGTAAATTGGATTTTTTAAAGGGTCTGGAAAGTTTTTCTTTTTTA : 60
snR34 : TATATTTACCACATTTTTCTGATTTTTTTTTAAGAGGGTGAATCAAAAATTTATTTTTT : 60
snR11 : AATCTTTCTTTTCTTTTTTTTGGTTTTTTGAAAAAGGGTATAATGTCTTTTATAATAG : 60
snR3 : TTGTTAATTCTATTTTTGAAGTTTCCTTCCGAAAAAAGTAACTTTGTCTTAAAGTACTA : 60
snR49 : TCCTTGATTTTACTTTCAATTTAGTTTACAATATCATGGGACTTTATCTCCATGACTATG : 60
snR5 : AACTTACTTCATTTGCAGGATCCTTCAGGATAAGAAACCCTCATTCAATAAACTGATCT : 60
snR31 : GCCCATTTTTACTTCTATTTCCATTTGAATTCAGAAAAGTGAAGCAAAAATTACACCATGA : 60
snR9 : TTCTCTCTTCTTTCTTCTTCTTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT : 60
          tt t           a a a

```

Allineamento delle sequenze attorno al sito d'inizio della trascrizione di 19 geni per snoRNA; il sito d'inizio è indicato dalla freccia ed è quello riportato nel “Yeast snoRNA database”

(<http://people.biochem.umass.edu/sfournier/fournierlab/snornadb/main.php>).

È comunque necessario notare che, in un recente lavoro (Zhang and Dietrich, 2005), il sito d'inizio per *SNR4* è stato mappato 3 bp a monte rispetto a quanto noto finora in letteratura, quello per *SNR31* 1 bp a valle e quello per *SNR33* 6 bp a valle.

Allineamento delle TATA box individuate in *S. cerevisiae*

```

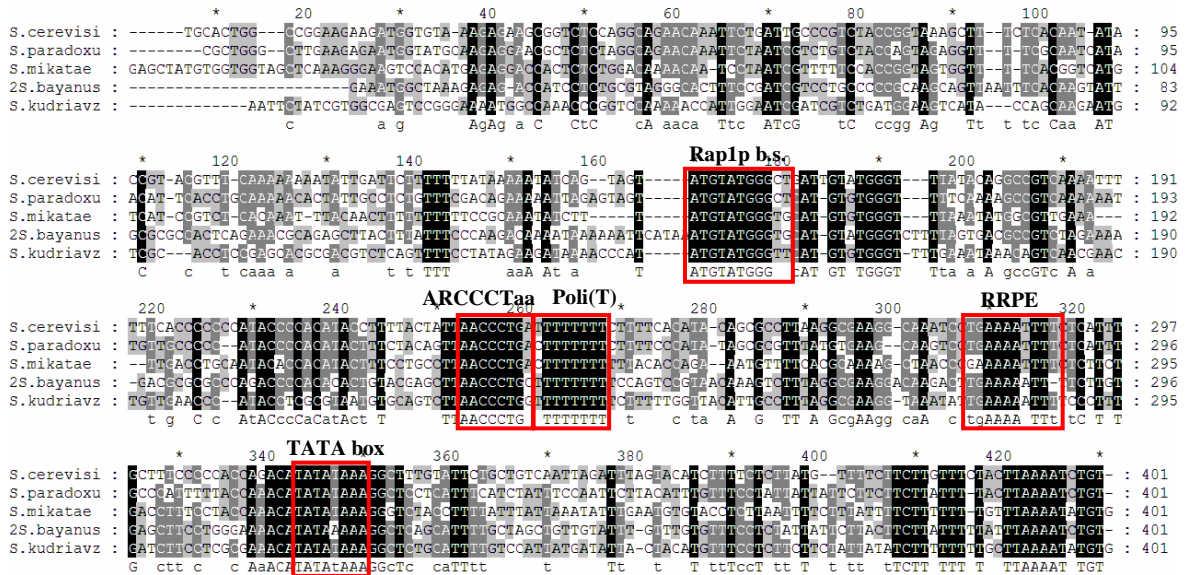
*
SNR48 : TTGTATTTTAAAGGCATC : 16
SNR31 : GTATATTTTAAAGCATG : 16
SNR67 : GAGTATTTTAAAGAACA : 16
SNR4 : ACCTATTTTAAATAGAA : 16
SNR65 : ACATATTTTAAACATAGG : 16
SNR68 : ACCTATTTTAAAGTGAA : 16
SNR189 : CTTTATTTTAAAGTGAGAT : 16
SNR17b : ACATATTTTAAATAGTGAG : 16
SNR64 : GAGTATTTTAAAGGGAGG : 16
SNR69 : GACTATTTTAAATGGAAG : 16
SNR39 : AAGTATTTTAAAGGATG : 16
SNR13 : ACATATTTTAAAGGGAAG : 16
SNR47 : TCAATATTTTAAAGGTGCC : 16
SNR56 : TCTATATTTTAAAGGGCGC : 16
SNR57 : TCATATTTTAAAGGATCT : 16
SNR40 : AAGTATTTTAAAGGCGGA : 16
SNR83 : CACTATTTTAAAGGCGAA : 16
SNR42 : GACTATTTTAAAGCCGTG : 16
SNR8 : TTGTATTTTAAAGCCGC : 16
SNR11 : TGGTATTTTAAAGCGGT : 16
SNR45 : AGGTATTTTAAAGCGAA : 16
SNR78 : GGGTATTTTAAAGAGGA : 16
SNR5 : GAGTATTTTAAAGCGAT : 16
SNR35 : AACTATTTTAAAGAGAT : 16
SNR37 : AAGTATTTTAAATAGAGCT : 16
SNR62 : CACTATTTTAAAGCAC : 16
SNR32 : CACTATTTTAAAGTCGA : 16
SNR9 : CACTATTTTAAAGAGAG : 16
SNR43 : CACTATTTTAAAGAGAA : 16
SNR49 : CGGTATTTTAAAGCAAAGT : 16
SNR85 : GTCTATTTTAAAGCTAGT : 16
SNR3 : AAGTATTTTAAAGGAAGC : 16
SNR10 : AGCTATTTTAAAGACAGA : 16
SNR81 : AGCTATTTTAAAGAGCCT : 16
SNR71 : AAGTATTTTAAAGAGGCC : 16
SNR41 : ACGTATTTTAAAGCACAA : 16
SNR60 : CCAATATTTTAAAGAAACAT : 16
SNR17a : ACATATTTTAAAGGCTT : 16
SNR161 : AGATATTTTAAAGCTTTA : 16
SNR30 : GCCATATTTTAAAGCACTT : 16
SNR190 : ATGTATTTTAAAGTACTGGA : 16
SNR58 : GCAATATTTTAAAGACTIAC : 16
SNR34 : GAATATTTTAAAGATACTA : 16
SNR63 : ATATATTTTAAAGTATATAT : 16
SNR80 : TTCTATTTTAAAGTATITC : 16
TATA aa

```

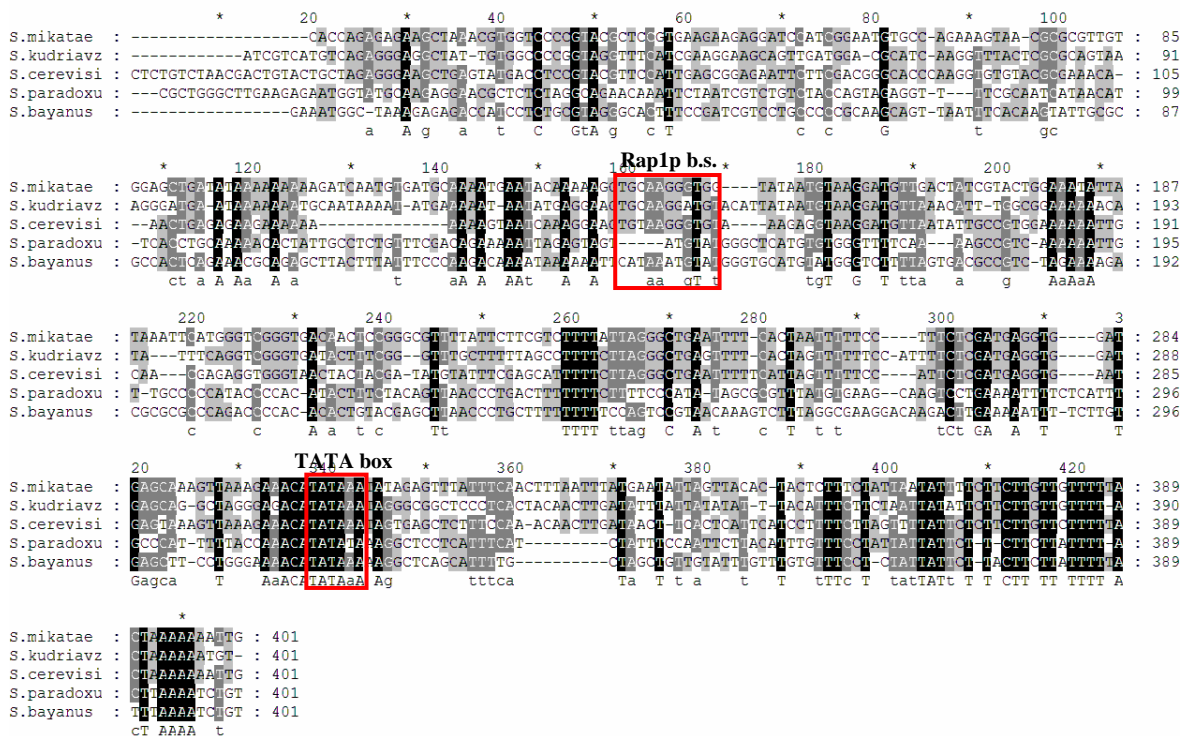
Sono state allineate le TATA box individuate a monte di 45 geni per snoRNA. Il consenso che si ricava è TATA(A/T)A(A/T)(A/G)(A/G). La base che precede la prima T è tendenzialmente una base diversa da T e la stessa tendenza si riscontra anche nelle altre quattro specie di Saccaromiceti.

Allineamento delle regioni a monte dei geni per snoRNA

SNR17A



SNR17B



SNR66

```

*      20      *      40      *      60      *      80      *      100
S.paradoxu : ----AGTAAGCAAAATACGTTTATCTATCCGA--GTTTCGGACCAAGACTTACCAAAATGTTTCACGACTATTAAATTTTTCCTACTTTAA--AGTCTTCAA : 97
S.mikatae : AAAAAGGAGTATCAAAATACGATTCTTAGCTTTG--IAGCAC--TCAACCAATTAACACTCT--TTTCAGTGTCTATCGGTTTACTTCTAATTAA--FAGTCTTCAA : 101
S.cerevisi : ----AGGAGTACGAAATACGTTTGTATCTCTA--GTTTCGGTAAAGACTATATAGTGGCTTTCACCTGTTTTCTAA--TTTCTTCCTTTAA--AGTCTTCAA : 99
S.kudriavz : ----AATAAGAAAACGAGCTTAAATCCCAACTTAAAGTAAAGACTTCAACA--ATT--CITTTAAGCTTATTATTAACCTTAGGGTGA--AGTCTTCAA : 96
S.bayanus : ----AAGAAAACACCTTAGAATTTCTAGAAAATGGAATGAGCTTCCCTGTGTATT--TTAATACTTCCCAATTAAATTTTGATTTAGAGTCTTCAA : 95
      a AaGAAA aAC TT t c t g a aa t aa T tt tT g T a T cTT t TttaA AgtTCTCAA

NME1 →
*      120      *      140      *      160      *      180      *      200      *
S.paradoxu : TCCATACCAAAGAATCGTCACAAATCCAAGCTTACAAAA--TGGAGTGAAT-----TTTTCCTCTCAATATA--TGCTTTGGGTTGAAAGTCTCCCAACAATT : 195
S.mikatae : TCCATACCAAAGAATCGTCACAAATCCAAGCTTACAAAA--TGGAGTGAAT-----TATTT--CGCTCAGTARTA--TGCTTTGGGTTGAAAGTCTCCCAACAATT : 198
S.cerevisi : TCCATACCAAAGAATCGTCACAAATCCAAGCTTACAAAA--TGGAGTGAAT-----TTTTCCTCTCAATATA--TGCTTTGGGTTGAAAGTCTCCCAACAATT : 197
S.kudriavz : TCCATACCAAAGAATCGTCACAAATCCAAGCTTACAAAA--TGGAGTGAAT-----TTCCTTTTCTCAATATA--TGCTTTGGGTTGAAAGTCTCCCAACAATT : 196
S.bayanus : TCCATACCAAAGAATCGTCACAAATCCAAGCTTACAAAA--TGGAGTGAAGGAATTTTTCCTCTCAATATA--TGCTTTGGGTTGAAAGTCTCCCAACAATT : 201
      TCCATgACCAAAGAATCGTCAC AAATCCAAGCTTACAAAA TGGAGT gaa t tttt CTCA TA TA TGCTTTGGGTTGAAAGTCTCCCAACAATT

*      220      *      240      *      260      *      280      *      300      *      3
S.paradoxu : CATATGCGGAAAACGTAATGAGATTTAAAAATTTTAAATGTTTAAATCAACTCATTAAAGGAGGATGTCCTTGGATATTCTCTTCTTGACCTGGTACCTCTATT : 301
S.mikatae : AATATGCGGAAAACGTAATGAGATTTAAAAATTTTAAATGTTTAAATCAACTCATTAAAGGAGGATATCTCTTGGATTTCCCTCTTGACCTGGTACCTTTTATT : 304
S.cerevisi : CATATGCGGAAAACGTAATGAGATTTAAAAATTTTAAATGTTTAAATCAACTCATTAAAGGAGGATGTCCTTGGGATTTCTCTTCTTGACCTGGTACCTCTATT : 303
S.kudriavz : CATATGCGGAAAACGTAATGAGATTTAAAAATTTTAAATGTTTAAATCAACTCATTAAAGGAGGATACCTTGGGATTTCTCTTCTTGACCTGGTACCTTTTATT : 301
S.bayanus : CATATGCGGAAAACGTAATGAGATTTAAAAATTTTAAATGTTTAAATCAACTCATTAAAGAGATATCTTTTAAATTTCTCTTCTTGACCTGGTACCTCTTATT : 306
      caATATGCGGAAAACGTAATGAGATTTAAAAATTTTAAATGTTTAAATCAACTCATTAAAG ag AT cctTgg TATTc CTCTTGACCTgTAC t T TTA

*      340      *      360      *      380      *      400      *      420
S.paradoxu : CAGGCTGCTGGTGTCTTCTTCGGTATGGATTCCCTATAATGGAATCTAAACCATAGTTATGACGATTGCTCTTCCCAGCTGGATGAGTAACCCAATGGAG : 407
S.mikatae : CAGGCTATCATGTTTCTTCGGTATGGATTCCCTATAATGGAATCTAAACCATAGTTATGACGATTGCTCTTCCCAGCTGGATGAGTAACCCAATGGAG : 410
S.cerevisi : CAGGCTGCTGGTGTCTTCTTCGGTATGGATTCCCTATAATGGAATCTAAACCATAGTTATGACGATTGCTCTTCCCAGCTGGATGAGTAACCCAATGGAG : 408
S.kudriavz : CAGGCTGCTGGTGTCTTCTTCGGTATGGATTCCCTATAATGGAATCTAAACCATAGTTATGACGATTGCTCTTCCCAGCTGGATGAGTAACCCAATGGAG : 407
S.bayanus : CAGGCTGCTGGTGTCTTCTTCGGTATGGATTCCCTATAATGGAATCTAAACCATAGTTATGACGATTGCTCTTCCCAGCTGGATGAGTAACCCAATGGAG : 412
      CA G g ctgGTGTTTTCTTCGGTA TGGATTCC TAtatATGGAATCTAAACCATAGTTATGACGATTGCTcttTCCCAGCTGGAT GAGTAACCCAATGGAG

*      440      *      460      *      480      *      500      *      520      *
S.paradoxu : CTTACTATCTTGGTCCATGGATTACCCCTTTTAACTTTTGAACAGTTTGTGTTTACGCTT--TTTTTAAAGCTAAAACCTTTTGCATAGTTAAATATGCCGG : 512
S.mikatae : CTTACTATCTTGGTCCATGGATTACCCCTTTTAACTTTTGAACAGTTTGTGTTTACGCTT--TTTTTAAAGCTAAAACCTTTTGCATAGTTAAATATGCCGG : 511
S.cerevisi : CTTACTATCTTGGTCCATGGATTACCCCTTTTAACTTTTGAACAGTTTGTGTTTACGCTT--TTTTTAAAGCTAAAACCTTTTGCATAGTTAAATATGCCGG : 514
S.kudriavz : CTTACTATCTTGGTCCATGGATTACCCCTTTTAACTTTTGAACAGTTTGTGTTTACGCTT--TTTTTAAAGCTAAAACCTTTTGCATAGTTAAATATGCCGG : 509
S.bayanus : CTTACTATCTTGGTCCATGGATTACCCCTTTTAACTTTTGAACAGTTTGTGTTTACGCTT--TTTTTAAAGCTAAAACCTTTTGCATAGTTAAATATGCCGG : 511
      CTTACTATCTTGGTCCATGGATTACCC TTTT t TTT AaCaGTTTGTGTTTAC Cta t ta AG GTAAAACCTTTTGCATAGTTAAATATGCC a

*      540      *      560      *      580      *      600      *      620
S.paradoxu : ACTGACAGCAATA--AAGAGCTGAATA--CTCAGAGAGTTATTATTAGTACTTGTGTAAGA--AGCAGATGGATGGTGAATAACT : 600
S.mikatae : ACTGACAGCAATA--AAAAGCTGAATA--CTAAGGGAGTTATTAAAGTAACTTGTGCAAGA--ACAAGATGGATAGGAGTAATA : 600
S.cerevisi : ACTGACAGCAATA--AAAAGCTGAATA--CTCAGAGAGTTATTATTAGTAACTTGTGTAAGA--AGCAGATGGATAGGATAACT : 600
S.kudriavz : ACTGACAGCAATA--AAAAGCTGAATA--CTCAGAGAGTTATTATTAGTAACTTGTGTAAGA--AGCAGATGGATAGGATAACT : 600
S.bayanus : ACTGATTAGCAATA--AAAAGCTGAATA--CTCAGAGAGTTATTATTAGTAACTTGTGTAAGA--AGCAGATGGATAGGATAACT : 600
      ACTGA a AATA AaAaG CTGAA Aa CTCAGA GAGTTATTCatt GT a CTTGT TAAGA Ag A gATGG ata tGAaTAAT t

```

Negli ultimi due blocchi è presente la sequenza codificante per la sequenza "leader" di snR66.

SNR84

```

*      20      *      40      *      60      *      80      *      100
S.cerevisi : -----GACTAGATATTACTGCTCGAAGGGCCATCTCCATCTTTATCACACATGGTGGCTGCGCGPAAATAGTAAAGGCTGCGGTAAATAAGCTTT----GAGAA : 96
S.mikatae : -----CATATTA-TGCT-TCANAAGTCAATTCACGGGTATCGGCTAAGGT-TTCGGGAGTAA-TAGTGTAAATACAAATAPAAATATCTCTCGGAGAT : 91
S.paradoxu : ATGATAATCCGGATATACTAATGTCCGGGGCCACTGCAATTATACACAGGTSTTCTCGCGGAAA-TGAAAGACTATACCAATGAAATTTCTAGGAGAAA : 105
S.kudriavz : -----TGA-CTGTGTTCCATCTCC-CCCAATCAACAGGTGCACACGTAGAGCAG-AAAAGTTTCCACAAA-CAAAATTCATTCAAG : 82
S.bayanus : -----TATCT-CCATAACACACAG-TACTTGGGTGGGGGAGGCGGATGAGTTTTTAA-AGCTTAATAAAT : 69
          a t      gat Tgc cg Atc acAaGgT t g G G a a t t aat t cAA a t T A A

*      120      *      140      *      160      *      180      *      200
S.cerevisi : TCT--TCCAGTG----GATAAATGTTACCTTGCATTT----GTCTGPAATTTCAGGTGACAAGAGAGACCACGAAATCCATAGAGGT--TAGTA--AAA : 186
S.mikatae : TTTCGGTCGACTTG--TCAAGTATACTAACCTGGAGTGGG-GTGATTIACAGGCTCGTGTGCAATTTATTCGCCAATAAATCAAGTGTAAATGGAGCTTTG : 194
S.paradoxu : TG--TCCAGCA--CAAATCTATTACCTTATATATC-GTGATAGAAATCTTCAAGTCT-----AAATCCAACAGGT--TAGTT--AAA : 182
S.kudriavz : CTTIAGACAATAAGCTTTCATGCCATTACGCTTAPAAATAG-ATGGCGGTAAACAGCAGTCTT-GAATAGCCACTACTGGGTAGTATTGGCCGCTCGCGG : 186
S.bayanus : GTATCCCCCATGAG---ATATATCACGCCCTAACTAGCCATGCTCTTCTACTACAGTCTTT-TAGAGCCCTAATGCGCCCTGCCCCGGATTGCGCTTACGCG : 170
          C g      A T      Cc t a T a c tGa a cT GacGTG t g ca a c t aG t t G c

220      Reb1p b.s.      240      * RRPE 260      *      280      * PAC      300      *      3
S.cerevisi : AGAA--TCCAGCGGGTATCCTTACTAGTCAAAAGTGA---AAAATTTTAAACTATAAATAAAGCAAGCATGATCA-GAAAGAGCATG-----TAA : 273
S.mikatae : CCAI--CCDAGCGGGTATCCCPAAATATCTCAAAAAA-----TTTCTTAACTATAATA-AAAGCAAGCATCA-GAAAGATTACA-----TAC : 275
S.paradoxu : TAA--TCCAGCGGGTATCCCPAAAGAGTCAAAATTT-----TTTCTTAAATGATAAATAAAGCAAGCATCA-GAGAGCATGAGACTTACGACGAAA : 274
S.kudriavz : CTG-GTTCAGCGGGTATCCCPAAATTTG-TGTCGTGA---AAAATTTTCCPAAAGTATAAATAAAGCATGATTA-GAGPATG-----TAA : 272
S.bayanus : CAAAGCTTACCGGGTATCCCGAAGCCATCCCGCAGTCGTCAAAATTTTCCPAAATGATCAAAATAAAGCATGATGAAGATTAAC-----A : 263
          at tccAGCGGGTAT C CAat g Tg tttt Aa ta TA A AAA At SCGATGA GA GA Atg t a

20      *      340      *      360      *      380      *      400      *      420
S.cerevisi : CTACTAATCAAGCTACTTA-CAACGAAAGAGCA-----GTCAGGATCCCGAGTTTAAAGTTTCAATAGAGACTAG-TTAATGAGAGTCTCAT : 363
S.mikatae : GAACA--TACCTTCCG-CAAAAGAAAGAAAT-----GTCAAAATCTCCAGTTTAAAGTTTCAATAGAAACTTCC-TTAATGAGAGTCTCAT : 362
S.paradoxu : CTAACCTCGAACHAATA-CAAAAGAAAGAAAT-----GTCAGGCTCCGAGTTTAAAGTTTGAATAGAAACTAG-TCCATGAGAGTCTCAT : 363
S.kudriavz : ACTTAGTAAACTTCTACTCAACGAAAGAAAC-----GTCAAAAGTCTAAGTTTAAAGTTTCAATAGAAAGTCCATTAATGAGAGTCTCAT : 363
S.bayanus : CAAATTTTCACTGTCACTAAAGAAAGAAATAAAAAACAACCGTGTCAAGTCCGAGTTTAAATTTGATGATAAAGTTTAA-TTAATGAGAGTCTCAT : 368
          a t a cT c a cAAA GAA A Aa GTCa a t C gaAGTT AaAgTtTGaaT A AaAg T G TtaATGAGAgT CTCAT

*      440      *      460
S.cerevisi : TATCAGACTATAGCTTTATATGGATATG-CTGATGTC- : 400
S.mikatae : TATCAGACTATAGCTTCTGTATATGTAAACTGACATG : 401
S.paradoxu : TATCAGACTATAGCTTTATATGGATATA-CTGACATC- : 400
S.kudriavz : TATCAGACTATAGCTTTATATGGAGAAA-TTGCATG : 401
S.bayanus : TACCGTATATAGTTT-TATGGCAAC-TCGAC---- : 400
          cAtCa AcTAAaGcTt TATgga A tgAc tg

```

Gli ultimi due blocchi dell'allineamento comprendono probabilmente la sequenza codificante per la sequenza "leader" di snR84.

SNR87

```

S.cerevisi : -----*-----20-----*-----40-----*-----60-----*-----80-----*-----100-----
S.cerevisi : -----CTAGGCTCTCTCCGTTGCCAATGGGCT--GCAATGGAAACAGTCCGGGCTGAAGGAAACCGATTAATGGATA--AATGCCGGAAGCT : 86
S.paradoxu : -----GGTACTCTCTCCCTGGTTACTTGGGCT--GACAGTAGAGAGCGTGTGGCTGGGACGAATCC--GATGATGGAG--AATGCCCTCAATC : 85
S.mikatae : -----CGCCCTCCATGCTTGGGGTAAAGATGCTAGGGAAGAGAGGAAACCGC--TGCTGGAG--AGTTCCGCCAACT : 72
S.kudriavz : -----GCTAGGCT--GGCATAAGACATGCTTGGACTAAGAGGAATG--GATCTGGAG--AGCGGTTCAACT : 65
S.bayanus : ACAAAGTCTGCCAAGGCTCACTCCCGCTCCAGCCAGTAAAGCT--GGTGTAGATGGATCCGTCACAGAGGGAAGTAAATCTACAGGAAATGTCACCAACAG : 104
          g          T gGCT g G TagA A g Tg Gc C gagGaaA c gaTg TgSaag A TgC c caac

S.cerevisi : -----*-----120-----*-----140-----*-----160-----*-----180-----*-----200-----*-----
S.cerevisi : -----GCPAGAGG--AATCTCTGGTAAACCGCTTTCGGGGGGAAGTAAAGCGACACCCCTTGAAGCTTCAATPAATGGAAA--TCCAAAATAA--TGATPAC : 187
S.paradoxu : -----GCAAGGAG--AATCTCCCAAAAGGCGCTTTCGGGGGGAAGTGGGGAGTACC--CTTTGA--ACTTCCATAAACGAAATCCGAAAATAAATGATPAC : 189
S.mikatae : -----GCGAGGAG--AACA--CTATACAGCTTCTTCCGGGGGAAGT--ACAAGGCTGCC--CTTGGAGCCAGTAAAGTGAAPAACTAATAAATAATGATPAC : 173
S.kudriavz : -----ACGAGGAG--AGCTCCCAAAAGTTCCTCTCCGGGGGGAAGT--GTGGAGGAAAGT--GGCGGCTGATTTGAAGGAAAGTAAATTTGATPAGT : 164
S.bayanus : -----GGGAAAGCTATTTCCCAAAAGGCGCTTCTCCGGGGGGAAG--GGCACCAACACAC--AGCCCAAGTATGAA--GCAAAATAA--TCATCGG : 197
          gca gCA C A ccca AaAgc ctCT T C GgGCGG aAGT G G g t agC gt a GcAa a c AaAaAa TGA a c

S.cerevisi : -----*-----220-----*-----240-----*-----260-----*-----280-----*-----300-----*-----3
S.cerevisi : -----GCAACTCAATTTGAAATTTACACATAATGATGATGGTAAAGATGTA----AAA--TTGTGAAAGGGTCTATTTGGCCCTGGCAATTTACTCAGTCA : 287
S.paradoxu : -----CAGATTTCAATTAAGAAATTTCTTAAATAATAGGGGTAAAGATGTAAT--GATG--ATGTGAAAGGATGTAATGGCGCAATGATTTACTCAGA : 291
S.mikatae : -----AATAAAACGACACGCTTTTCTCA--ATAA--ATGGGTAAAGATGTAATGAACT--TTGTGAAAGGGTCTATTTGGCTGGCCCTTTCT : 266
S.kudriavz : -----GAG--CAGATGTAATTTT--CTAATGAG--TGGGGTGATGATGTAAGGCGCT--TCGTGAAAGGGTGC--AATTTGCTGGCGCAATTTGTTGAA : 261
S.bayanus : -----GAAAGCAAACTACGGGCTTGTACTTTAGATGCGGGTGGGATGTAATGGGCATC--ATGAAAGGGTGC--TCAATTTGCTGCTTT : 292
          g a a c A a t tttt t At aTgTgTGT aGATGTA a ga ttTG ATgGgTG T T TgctGtc tt t

S.cerevisi : -----*-----20-----*-----340-----*-----360-----*-----380-----*-----400-----*-----420-----
S.cerevisi : -----GATAACTACCTTCCGCAAAATTTGGTGTGTTT-----TTTACTTCT-----ACACTAAAGCTGGATT--TTTTTTTAA--PACCTACTTCTTC : 378
S.paradoxu : -----AATAACGCGCTTTTACAGA--TTACTGGCTTCTCTTT-----TTTTTTTCT--GACACTAAACGAAATAA--ATCTTTTAA--PACCTAGGTTTTC : 384
S.mikatae : -----AATGCTTCTCT--GACAGGATTTATTTGTTCTA--TTTT-----TTCAITTTGATACACTGAAACCAACGGG--TACTTTGCCCAATAGTACCTA : 360
S.kudriavz : -----AATGAACTAATC--GCCAAGGTTTTTTTTTTGTTCTCTGGTTTGTCTTCTTTTCTTTTCACTCGACTATTC--CGCCCTAACAAATGAATCTTAATA : 363
S.bayanus : -----AATTTCTA--GCCACATAATCTTTTAACTGGTA-----TTTTTTTTTACATTTGATA--TGAATGCTAGCTGAGGAGGACTGCTAATA : 383
          t Ag t a G CAa Tt T gtTt t tT ttt tttt cACA A t a g tta A a Tag T t T

S.cerevisi : -----*-----440-----*-----460-----*-----480-----*-----500-----*-----520-----*-----
S.cerevisi : -----AGT--GACTCCAAG-----TG--TATCTCTCTTGAATTTAAAAGCTTCAATTT--TAAATCCGATGACTGTTATGCGCAGCTGCTGGAGATGACAT : 470
S.paradoxu : -----CGGT--AAACTAAATA-----CC--TAGCCGCTTCCGATTCGATAATCTTA--TGA--TAAATGCTGACATGATTAATGACCTGTGAAATTTAGATGAT-- : 474
S.mikatae : -----AAGCTAGACATTAAGGTTTGTTT--TAGATATGCCCGTTAAAGTCTTTTCACTCTGCTCCATAGTGCTACTGTTCCCATATGACGTAACCTCGGATTTTT : 464
S.kudriavz : -----AAGCTAGATTTCCAACTCTATCAAAATTTTAAATCTTCCCTTGGCCACTCTCTAATTAATCCTTTTCACTGAGTAAATGACAGAGATAAATGGGC : 469
S.bayanus : -----AAGCTTTCTTTCA-----CAGATATCAATTC--TTGGAAATATATTTAAATATCCATTTAATAAACAATCCGAGGCTGAATTT--CCCAGGCTCAT : 476
          a A tt t TA t g Tt t t t a t Tt aaat c T a t at

S.cerevisi : -----*-----540-----*-----560-----*-----580-----*-----600-----*-----620-----*-----
S.cerevisi : -----CACAAAGTCTTCACTAGAAACAATAAATATATTACAAATCTGTTTCAGAAATGATATGATACAGTGGTATCATATTTGTTCTTTTTCAGGTACCAGTGTCT : 576
S.paradoxu : -----TCTCAATCTPAGCAAAAGTACATAGCAGCCATGTTAAGAAATGGTATGATACAGTGGTATCATATTTCTCTCCGATTCAGACCGCCAGTGTCT : 573
S.mikatae : -----GGTTAAGTCTTAATAGCAAAAAGTCAATGAGGACCGTTPAGAAATGATATGATACAGTGGTATCATATTTTCTTTTAACTTGAACCTTGTAGTCTCT : 570
S.kudriavz : -----TACCCAGCATTGAATAGGCAAAAAGTAAATGTGAGACCGCTTGAATGACTGATGATACAGTGGTATCATATTTCTCTCTTTAATGAGTAAATGAGTGTCT : 574
S.bayanus : -----CATTGTGTTGGATAGATAAATACTGATCTGAAACGGTGTGAATGACTACATGATACAGTGGTATCATATTTCTCTCTTTAATGAGTAAATGAGTGTCT : 579
          gtgtT aaTga cAtAAA t ATA acg GTTaaGATGA tATGATACAGT GIATCaTaTTTctTcttaa Tg a gtAGTCT

S.cerevisi : -----*-----640-----*-----660-----
S.cerevisi : -----TATCT--AGACATTCATGCAAG----- : 600
S.paradoxu : -----TGTTCAPAGGCAATTCGCAAGAG----- : 600
S.mikatae : -----TTTCAT--AAGGCGCCGCAAGPAGGTATT : 600
S.kudriavz : -----TATTAC--AGTGGACCCGCAAGCAATC---- : 600
S.bayanus : -----TAT-----AAGCTCTCAGAGAGAC----- : 600
          T T a A aG Ca aag Aa
    
```

Gli ultimi due blocchi dell'allineamento comprendono probabilmente la sequenza codificante per la sequenza "leader" di snR87.

RINGRAZIAMENTI

A più di tre anni dai ringraziamenti della mia tesi di laurea, ecco quelli della tesi di dottorato: forse saranno un po' banali ma ho pensato che non sarebbe stato giusto non farli proprio. Quindi...

Per primo ringrazio il mio capo, il prof. Giorgio Dieci: chiunque sarà nei prossimi anni il mio superiore avrà un modello difficile da eguagliare.

E ora tutte le mie tesiste (in ordine strettamente cronologico) e cioè Eleonora, Barbara, Chiara, Daniela e Roberta: è anche per merito loro se questi anni sono volati, tra Northern, lieviti e chiacchiere. Ragazze, spero di essere stata una buona insegnante in lab! Senza di voi questo lavoro non sarebbe stato possibile!

Ringrazio tutti i dottorandi (e dottori di ricerca) del dipartimento, ma in particolare quelli del mio lab, Gloria, Andrea e Laura, che ora avranno più spazio (salvo invasioni da parte di orde di tesisti); approfitto per ringraziare anche i dottorandi che mi hanno preceduta, Priscilla, Roberto ed Elisa che mi hanno fatto da "balie" durante il periodo in lab prima della laurea e durante il primo anno di dottorato. Ringrazio poi anche Roberta e Barbara, i dottori di ricerca che mi hanno aiutato in alcuni punti del mio lavoro sugli snoRNA. Non posso dimenticare neanche Ilaria, un po' mia ex tesista "ufficiosa" e anche mia ex coinquilina (con la Gloria), e la Roby: da quando è a San Diego le mie serate latino-americane si sono drasticamente ridotte!

Prima ho parlato di chiacchiere quindi come non ringraziare Chiara B., con cui io e la Chiara abbiamo abbondantemente spettegolato in quest'ultimo anno.

E ora la Gessica: abbiamo condiviso questi quattro anni in dipartimento e mi considero molto fortunata di averla come amica. La sua amicizia è sempre stata preziosa per me, nei momenti felici e in quelli che lo sono stati un po' meno; penso che la piccola Oriele (che è nata una settimana fa, ma a cui voglio già bene) sia fortunatissima ad averla come mamma!

Vorrei ringraziare i ragazzi che dopo la laurea sono andati a lavorare lontano da Parma e cioè Danilo e Francesco (anche se ormai ci vediamo poco) e poi Matteo che mi permette di "evadere" ogni tanto da Parma: non credo che ci siano parole per dire quanto questo sia importante per me.

Grazie ai miei amici Ilaria, Max ed Elisa per la loro preziosissima amicizia.

Infine grazie ai miei genitori per tutto quello che hanno fatto per me.