Mol Genet Genomics (2011) 286:395–410
DOI 10.1007/s00438-011-0635-y

ORIGINAL PAPER

RNA splicing and debranching viewed through analysis of RNA lariats

Zhi Cheng · Thomas M. Menees

Received: 11 March 2011 / Accepted: 30 June 2011 / Published online: 8 November 2011

Abstract Intron lariat RNAs, created by pre-mRNA splicing, are sources of information on gene expression and structure. Although produced equivalently to their corresponding mRNAs, the vast majority of intron lariat RNAs are rapidly degraded. However, their levels are enhanced in cells deficient for RNA debranching enzyme, which catalyzes linearization of these RNAs, the rate-limiting step in their degradation. Furthermore, RNA lariats are resistant to degradation by the 3′ exonuclease polynucleotide phosphorylase (PNPase), providing a means to enrich for lariat RNAs. Working with the yeast *Saccharomyces cerevisiae* as a model organism, our goal was to develop novel combinations of methods to enhance the use of intron lariat RNAs as objects of study. Using RT-PCR assays developed for detecting and quantifying specific lariat RNAs, we demonstrate the resistance of RNA lariats to degradation by PNPase and their sensitivity to cleavage by RNA debranching enzyme. We also employ sequential treatments with these two enzymes to produce characteristic effects on linear and lariat RNAs. We establish the utility of the methods for analyzing RNA debranching enzyme variants and in vitro debranching reactions and discuss several possible applications, including measuring relative rates of transcription and combining these methods with non-gene-specific RNA sequencing as a novel approach for genome annotation. In summary, enzymatic treatments that produce characteristic effects on linear and lariat RNAs, combined with RT-PCR or RNA sequencing, can be powerful tools to advance studies on gene expression, alternative splicing, and any process that depends on the RNA debranching enzyme.

Keywords Intron RNA lariats · Debranching enzyme Dbr1 · Polynucleotide phosphorylase · mRNA splicing · Yeast Ty1 retrotransposon

Communicated by A. Aguilera.

Electronic supplementary material The online version of this article (doi:10.1007/s00438-011-0635-y) contains supplementary material, which is available to authorized users.

Z. Cheng · T. M. Menees (✉)
School of Biological Sciences, University of Missouri-Kansas City, Kansas City, MO 64110, USA
e-mail: meneest@umkc.edu

Present Address:
Z. Cheng
The Pediatric Surgery Laboratory, Department of Surgery, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

Introduction

Pre-mRNA introns play an important role in the regulation of gene expression for many eukaryotes because their presence allows for the occurrence of alternative splicing, which results in the creation of multiple proteins from a single gene, many of which are expressed in cell- or tissue-specific patterns (Hallegger et al. 2010; Nilsen and Graveley 2010). Introns are excised in a lariat conformation (Abelson 2008; Smith et al. 2008; Wahl et al. 2009) and, following excision, the 3′ tails of the lariats undergo exonucleolytic degradation up to the lariat branch point (Chapman and Boeke 1991; Salem et al. 2003). The predominant pathway for further exonucleolytic degradation requires cleavage of the 2′–5′ bond at the branch point by RNA debranching enzyme, a 2′–5′ phosphodiesterase found in all eukaryotes (Ooi et al. 2001).

Although intron RNA sequences contain information necessary for their removal from pre-mRNAs, some introns contain additional information. In most eukaryotes microRNAs (miRNAs) (Rodriguez et al. 2004; Kim and Kim 2007; Tang and Maxwell 2008) and small nucleolar RNAs (snoRNAs) (Filipowicz and Pogacic 2002; Lestrade and Weber 2006) are encoded within introns. In studies with human cells it has been found that the vast majority of intronic miRNAs are excised from pre-mRNAs by Drosha (Kim and Kim 2007; Kataoka et al. 2009). However, mirtrons, which have been found in fruit flies, worms, and vertebrates, are miRNAs processed from introns that have been excised from pre-mRNAs by the spliceosome (Berezikov et al. 2007; Okamura et al. 2007; Ruby et al. 2007). Intronic sno-RNAs are also processed from excised introns, as determined in baker's yeast, humans, and other eukaryotes (Ooi et al. 1998; Filipowicz and Pogacic 2002; Lestrade and Weber 2006).

Debranching and subsequent degradation of most intron RNAs are rapid, resulting in low steady state levels of intron RNAs relative to the levels of the corresponding mRNAs, as determined by studies in yeast (Chapman and Boeke 1991; Salem et al. 2003). The exceptions are intron sequences corresponding to RNAs with additional functions (e.g., snoRNAs). Studies in many different organisms have determined that cleavage of the 2′–5′ bond by RNA debranching enzyme is important for maturation of intron-encoded snoRNAs and mirtrons (Ooi et al. 1998; Kiss et al. 2006; Berezikov et al. 2007; Okamura et al. 2007; Ruby et al. 2007). For mirtrons, RNA debranching enzyme acts instead of Drosha (Berezikov et al. 2007; Okamura et al. 2007; Ruby et al. 2007).

Genome-wide studies analyzing excised intron RNAs in fruit flies and yeast have identified introns and (for flies) alternative splicing patterns (Conklin et al. 2005; Juneau et al. 2007; Zhang et al. 2007). These analyses relied on creating cell populations that accumulate excised intron RNAs at elevated levels either due to mutation of the gene encoding debranching enzyme in yeast (Juneau et al. 2007; Zhang et al. 2007) or knock down of debranching enzyme expression with siRNA in flies (Conklin et al. 2005). Cells defective for RNA debranching activity accumulate excised introns in their lariat forms, with shorted 3′ tails, as described in studies with yeast mutants (Chapman and Boeke 1991; Salem et al. 2003). Therefore, information on the 3′ intron–exon junction is not obtainable from intron lariat RNA sequences. However, unlike other methods of identifying intron sequences, studies in plants and human cells have shown that the positions of RNA branch points can also be deduced from analyzing intron RNA lariats (Vogel et al. 1997; Gao et al. 2008).

Our interest in intron RNA lariats comes from the observation that RNA debranching enzymes are host factors for retroviruses and retrovirus-like elements. The human RNA debranching enzyme, Dbr1, is a host factor for the retrovirus HIV-1, acting to promote retroviral reverse transcription (Ye et al. 2005; Bushman et al. 2009). Previously, the *S. cerevisiae* RNA debranching enzyme, Dbr1p, had been found to be a host factor for both the Ty1 and Ty3 retrotransposons, which are related to animal retroviruses in their genomic structures and replication cycles (Chapman and Boeke 1991; Karst et al. 2000; Griffith et al. 2003; Irwin et al. 2005). For Ty1, Dbr1p also appears to promote reverse transcription (Karst et al. 2000; Griffith et al. 2003). In fact, we have presented evidence that Ty1 RNA forms a lariat and we proposed a model in which the formation of the lariat and its later debranching facilitate both the minus strand transfer reaction and extension of the minus strand that occur during retroviral reverse transcription (Cheng and Menees 2004; Goff 2007; Engleman 2010). However, subsequent work by others has raised concerns about the interpretation of our results and our model for the role of Dbr1p in retrotransposition (Coombes and Boeke 2005; Pratico and Silverman 2007). Although siRNA knockdown of human Dbr1 interferes with production of HIV-1 cDNA at the precise stage of reverse transcription predicted by our disputed model (Ye et al. 2005; Bushman et al. 2009), there is currently no further evidence that either HIV-1 RNA or Ty1 RNA forms a lariat.

Since RNA debranching enzymes act predominantly on excised intron lariats, we are interested in developing tools for studying these lariat RNAs, which may hold the key to understand the role of these enzymes as host factors for Ty1 and HIV-1. *S. cerevisiae* cells lacking a functional *DBR1* gene, which encodes the RNA debranching enzyme, accumulate excised intron RNA lariats (Chapman and Boeke 1991; Salem et al. 2003). Since *S. cerevisiae dbr1* mutants exhibit nearly wild-type growth capabilities, the mutant cells are a robust source of excised intron lariats. Human and *Drosophila* cells, in which expression of RNA debranching enzyme has been knocked down with siRNA, have also been used as sources of intron RNA lariats (Conklin et al. 2005; Ye et al. 2005).

In this work we have developed methods for detecting and quantifying specific RNA lariat species and for treating RNA populations to enrich them for RNA lariats. We establish the utility of these methods by analyzing the functionality of Dbr1p mutants and to observe time courses of debranching reactions. The techniques described can be adapted to quantify transcription levels for specific genes as well as to perform genome-wide analyses to identify introns, including their branch point sequences, and analyze alternative splicing patterns.

Materials and methods

Yeast and bacterial strains, plasmids, and general procedures

The following yeast strains were used: TMY30 (*MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1*), TMY60 (TMY30 *dbr1::neo*^{r}), TMY497 [=TMY30 mutated to *dbr1* (D180Y allele)], TMY498 [TMY30 mutated to *dbr1* (G84A allele)], TMY499 [=TMY30 mutated to *dbr1* (Y68S allele)]. TMY453, a *dbr1Δ::hisG* version of sigma strain 10560-23C (Loeb et al. 1999), was used for *FLO8* RT-PCR experiments (sigma strain 10560-23C = *MATalpha ura3-52 his3::hisG leu2::hisG*). The *dbr1Δ::hisG* allele was created using pTM513, a *DBR1* gene blaster plasmid (Alani et al. 1987) containing *dbr1Δ::hisG-URA3-hisG*, and targeted to replace *DBR1* chromosomal sequences by digestion with *Pvu*II.

The following *Escherichia coli* strains were used: Rosetta DE3 (Novagen) [F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (*DE3*) *pLysSRARE* (*Cam*^{R})]; XL1 Blue [*F′::Tn10 proA*$^+$*B*$^+$ *lacI*^{q} Δ(*lacZ*)*M15/recA1 endA1 gyrA96* (*Nal*^{r}) *thi hsdR17* (r_k^- m_k^+) *supE44 relA1 lac*]; JM109 [*F′ traD36 lacI*^{q} Δ(*lacZ*)*M15 proA*$^+$*B*$^+$*/e14*$^-$(*McrA*$^-$) Δ(*lac-proAB*) *thi gyrA96* (*Nal*^{r}) *endA1 hsdR17* (r_k^- m_k^+) *relA1 supE44 recA1*]; ES1301 [*lacZ53 thyA36 rha-5 metB1 deoC IN*(*rrnD-rrnE*) *mutS201::Tn5*]; and TOP10 [*F-mcrA* Δ(*mrr-hsdRMS-mcrBC*) *Φ80lacZΔM15* Δ*lacX74 recA1 deoR araD139 D*(*ara-leu*)*7697 galU galK rpsL* (*Str*^{R}) *endA1 nupG*].

The following plasmids were used for this study: pET16b-DBR1 was a gift of Beate Schwer and was used to express Dbr1p in *E. coli* (Martin et al. 2002). pRS306 (Sikorski and Hieter 1989) was used as a *URA3* template for making a PCR fragment to create a *dbr1Δ::URA3* allele at the *DBR1* locus. YEp351 (*LEU2*) (Hill et al. 1986) was used in cotransformations with the PCR fragment that resulted in the creation of a *dbr1Δ::URA3* strain. This strain was an intermediate in the creation of *dbr1* point mutants (see section below on the creation of *dbr1* mutant strains). pTM431, pTM432, and pTM435 were all created by random mutagenesis of pYES2/GS-DBR1 from Invitrogen and encode Dbr1p D180Y, Dbr1p G84A, and Dbr1p Y68S, respectively (Salem et al. 2003) (Salem and Menees, unpublished data). The *DBR1* gene blaster plasmid pTM513 was created in three steps. First, the 3.8 kbp *Bam*HI-*Bgl*II fragment from pNKY51 (Alani et al. 1987), containing *hisG-URA3-hisG*, was ligated into the *Bam*HI site of pBluescript (Invitrogen) to create pTM509. Second, the 5′ UTR of *DBR1* was amplified from genomic DNA using oligonucleotide primers 331 and 332, then the PCR product was trimmed with *Eco*RI and *Bam*HI and ligated into *Eco*RI and *Bam*HI sites of pTM509 to create pTM511. Third, the 3′ UTR of *DBR1* was amplified from genomic DNA using oligonucleotide primers 333 and 336, then the PCR product was trimmed with *Xba*I and *Not*I and ligated into *Xba*I and *Not*I sites of pTM511 to create pTM513.

When not specifically described, general molecular techniques (Ausubel et al. 2003) as well as standard yeast media and general procedures (Kaiser et al. 1994) were used. Oligonucleotides are listed in Tables 1 and 2.

RNA extraction

Yeast strains were grown to mid-logarithmic phase prior to isolating total cellular RNA. In some cases yeast cells were used directly for RNA preparation after cell growth was complete. In other cases yeast cells were pelleted and flash frozen in a dry ice ethanol bath and stored at −80°C prior to RNA preparation. We found no difference in results for RNAs prepared from cells processed in these two ways. Total yeast RNA was prepared by the hot acid phenol method (Ausubel et al. 2003) or by a column purification method (RNeasy kit, Qiagen) from small cultures (10 ml) grown to mid-logarithmic phase ($OD_{600} = \sim 1$). RNA samples were treated with RNase free DNase I (Fisher) to remove DNA contamination. RNA concentration was measured spectrophotometrically by reading OD_{260}. The OD_{260}/OD_{280} ratio was used as an RNA quality assessment.

Preparation of Dbr1p enzyme from *E. coli*

The pET16b-DBR1 expression plasmid encodes yeast Dbr1p as a N-terminal 10×-histidine-tagged protein (Martin et al. 2002). Expression and purification of the histidine-tagged Dbr1p were performed mostly as described previously (Martin et al. 2002). Rosetta DE3 *E. coli* cells (Novagen) were used for expression of Dbr1p instead of *E. coli* strain BL21-Codon Plus(DE3)RIL (Stratagene). Sonication of cells was performed on ice for 60 s, in 1 s pulses, with a large probe at 50% power. Triton X-100 was added after sonication to a final concentration of 0.1%. The tagged Dbr1p was purified from *E. coli* extracts by binding to and eluting from Nickel–nitrilotriacetic acid-agarose columns, as described previously (Martin et al. 2002), and fractions were assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Peak fractions from the elution were dialyzed against debranching buffer (Ooi et al. 2001) (20 mM HEPES KOH, pH 7.9; 125 mM KCl; 0.5 mM $MgCl_2$; 1 mM DTT; 10% glycerol). In some cases, Dbr1p was concentrated by spinning through a Microcon YM-30 spin concentrator at 14,000×*g* for 40 min at 4°C in a Beckman Allegra 25R centrifuge (TA-15-1.5 rotor). The concentrations of Dbr1p preparations were 50–100 ng/μl.

Table 1 Oligonucleotides

Primer	Sequence	Position[a]
146	cactctcccataacctccta	*ACT1* intron nt 100–119
215	ctcaaaccaagaagaaaaagaa	*ACT1* nt −128 to −107
216	tgataccttggtgtcttggtct	*ACT1* nt 130–109
331	aggatgtttccgtctttagaa	−761 to −741 upstream of *DBR1* ORF
332	gaggatcctgataaatgtctgcccatctt	−10 to −30 upstream of *DBR1* ORF; *Eco*RI site added at 5′ end
333	gctctagaacgaatgcagacggaattaga	16–30 after *DBR1* stop codon; *Xba*I site added at 5′ end
336	ataagaatgcggccgcaaagggatccaatgtggtga	779–760 after *DBR1* stop codon; *Not*I site added at 5′ end
363	gcaagcgctagaacatacttag	*ACT1* intron nt 18–1, 265–262
372	agtgaatagttcgtatccagattc	*FLO8* nt 12–35
373	catacaaaaagccttgaggtg	*FLO8* nt 418–398
374	ggtagcaaatattctgggacatct	*FLO8* nt 422–445
375	attctgggttggccctacattt	*FLO8* nt 837–816
376	agtcaaaacgttactggctgg	*FLO8* nt 841–861
377	tgcttgattgcggaagttag	*FLO8* nt 1260–1241
378	ttggcgaggaagatatttattc	*FLO8* nt 1268–1289
379	aagataatggactggatacagccg	*FLO8* nt 1675–1652
380	ttcgatccagaaagtggcaa	*FLO8* nt 1693–1712
381	ttttcctctggagtagataatgtg	*FLO8* nt 2036–2013
382	atcaaggatatgattttgacgc	*FLO8* nt 2054–2075
383	cagccttcccaattaataaaattg	*FLO8* nt 2399–2376
408	taaatagcttggcagcaacagg	*URA3* nt 67–46
417	ttgcgaattgctgtacaagg	*DBR1* nt 10–29
418	caagtcatgaatttagagataaatgc	*DBR1* nt 1217–1192
443	tgctgtcatggtcagctaaaccaaatttataaagaagtgt…	5′ 40 nt = *DBR1* nt 31–70
	…taactatgcggcatcagagc	3′ 20 nt = *URA3* flank in pRS306
444	gataaatgctttagtttgtcgtacttcatctttctgaata…	5′ 40 nt = *DBR1* nt 1200–1161
	…cctgatgcggtattttctcc	3′ 20 nt = *URA3* flank in pRS306

[a] For the *ACT1*, *FLO8*, *URA3* and *DBR1* genes, the nucleotide positions are relative to the first nucleotide of the coding sequence, except for the *ACT1* intron, where positions are relative to the first nucleotide of the intron

Mass spectrometry of purified Dbr1p was performed at the UMKC School of Biological Sciences Proteomics Facility.

Enzymatic treatments of RNA

Bacillus stearothermophilus PNPase was acquired from Sigma (St. Louis, Missouri) and a stock of 3.5 units/ml was prepared by dissolving the protein in water, then adding Tris HCl, pH 8.5, to a final concentration of 50 mM. PNPase reactions were performed in PNPase buffer (50 mM Tris HCl, pH 8.5; 1 mM 2-mercaptoethanol; 1 mM EDTA; 20 mM KCl; 1.5 mM $MgCl_2$; 10 mM Na_2HPO_4, pH 8.3) on 20–1,000 ng of total yeast RNA in 20 µl reactions for 1 h at 60°C, using 1 µl of the PNPase stock. Upon completion of reactions, samples were heated to 85°C for 10 min, then either used directly in RT-PCRs or ethanol precipitated. Mock treatments were performed in the same way, minus PNPase.

Approximately 50–100 ng of yeast Dbr1p prepared from *E. coli* was used for in vitro debranching reactions of 20–200 ng of RNA. Reactions were performed at 30°C for 45 min in a 20 µl volume containing 1× debranching buffer (20 mM HEPES–KOH pH 7.9, 125 mM KCl, 0.5 mM $MgCl_2$, 1 mM DTT and 10% glycerol). Reactions were stopped by heating at 65°C for 10 min. Mock treatments were performed in the same way, minus Dbr1p.

For sequential enzymatic treatments, RNA samples were phenol/chloroform extracted and ethanol precipitated after the first treatment (PNPase or Dbr1p) then resuspended and treated with the second enzyme.

RT-PCR methods

RT-PCRs of lariat and linear RNAs were performed with QIAGEN one-step RT-PCR kit (Valencia, CA) under the following general conditions: 50°C, 30 min; 95°C, 15 min;

Table 2 Primers and probes for qRT-PCR

Target and primers/probe	Sequence	Position[a]
ACT1 mRNA		
FWD primer	TCCCAAGATCGAAAATTTACTGAAT	−30 to −6
REV primer	TTTACACATACCAGAACCGTTATCAAT	54 to 28
TaqMan probe	VIC-TGAATTAACAAGGTTGCTGCT-MGBNFQ	−4 to –26
ACT1 intron		
FWD primer	ATTTTTCACTCTCCCATAACCTCCTATA	94 to 121
REV primer	TTTCAAGCCCCTATTTATTCCAAT	173 to 150
TaqMan probe	6FAM-TGACTGATCTGTAATAACCA-MGBNFQ	123 to 142
RPP1B mRNA		
FWD primer	AGGCCGCTGGTGCTAATG	89 to 106
REV primer	TCCAAAGCCTTAGCGTAAACATC	146 to 124
TaqMan probe	VIC-CGACAACGTCTGGGC-MGBNFQ	108 to 122
RPP1B intron		
FWD primer	AATGCAACCTAAAACGACTTTGTG	12 to 35
REV primer	TTTCTCGGGACGATTGTTGTC	77 to 57
TaqMan probe	6FAM-ACTACGAAGAGAAAGATT-MGBNFQ	38 to 55
YRA1 mRNA		
FWD primer	AGGTTTGCCAAGGGACATTAAG	249 to 270
REV primer	ACACCACCTACTTGAGATGCAAAA	314 to 291
TaqMan probe	VIC-AGGATGCTGTAAGAGAAT-MGBNFQ	272 to 289
YRA1 intron		
FWD primer	CGCATCGTCTCGTGTGGAT	42 to 60
REV primer	GATCAAAAGCGTGTGCCATATC	107 to 86
TaqMan probe	6FAM-CGAGAAATATTCTTTGTAAGGAA-MGBNFQ	62 to 84

The primers and probe for human GAPDH qPCR were designed by Applied Biosystems and obtained from the company

[a] Relative to start of coding sequence for mRNA primers and probes. Relative to start of intron sequence for intron primers and probes

nine cycles of 94°C for 30 s, 54°C for 30–60 s [touchdown to 46°C (−1°C per cycle)], 72°C for 30 s; X cycles (see below) of 94°C for 30 s, 46°C for 30 s, 72°C for 30–45 s; 72°C for 5–10 min; 4°C hold. The number of cycles in the post-touchdown phase of different RT-PCRs (X cycles above) varied with the experiment and are reflected in the following reaction profile names: ACT1-1, 29 cycles, post-touchdown; ACT1-2, 24 cycles, post-touchdown; ACT1-3, 19 cycles, post-touchdown; ACT1-4, 15 cycles, post-touchdown; and ACT1-5, 11 cycles, post-touchdown. RNA amounts between 2 and 50 ng were used in RT-PCRs. RT-PCRs were analyzed by either PAGE or agarose gel electrophoresis.

Real-time RT-PCR (qRT-PCR) of lariat and linear RNAs

Primers and probes for qPCR were designed using Sequence Detection Systems software from Applied Biosystems and are listed in Table 2. All probes and primers for qRT-PCR were purchased from Applied Biosystems. Validation experiments were performed that demonstrated that the efficiencies of target and reference PCRs were approximately equal.

For total RNA samples (untreated or treated with Dbr1p/PNPase, as described above), 20–1,000 ng of RNA was reverse transcribed into cDNA using random hexamers in a 100 µl reaction at 45°C for 60 min.

PCR MasterMix reagents from Applied Biosystems were used for qPCR reactions, which were performed in triplicate for each sample. Reactions were prepared and run according to a standard protocol established by Applied Biosystems on an ABI 7500 real-time PCR machine. Briefly, reactions contained 2× PCR MasterMix, 900 nM forward primer, 900 nM reverse primer, 250 nM TaqMan probe, and cDNA (~20 ng). Reactions were incubated for 2 min at 50°C and then 10 min at 95°C and before proceeding through 40 cycles of a 30 s incubation at 95°C and a 60 s incubation at 60°C. Completed reactions were held at 4°C.

Relative quantification (RQ) of results was performed using the comparative C_T method ($\Delta\Delta C_T$) (Schmittgen and Livak 2008). The amplification of each target intron sequence was compared to amplification of the corresponding mRNA sequence and a ΔC_T was determined. To

compare the different samples to each other, the wild-type sample was used as the calibrator sample. Therefore, the ΔC_T of the wild-type sample was subtracted from the ΔC_T for each sample to determine $-\Delta\Delta C_T$ values. In Fig. 7, $RQ = 2^{-\Delta\Delta C_T}$ for each $-\Delta\Delta C_T$ and represents the fold-difference in intron levels between a given sample and the wild-type sample (*DBR1*).

In vitro debranching time course

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) cDNA, the exogenous control for qPCR in these experiments, was generated by reverse transcribing 600 ng of human RNA at 45°C for 1 h using the RT kit from Applied Biosystems. A debranching reaction mix was set up on ice and contained 5,600 ng of total RNA from TMY60 (*dbr1*) cells, ~6 ng GAPDH cDNA, 140 µl of purified Dbr1p, and 350 µl 2× debranching buffer in a final volume of 700 µl. Seven 100 µl aliquots of this mix were distributed to 0.2 ml PCR tubes. The debranching reaction was directly inactivated in one tube (0 min reaction time) by raising the temperature to 95°C, followed by phenol/chloroform extraction and ethanol precipitation. The remaining six tubes were incubated at 30°C and individual reactions were stopped after 2.5, 5, 10, 15, 30, and 60 min. Reactions were stopped by raising the temperature to 95°C, followed by phenol/chloroform extraction and ethanol precipitation. RNAs were then treated with PNPase, as described above, to degrade intron lariats linearized by Dbr1p. Reverse transcription of the RNAs remaining from the different debranching reactions was performed using the RT kit from Applied Biosystems and random hexamer primers. qPCRs using these cDNAs were performed as described above, amplifying a volume of cDNA roughly corresponding to ~20 ng of starting total RNA, using primers and probes for yeast *ACT1*, *YRA1*, and *RPP1B* introns as well as human GAPDH. GAPDH cDNA was the exogenous control because it is insensitive to PNPase and remained at a constant level in each reaction.

Creation of *dbr1* point mutant strains

Mutants were created using modifications of the *delitto perfetto* method (Storici et al. 2001) and the site specific genomic (SSG) method (Gray et al. 2004). Initially, a *dbr1Δ::URA3* strain was created to facilitate the introduction of point mutant alleles of *dbr1* into the *DBR1* locus. Yeast strain TMY490, containing a *URA3*-marked deletion of 1,090 bp of the 1,215 bp *DBR1* coding sequence (nts 71–1,160 deleted), was constructed by transformation of TMY30 with a PCR fragment containing the *URA3* gene from pRS306 flanked by ends corresponding to 5′ and 3′ segments of the *DBR1* coding region.

The fragment used for making the *dbr1Δ::URA3* allele was created by PCR of pRS306 with oligonucleotides 443 and 444, the 3′ 20 nt of which anneal to the ends of the *URA3* gene on pRS306 and the 5′ 40 nt of which correspond to *DBR1* sequences (see Table 1).

The *dbr1Δ::URA3* disruption on yeast chromosome XI was created by homologous recombination between the *DBR1* locus and the *dbr1Δ::URA3* PCR fragment. Briefly, TMY30 was transformed with the *dbr1Δ::URA3* PCR fragment and transformants were selected on SD-Uracil plates. Transformants were screened by PCR with primer pairs 401/402, which anneal within the *DBR1* sequences that are deleted in the *dbr1Δ::URA3* allele, and 417/418, which anneal outside the *DBR1* sequences that are deleted in the *dbr1Δ::URA3* allele. Transformants containing the *dbr1Δ::URA3* allele template a 417/418 PCR product but not a 401/402 PCR product. DNA sequencing of PCR products was performed to verify the presence of the *dbr1Δ::URA3* allele.

Replacement of the chromosomal *dbr1Δ::URA3* allele with *dbr1* point mutations was accomplished by transformation. TMY490 (*dbr1Δ::URA3* strain) was co-transformed with YEp351(*LEU2*) and PCR fragments of *dbr1* point mutants. The PCR fragments were generated from plasmids pTM431, pTM432, and pTM435 with PCR primer pairs 417/418. Transformants (with YEp351) were selected in SD-leucine liquid media during a 48 h incubation period at 30°C (with shaking). After this selection period, cells were spread onto 5-fluoroorotic acid plates to select for cells that lost function of the *URA3* gene within the *DBR1* locus. Recombinants within the FOA^r population that have replaced the *dbr1Δ::URA3* allele with a *dbr1* point mutant allele were identified by PCR screening. Positive clones were identified as those that template a 417/418 PCR product but not a 417/408 PCR product (specific for the *dbr1Δ::URA3* allele). DNA sequencing of PCR products was performed to verify the presence of a *dbr1* point mutant allele.

Results

RT-PCR detection of lariat RNAs

S. cerevisiae ACT1, which encodes actin, is a robustly expressed gene that contains an intron of 308 nt. The first example of a spliceosomal intron discovered in yeast (Gallwitz and Sures 1980; Ng and Abelson 1980), the *ACT1* intron contains all the canonical features of yeast introns and is efficiently spliced from pre-mRNA, producing an excised lariat with a 265 nt circle. We chose this

well-characterized gene to assess intron levels as we developed and tested tools for detecting and enriching excised intron lariats. Primers were designed for use in RT-PCR to detect the lariat form of the *ACT1* intron RNA and, as a control, *ACT1* mRNA (Fig. 1a). RT-PCR of total yeast RNA using primers that flank the *ACT1* exon–exon junction (primers 215 and 216) amplifies a 285 bp product from *ACT1* mRNA. Primer 363 spans the *ACT1* intron lariat branch point and is used in combination with primer 146, which anneals to sequences complementary to the *ACT1* intron upstream of the lariat branch point, in an RT-PCR that amplifies a 184 bp product from the lariat form of the *ACT1* intron RNA. As expected, when RT-PCRs are performed using total RNA samples from wild-type (TMY30) and *dbr1* mutant yeast cells (TMY60), the amounts of *ACT1* mRNA products are similar when using equivalent amounts of RNA from the two cell types (Fig. 1b, lanes 1 and 3 as well as lanes 5 and 7). However, the *ACT1* intron RNA lariat product is much more readily produced from *dbr1* cells, also as expected (Fig. 1b, lane 4 vs. 2 and lane 8 vs. 6). These data clearly show that a *dbr1* mutant strain or, where appropriate, a Dbr1p knock-down strain contains a rich source of expressed intron sequences. It is also evident that the use of intron-specific RT-PCR could be used to detect excised introns from genes expressed at very low levels. For studies on alternative splicing, the use of RT-PCR on RNA from Dbr1p-deficient cells can allow detection of rare splice variants, an approach that has already been exploited (Conklin et al. 2005).

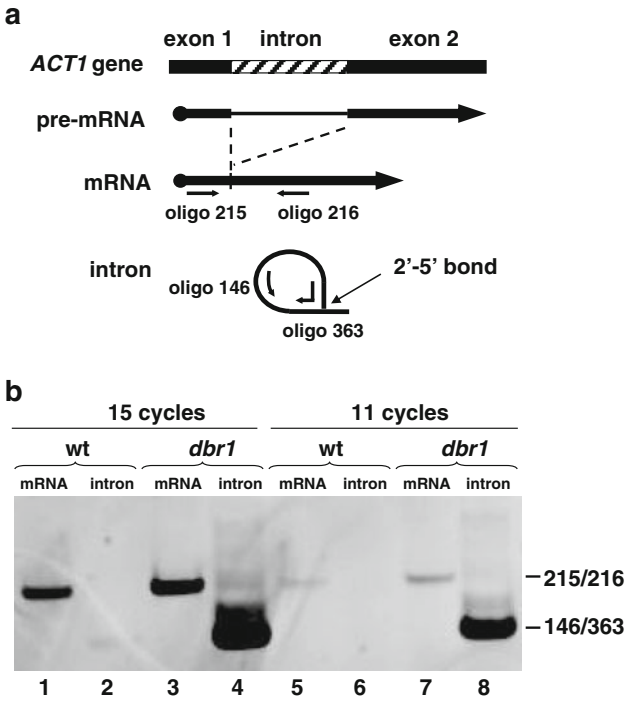

Fig. 1 RT-PCR detection of lariat RNAs. **a** Annealing positions of primers for RT-PCR detection of *ACT1* intron lariat RNA and mRNA. **b** Agarose gel analysis of RT-PCRs for *ACT1* RNAs. *Lanes 1–4* contain reactions run 15 cycles after the touchdown phase of PCR; *lanes 5–8* contain reactions run 11 cycles after the touchdown phase of PCR. The different numbers of cycles were run to show the linearity of the PCRs. Primer pairs for PCRs are indicated on *right*. Reactions using wild type (*DBR1*) RNA samples are in *lanes 1*, *2*, *5*, and *6*; reactions using *dbr1* mutant RNA samples are in *lanes 3*, *4*, *7*, and *8*

A previous report described the use of radiolabeled primers spanning intron RNA branch points for analyzing intron populations by primer extension (Spingola et al. 1999). The RT-PCR method we describe could be modified to survey intron lariats containing specific sequences at intron 5′ ends and branch points, as described previously (Spingola et al. 1999). RT-PCR has added utility because the products can be cloned and sequenced to identify the individual introns represented in a lariat population.

Insensitivity of lariat RNAs to the 3′ exonuclease PNPase

Linear and lariat RNAs have different sensitivities to 3′ exonucleases, including PNPase, a component of bacterial RNA degradation systems (Suzuki et al. 2006). PNPase degrades linear RNAs but does not proceed past the 2′ branch present in intron RNA lariats (Suzuki et al. 2006). Therefore, treatment of RNA samples with PNPase should result in a vast enrichment of excised intron lariats in the RNA that remains intact after treatment. This difference should be evident in the results of the RT-PCR assay described above when amplifying PNPase-treated RNA samples. Since RNA secondary structures reduce the efficiency of PNPases (Guarneros and Portier 1990; McLaren et al. 1991), reactions were performed at elevated temperature (60°C) using PNPase from *B. strearothermophilus* to circumvent this problem. Total RNA samples from a *dbr1* mutant strain (TMY60) were treated with a range of PNPase concentrations and then subjected to RT-PCR to detect *ACT1* intron RNA lariats as well as the linear mRNA (Fig. 2). Results are consistent with expectations, as observed previously (Suzuki et al. 2006) the use of PNPase selectively preserves RNA lariats.

The high-temperature reaction using PNPase from a thermophile appears to be much more efficient than the reported reaction with the *E. coli* PNPase at 37°C (Suzuki et al. 2006). To eliminate the RT-PCR product from the *ACT1* mRNA, PNPase must degrade, at the very least, the RNA corresponding to the binding site for the downstream primer (oligonucleotide 216). To accomplish this, PNPase must degrade all the RNA that lies to the 3′ side of the

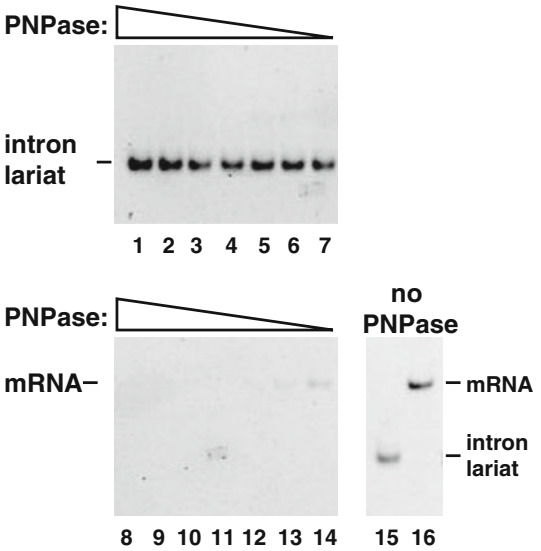

Fig. 2 PNPase degrades linear RNAs but not lariat RNAs. *Upper panel* (*lanes 1–7*). Agarose gel analysis of RT-PCRs for *ACT1* intron lariat RNA following a series of PNPase treatments. PNPase reactions were performed with diminishing concentrations of enzyme as indicated by the wedge at the top of the gel image (RT-PCR of highest concentration PNPase reaction is in *lane 1*; RT-PCR of lowest concentration PNPase reaction is in *lane 7*). RT-PCRs for *ACT1* intron lariat RNA were performed with primers 146 and 363 (see Fig. 1) and run for 15 cycles after the touchdown phase of the reaction. *Lower panel* (lanes *8–14*). Agarose gel analysis of RT-PCRs for *ACT1* mRNA (linear) from the same series of PNPase treatments described above. PNPase reactions were performed with diminishing concentrations of enzyme as described for *upper panel*. RT-PCRs for *ACT1* mRNA were performed with primers 215 and 216 (see Fig. 1) and run for 24 cycles after the touchdown phase of the reaction. *Lanes 15* and *16* contain *ACT1* intron lariat and *ACT1* mRNA RT-PCRs, respectively, of RNA samples that did not undergo PNPase treatment. The products in these lanes serve as size markers for the intron lariat and mRNA products in *lanes 1–14*

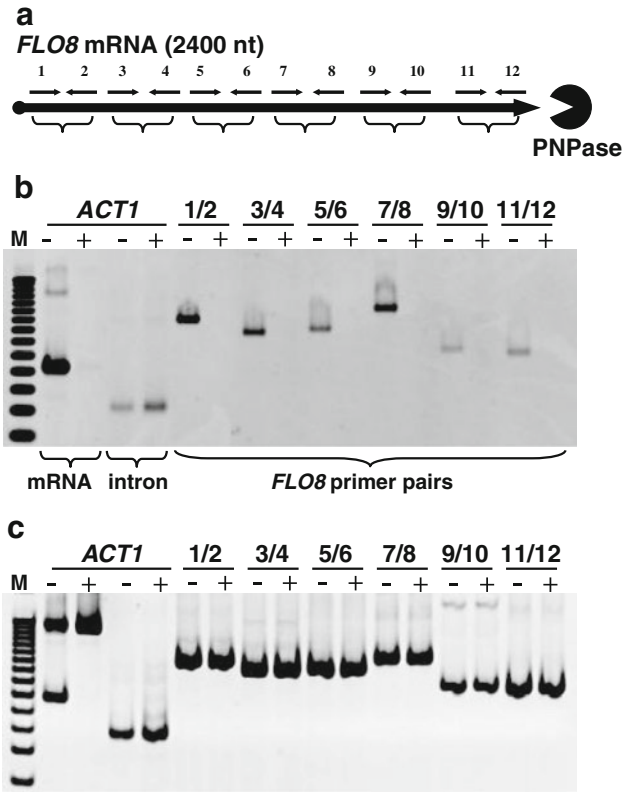

Fig. 3 Processivity of PNPase on *FLO8* mRNA. **a** Primer pairs for amplifying different segments along the length of *FLO8* mRNA: 1/2 = primers 372 and 373; 3/4 = primers 374 and 375; 5/6 = primers 376 and 377; 7/8 = primers 378 and 379; 9/10 = primers 380 and 381; 11/12 = primers 382 and 383. Primers are listed in Table 1. **b** PAGE analysis of RT-PCRs for *FLO8* mRNA segments following PNPase treatment (+ *lanes*) and mock treatment (− *lanes*) of a total cellular RNA sample that had been pretreated with DNase I. Lanes containing the various *FLO8* RT-PCRs are indicated *below* the gel image; the *FLO8* primer pairs are indicated *above* the gel image. RT-PCRs for *ACT1* RNAs are in the *four lanes* under the *ACT1* title and serve as controls that indicate the PNPase reactions proceeded as expected. The RT-PCRs for *ACT1* mRNA and intron RNA are indicated *below* the corresponding lanes; these reactions used primer pairs 215/216 and 146/363, respectively. The lane marked "*M*" contains a DNA molecular weight standard (50 bp ladder). **c** PAGE analysis of RT-PCRs as described for **b** except that total cellular nucleic acid samples were not treated with DNase I prior to PNPase treatment and RT-PCRs. For all RT-PCRs in **b** and **c**, reactions were performed with 24 cycles after the touchdown phase

oligonucleotide 216 binding site, which includes 998 nt of the *ACT1* coding sequence plus the 3′ UTR and the polyA tail. To further examine the processivity of *B. strearothermophilus* PNPase, the degradation of *FLO8* mRNA was assessed. *FLO8* mRNA is >2.4 kb in length. Primer pairs were designed to amplify different portions of this mRNA along its length (Fig. 3a). Total nucleic acid samples and RNA samples (DNased total nucleic acid samples) were treated with PNPase and subjected to RT-PCR to detect the various segments of *FLO8*. As shown in Fig. 3b, PNPase readily degrades every segment of *FLO8* mRNA assayed. As expected, PNPase has no effect on *FLO8* DNA present in the total nucleic acid samples (Fig. 3c). Other enzymes reported to work as well as *B. strearothermophilus* PNPase in our studies are *Thermus thermophilus* PNPase at 65°C (Falaleeva et al. 2008) and *E. coli* RNase R at 37°C (Suzuki et al. 2006; Vincent and Deutscher 2006), although direct comparisons have not been made.

Sensitivity of lariat RNAs to Dbr1p

Linear and lariat RNAs also have different sensitivities to RNA debranching enzyme, which can be exploited to confirm that an RNA species has a lariat conformation. The RT-PCR strategy employing a primer that spans a lariat branch point, as described above for the *ACT1* intron, can be used to demonstrate the cleavage of the 2′–5′ bond. This is due to the fact that after Dbr1p treatment the binding site for the primer that spans the *ACT1* intron branch point (oligonucleotide 363) is split into two non-contiguous

sections, with the section that anneals to the 3′ end of the primer being only 3 bp in length. After debranching of the lariat, the critical 3′ end of the primer will not effectively anneal to the intron RNA to prime RT-PCR. Dbr1p treatment has no effect on *ACT1* mRNA, which should still be readily detected by RT-PCR.

To perform Dbr1p treatments, *S. cerevisiae* Dbr1p was expressed in *E. coli* and purified by metal affinity chromatography (Fig. 4a). Although histidine-tagged Dbr1p is expected to have a mass of ~50 kDa, the mobility of the main product in SDS-PAGE is ~45 kDa. Others have observed this anomalous mobility for histidine-tagged Dbr1p (Ooi et al. 2001) and have speculated that the protein may undergo limited proteolysis in *E. coli*. However, mass spectrometric analysis of the main band in the stained gel shows it to be the expected molecular mass of the histidine-tagged Dbr1p (50,062 Da) (Fig. 4b), indicating that the protein is intact and must run anomalously in SDS-PAGE because of its physical properties.

Using the Dbr1p enzyme preparation, debranching reactions were carried out on total RNA samples from a *dbr1* mutant strain. RT-PCR analysis reflects the differential sensitivity of linear and lariat RNAs to Dbr1p. After Dbr1p treatment, RT-PCR detection of *ACT1* RNA lariat is greatly decreased (Fig. 5, lane 4 vs. 2). On the other hand, the product indicative of *ACT1* linear mRNA is still readily detectable after Dbr1p treatment (Fig. 5, lane 3 vs. 1).

Combinations of PNPase and Dbr1p treatments

PNPase and Dbr1p treatments can be used in combination when exploring the properties of a particular RNA species. Sequential enzymatic treatments can also be used to enrich for RNA lariats and then linearize them for further manipulations. To demonstrate this, *ACT1* RNA species present within a total RNA sample from a *dbr1* mutant strain were analyzed by RT-PCR following sequential PNPase and Dbr1p treatments. As shown in Fig. 6a, lanes 1–4, initial treatment of the RNA sample with PNPase degrades the linear mRNA (lanes 1 and 3), but leaves lariat RNA intact (lane 2). Subsequent treatment with Dbr1p shows that the resistant RNA is a lariat (lane 4). As shown

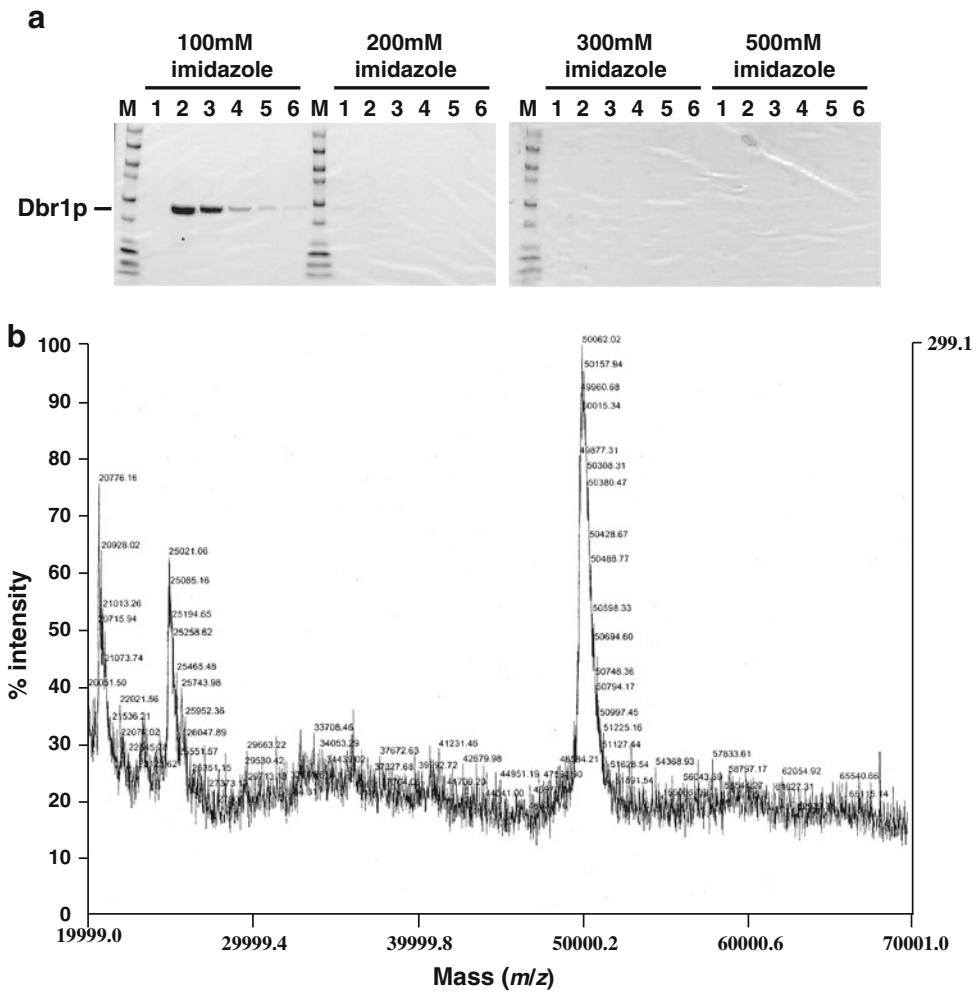

Fig. 4 Purification of Dbr1p. **a** Elution profile of histidine-tagged yeast Dbr1p purified from *E. coli*. Dbr1p bound to a nickel–nitrilotriacetic acid (nickel–NTA) column was eluted with increasing concentrations of imidazole (indicated at *top*). Six, ~1.5 ml fractions were collected for each imidazole concentration. Fraction numbers are listed *above each lane*. Protein molecular weight standards (Precision Plus Protein Kaleidoscope standards, Bio-Rad) were run in lanes indicated by "*M*." Dbr1p runs between the 50 and the 37 kDa standards. **b** Matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) mass spectrometry analysis to assess the molecular mass of the main elution product in fractions 2–6 of the 100 mM imidazole elution

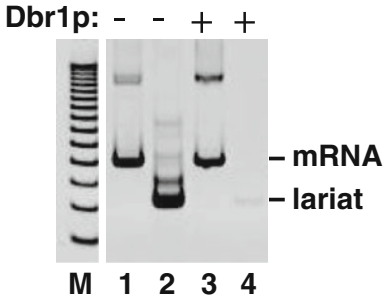

Fig. 5 In vitro debranching reaction. Agarose gel analysis of RT-PCRs for *ACT1* RNAs following treatment with Dbr1p (+ *lanes*) and mock treatment (− *lanes*) of a total cellular RNA sample. *Lanes 2* and *4* contain RT-PCRs for *ACT1* intron lariat RNA; *lanes 1* and *3* are RT-PCRs for *ACT1* mRNA. The lane marked "*M*" contains a DNA molecular weight standard (50 bp ladder). RT-PCRs for *ACT1* intron lariat RNA were run for 19 cycles after the touchdown phase of the reaction; RT-PCRs for *ACT1* mRNA were run for 24 cycles after the touchdown phase of the reaction

in Fig. 6a, lanes 5–8, skipping the initial PNPase treatment leaves the linear mRNA intact (lanes 5 and 7) as well as the lariat RNA (lane 6). The lariat RNA is then distinguished by its sensitivity to cutting with Dbr1p (lane 8). The order of the PNPase and Dbr1p reactions can be switched to generate a complementary set of predictable results (Fig. 6b).

Real-time RT-PCR measurement of lariat RNA levels

A real-time RT-PCR method (qRT-PCR), using the TaqMan detection system (Applied Biosystems), was developed to quantitatively compare the intron RNA lariat levels of different samples. We expanded our study to include not only the *ACT1* intron but also the *YRA1* and *RPP1B* introns to investigate the generality of the methods. *YRA1* encodes an RNA binding protein involved in mRNA export from the nucleus (http://www.yeastgenome.org) and is moderately expressed, although less than *ACT1* (Holstege et al. 1998) (2005 update at http://www.web.wi.mit.edu/young/pub/holstege.html). The *YRA1* intron is 765 nt in length, which is larger than the 300 nt average for yeast introns, and contains a non-canonical branch point sequence (http://www.compbio.soe.ucsc.edu/yeast_introns.html; http://www.embl.de/ExternalInfo/seraphin/yidb.html) (Lopez and Seraphin 2000). Furthermore, the intron is inefficiently spliced from pre-mRNA, which is important for the autoregulation of Yra1p protein levels (Preker et al. 2002; Rodriguez-Navarro et al. 2002; Preker and Guthrie 2006). *RPP1B* encodes a ribosomal protein and is even more highly expressed than *ACT1* (Holstege et al. 1998) (2005 update at http://www.web.wi.mit.edu/young/pub/holstege.html). The *RPP1B* intron is typical for yeast, 301 nt in length, with canonical sequences (http://www.compbio.soe.ucsc.edu/yeast_introns.html; http://www.embl.de/ExternalInfo/seraphin/yidb.html) (Lopez and Seraphin 2000).

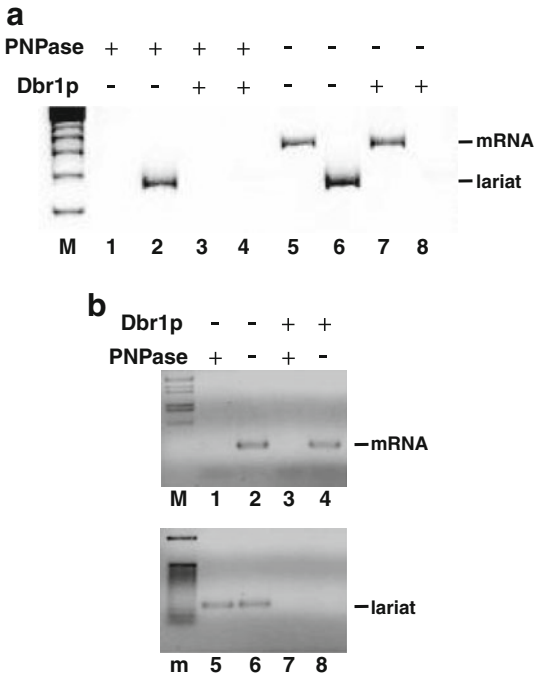

Fig. 6 Combinations of PNPase and Dbr1p treatments. **a** Agarose gel analysis of RT-PCRs for *ACT1* RNAs following treatment of a total cellular RNA sample from a *dbr1* strain with Dbr1p (+ Dbr1p) and PNPase (+ PNPase) as well as mock treatment (− treatment). In this experiment, PNPase treatment preceded Dbr1p treatment for samples that were treated with both enzymes. *Lanes 1*, *3*, *5*, and *7* contain RT-PCRs for *ACT1* mRNA of a total cellular RNA sample. *Lanes 2*, *4*, *6*, and *8* contain parallel RT-PCRs for *ACT1* intron lariat RNA. **b** Agarose gel analysis of RT-PCRs for *ACT1* RNAs following treatment of a total cellular RNA sample from a *dbr1* strain with Dbr1p and PNPase as well as mock treatment. In this experiment, Dbr1p treatment preceded PNPase treatment for samples that were treated with both enzymes. *Lanes 1–4* contain RT-PCRs for *ACT1* mRNA of a total cellular RNA sample. *Lanes 5–8* contain parallel RT-PCRs for *ACT1* intron lariat RNA. For both **a** and **b**, RT-PCRs for *ACT1* intron lariat RNA were run for 19 cycles after the touchdown phase of the reaction and RT-PCRs for *ACT1* mRNA were run for 24 cycles after the touchdown phase of the reaction. The lanes marked "*M*" and "*m*" contain DNA molecular weight standards ("*M*" = λ phage DNA cut with *Hin*DIII + *Eco*RI; "*m*" = 50 bp ladder)

Initially we used a strategy like the one used for RT-PCR of *ACT1* intron lariats described above, with one primer spanning the lariat branch point and serving as both the RT primer and the reverse primer for PCR. However, we switched to using random primers for the RT step to allow amplification of the different target sequences from a common pool of cDNA. Consequently, both PCR primers anneal upstream of the branch point for each target gene, with a TaqMan probe annealing between them (see Fig. 7a). Since these types of primers will also prime amplification of genomic DNA we ran control PCRs for each sample without a prior RT step to ensure that DNA

contamination was not contributing to the PCR product. The mRNA for each target gene served as the endogenous control for qRT-PCR (Fig. 7a, top). Using this strategy, intron sequences for *ACT1*, *RPP1B*, and *YRA1* were amplified from *dbr1* and wild-type yeast strains (TMY60 and TMY30). As shown in Fig. 7b–d [*DBR1* (wild type) vs. *dbr1* null mutant], the real-time method generated the expected results: the different intron RNAs accumulate at higher levels in the *dbr1* null mutant strain than in wild type.

qRT-PCR was also used to analyze mutant variants of Dbr1p. Previously, a set of point mutants had been created by random PCR mutagenesis and analyzed for intron RNA levels by an RNase protection assay (Salem et al. 2003) (Salem and Menees, unpublished data). In those experiments the *dbr1* mutant alleles were under the control of a strong, inducible promoter (p*GAL1*) and carried on a high copy plasmid. The yeast strain carried a *dbr1*Δ mutation [open reading frame (ORF) deletion] at the *DBR1* locus so the plasmid-borne *dbr1* mutant alleles were the only sources of Dbr1p. For the current study, three *dbr1* point mutants (D180Y, G84A, and Y68S) were analyzed by qRT-PCR to determine their levels of intron lariat RNA relative to wild-type (*DBR1*) and *dbr1*Δ. To make the analysis more biologically relevant, each of the *dbr1* mutant alleles was placed at the *DBR1* locus, replacing the

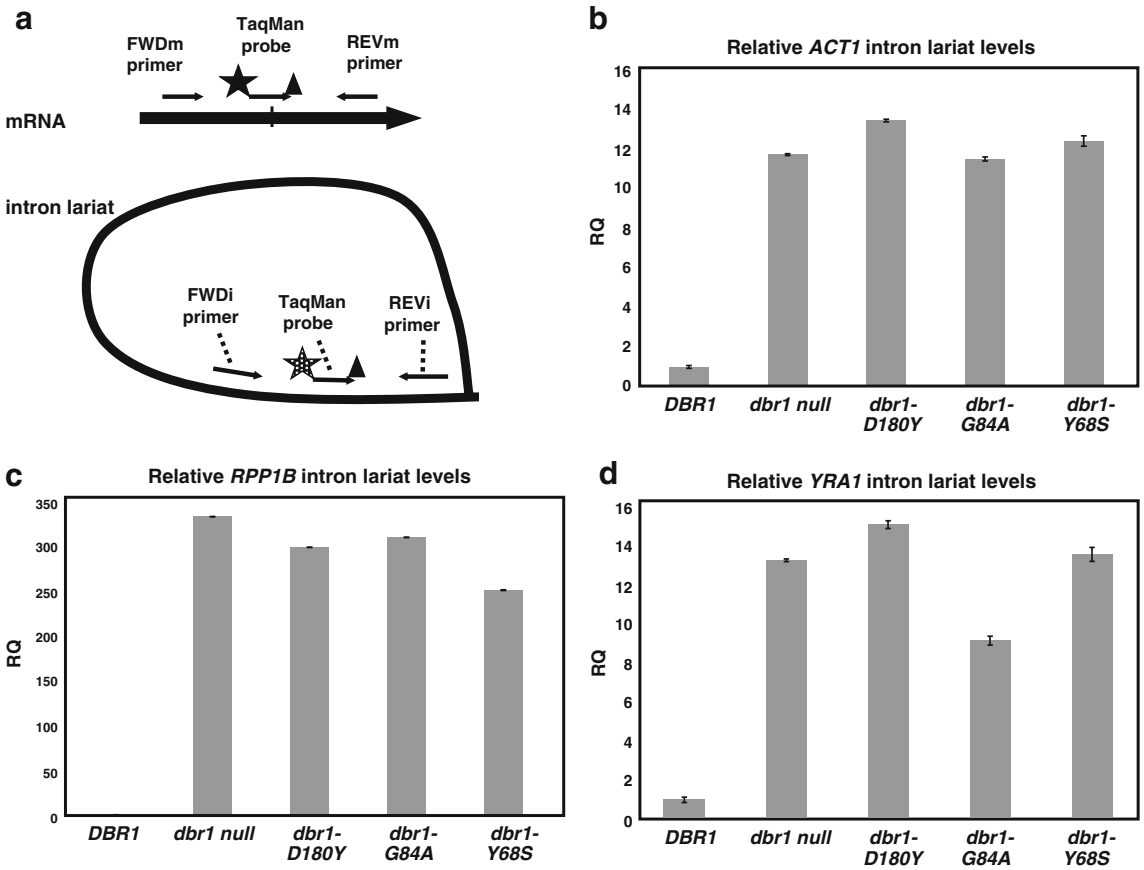

Fig. 7 Real-time RT-PCR (qRT-PCR) measurement of lariat RNA levels. **a** Annealing positions of primers for RT-PCR detection of mRNA and intron lariat RNA species. FWDm and REVm primers are designed to amplify the segment of mRNA spanning an exon–exon junction (indicated by *vertical line*). A TaqMan probe is designed to span the same exon–exon junction. The *star* and the *triangle* at opposite ends of the TaqMan probes represent the fluorescent reporter molecule and the quencher that are bound to the 5′ and 3′ ends, respectively. The TaqMan probes that anneal to a particular mRNA and lariat RNA pair contain different fluorescent reporter molecules, indicated by *solid* and *stippled stars*. Note that lariat RNA detection does not involve annealing of PCR primers or TaqMan probes across lariat branch points. **b** Relative quantification of *ACT1* intron lariat RNA in total RNA samples from different yeast strains. The allele at the *DBR1* locus for each strain is indicated *beneath* the *bar graph*. *RQ*, the relative quantification, is the ratio of intron RNA to mRNA for a particular sample relative to the ratio of intron RNA to mRNA for the *DBR1* (wild-type) sample at the *left end* of the *bar graph* (which sets the RQ for *DBR1* itself to 1). Quantification experiments were repeated three times and the qPCRs were performed in triplicate each time. The *standard error bars* display the calculated maximum (RQMax) and minimum (RQMin) expression levels that represent standard error of the mean expression level (RQ value), as described in the Applied Biosystems 7300/7500 product literature. **c** Relative quantification of *RPP1B* intron lariat RNA for the same RNA samples presented in **b**. **d** Relative quantification of *YRA1* intron lariat RNA for the same RNA samples presented in **b**

wild-type allele, and was under the control of the native *DBR1* promoter. After log-phase growth of cells, RNA samples from wild-type and mutant strains were harvested and subjected to qRT-PCR to amplify intron and messenger RNA sequences from *ACT1*, *RPP1B*, and *YRA1*. The three *dbr1* alleles tested show strong intron RNA accumulation phenotypes, comparable to the *dbr1*Δ knockout allele (Fig. 7b–d), which fits with information from previous experiments (Salem et al. 2003; Khalid et al. 2005) (Salem and Menees, unpublished data).

qRT-PCR analysis of a debranching time course

Using a combination of Dbr1p and PNPase treatments, in vitro debranching reactions of total cellular RNA from a *dbr1* strain were followed over time courses of 30 min. Debranching reactions were stopped at different times and the reaction products were treated with PNPase to degrade linearized intron RNAs. The remaining intron lariats were detected by qRT-PCR as described above. Because the PNPase treatment step degrades all linear RNAs, human GAPDH cDNA was added to the yeast RNA samples as an exogenous control. The GAPDH cDNA is insensitive to both Dbr1p and PNPase, remaining at the same level in the various samples (data not shown). Debranching of the *ACT1* and *RPP1B* intron lariats was almost complete within the first 5 min of the reactions (Figure S1). However, the debranching rate of the *ACT1* intron lariat appears to be only two-thirds the initial rate of the *RPP1B* intron lariat.

Discussion

The results reported here describe the development and application of (1) RT-PCR methods for sensitively and specifically detecting and quantifying lariat RNA species and (2) enzymatic treatments with characteristic, distinguishing effects on lariat and linear RNAs. We have explored some applications of these methods, such as testing the activity of Dbr1p mutants and following debranching reactions over time courses.

Our results using qRT-PCR to follow in vitro debranching reactions match well with the results reported previously using pure, synthetic substrates (Nam et al. 1994; Khalid et al. 2005). Nam et al. (1994) performed an in vitro debranching time course of a ^{32}P-labeled lariat RNA purified from an in vitro splicing reaction, using an extract from *E. coli* cells expressing yeast Dbr1p. Khalid et al. (2005) performed in vitro debranching time courses with a ^{32}P-labeled synthetic branched substrate and yeast Dbr1p purified from an *E. coli* expression system, using the same plasmid that we used in the current study. Khalid et al. (2005) reported that Dbr1p acted very quickly in reactions, with 79% of the branched substrate converted to linear product within 13 s (the first time point of the experiment). Our reactions, using actual yeast intron lariat RNAs within a total RNA sample, show that the debranching rates can vary from one intron lariat to another. The *ACT1* intron is debranched at only two-thirds the initial rate at which the *RPP1B* intron lariat is debranched. These data suggest the possibility that different intron lariats are debranched at different rates in vivo, which may be of functional significance. Slower rates of debranching may occur for introns that contain snoRNAs or mirtrons, reflecting the binding of additional factors to intron sequences or specific folding properties of the RNA. If so, the rate of debranching of introns could be used to predict which introns may contain additional information. Relative debranching rates can be inferred from quantitative analysis of intron RNA levels relative to mature mRNA levels for a given gene compared to a standard, rapidly debranching intron RNA. For these types of experiments RNA samples would have to be taken from a wild-type strain (*DBR1*), where lariat RNAs are not stabilized. Inefficient splicing would have to be ruled out before further study of candidate slow debranchers. As described above, *YRA1* is an example of a gene that uses splicing inefficiency to regulate protein levels.

We hypothesize that qRT-PCR of lariat RNAs can be used to determine the relative rates of transcription for different intron-containing genes, although there are many caveats. Currently, nuclear run-on assays are considered to be the best method for estimating a gene's transcription rate (Garcia-Martinez et al. 2004; Smale 2009). Rate estimates have also been calculated from combining information on an mRNA's steady state level and its half-life (Holstege et al. 1998). A new method, called dynamic transcriptome analysis (DTA) (Miller et al. 2011) uses metabolic labeling of cellular RNAs to measure RNA synthesis and decay rates. All these methods have different strengths and weaknesses. These differences are highlighted by the fact that there are differences in the relative transcription rates they predict (Table 3).

The use of intron RNA lariats as a novel data source for estimating relative levels of transcription for pre-mRNAs limits the utility to intron-containing genes, a notable limitation for *S. cerevisiae*. Furthermore, a Dbr1p-deficient strain would have to be used for introns lariats to be a stable record of transcription. Work with yeast *dbr1* mutants over the years has not found any significant perturbation of cellular physiology other than the accumulation of intron RNA lariats. In the experiments reported in Fig. 7b–d, the level of *RPP1B* intron RNA in a *dbr1*Δ strain relative to the level in wild type is much greater (~330-fold) than the corresponding levels of *ACT1* and *YRA1* intron RNAs (~13-fold). These data indicate that

Table 3 *ACT1*, *YRA1*, and *RPP1B* mRNA expression

Gene	Transcriptional frequency[a]		DTA[b]	Relative intron levels[c]
ACT1 (YFL039C)	45.5[d] (1)	7.2[e] (1)	63.2 (1)	1.0
YRA1 (YDR381W)	16.2[d] (0.4)	80.6[e] (11.2)	88.9 (1.4)	1.1
RPP1B (YDL130W)	120.0[d] (2.6)	23.0[e] (3.2)	192.7 (3)	28.1

[a] mRNAs/cell/hr; numbers in parentheses are levels normalized to *ACT1* level

[b] DTA = dynamic transcriptome analysis (Miller et al. 2011), measured as mRNAs/cell/cell cycle time (150 min); numbers in parentheses are levels normalized to *ACT1* level

[c] Derived from data in Fig. 7 for the *dbr1* null strain versus wild type for each gene and normalized to *ACT1* level

[d] Estimated from RNA expression levels and mRNA half-lives (Holstege et al. 1998)

[e] Estimated from genomic run on experiments (Pelechano and Perez-Ortin 2010)

the transcription rate for *RPP1B* is almost 30-fold greater than the rates for *ACT1* and *YRA1* (summarized in Table 3). These relative transcription rates are very different from estimates based on nuclear run-on assays, mRNA steady state levels plus half-lives, and DTA (Table 3). Therefore, development of new methods will provide additional insights into the issues related to determining transcription rates and contribute to the development of ideas about this important aspect of gene expression.

An area where the utility of excised introns is clearer is in determining relative rates of alternative splicing for a particular gene. Variable stabilities of different mRNAs confound estimates of their rate of synthesis, whether the synthesis that produces the mRNAs in question is transcription, as discussed above, or alternative splicing. The use of a Dbr1p-deficient strain, which stabilizes the alternatively excised intron lariats equivalently, results in intron RNA lariat levels that directly reflect the rate of alternative splicing.

The methods reported here can also be applied to genome-wide analysis of introns themselves and are an improvement on previous analyses that also directly analyzed intron RNA lariats. In one study (Zhang et al. 2007), researchers sought to identify introns on a genome-wide scale by analyzing lariat RNA populations within *S. cerevisiae* cells. They accomplished this by identifying RNA sequences that accumulate at elevated levels in a *dbr1* mutant strain compared to a wild-type strain. These researchers used a genomic tiling array with probe sequences spaced an average of 5 bp apart to map the origins of the elevated RNAs. The vast majority of RNAs with enhanced expression in the *dbr1* strain corresponded to the lariat portions of annotated introns but several new introns were also identified. However, introns of genes that exhibit low levels of transcription were not detected. In fact, <150 yeast introns were detected in their experiments (about 50% of yeast introns), even though almost all the undetected introns were in expressed genes (i.e., expressed in the growth condition used for the study). In another study of yeast introns that also compared RNAs accumulating in *dbr1Δ* and wild-type strains using high density tiling arrays (Juneau et al. 2007), the sensitivity of the method was enhanced by optimizing signal analysis by training the software on known intron and exon sequences. In this study, 76% of known introns were detected and several new introns were discovered.

High-throughput sequencing has recently been applied to the analysis of whole transcriptomes (RNA-seq) and has proven to be a powerful method for deducing the presence of introns and assessing alternative splicing patterns (Cloonan et al. 2008; Lister et al. 2008; Mortazavi et al. 2008; Nagalakshmi et al. 2008; Pan et al. 2008; Sultan et al. 2008; Wilhelm et al. 2008). Analysis of RNA-seq data from the *S. cerevisiae* transcriptome detected 60% more introns than Zhang et al. (2007), identifying 240 of the 306 known yeast introns (Nagalakshmi et al. 2008).

RNA-seq of intron RNA lariat populations prepared using PNPase can provide complementary information to RNA-seq of whole transcriptomes and may reveal new lariat sequences not evident from transcriptome analysis alone. Intron RNA lariat levels can be greatly enhanced by blocking the RNA debranching reaction, which increases the likelihood of detecting even rare splicing events. Because cells defective for RNA debranching activity accumulate excised introns in their lariat forms, with shorted 3′ tails (Chapman and Boeke 1991; Salem et al. 2003), information on the 3′ intron–exon junction is not obtainable from intron lariat RNA sequences. Nevertheless, lariat sequences provide information about branch points that is not obtainable from whole transcriptome sequencing (Vogel et al. 1997). Such information is especially useful for studies of introns in organisms whose branch point sequences are not as highly conserved as those in *S. cerevisiae* [e.g., humans (see Gao et al. (2008) and references therein)]. Finally, the absence of known intron sequences from an

RNA population enriched for RNA lariats can indicate that a gene is not expressed under the growth regimen employed. However, if an intron-containing gene is known to be expressed during the experiment, absence of intron sequences from the RNA lariat population could be an indication that the intron is removed by the hydrolytic splicing pathway observed for self splicing group II introns rather than the predominant branching pathway (Daniels et al. 1996; Vogel and Borner 2002).

A recent study of human branch points was performed to better establish their sequence characteristics (Gao et al. 2008). To this end, RT-PCR and sequencing of 367 individual intron lariat RNA branch points were performed. The method relied on the characteristic misincorporation of a single nucleotide during reverse transcription across the 2′–5′ bond as an indicator of the position of the branch point. However, branch points could be established for less than half of the introns (181 of the 367 introns) because RT-PCR products for the others lacked an apparent misincorporated nucleotide. Even though the work of Gao et al. (2008) represents an important advance, their study only analyzed introns from 20 housekeeping genes. The effect of this limited sampling on the resulting branch point consensus sequence is not known. High-throughput sequencing of enriched lariat RNAs from human cells will be useful for much more detailed analysis of human branch point sequences.

Recently, a computational model for identifying mammalian branch point sequences has been developed that relies, in part, on splicing factor 1 (SF1) binding affinity data (Pastuszak et al. 2010). However, in a test of the model, it only predicted 36% of the actual branch points identified by Gao et al. (2008). The study's authors concluded that other criteria are needed to enhance branch point sequence predictions. Even with better models, it is clear that validation of any predicted branch point will be required. The RNA-seq method proposed here will be of great utility in efforts to identify or validate branch point sequences.

We are currently applying the methods described here to test our model that Ty1 RNA forms a transient lariat during the reverse transcription step of retrotransposition. We are also interested to see if these methods can be used to identify any novel lariat or other non-linear RNAs in *S. cerevisiae*. Such RNAs, whether Ty1 RNA or another RNA, may be the key to unlock the mystery of how Dbr1p acts as a host factor for Ty1 retrotransposition.

Acknowledgments We thank Beate Schwer for the gift of pET16b-DBR1 and Haoping Liu for yeast sigma strain 10560-23C. Support was provided by National Science Foundation, the University of Missouri Research Board, and the University of Missouri-Kansas City School of Biological Sciences.

References

Abelson J (2008) Is the spliceosome a ribonucleoprotein enzyme? Nat Struct Mol Biol 15(12):1235–1237

Alani E, Cao L, Kleckner N (1987) A method for gene disruption that allows repeated use of *URA3* selection in the construction of multiply disrupted yeast strains. Genetics 116(4):541–545

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (2003) Current protocols in molecular biology on CD-ROM. Current Protocols Inc., Brooklyn

Berezikov E, Chung WJ, Willis J, Cuppen E, Lai EC (2007) Mammalian mirtron genes. Mol Cell 28(2):328–336

Bushman FD, Malani N, Fernandes J, D'Orso I, Cagney G, Diamond TL, Zhou H, Hazuda DJ, Espeseth AS, Konig R, Bandyopadhyay S, Ideker T, Goff SP, Krogan NJ, Frankel AD, Young JA, Chanda SK (2009) Host cell factors in HIV replication: meta-analysis of genome-wide studies. PLoS Pathog 5(5):e1000437

Chapman KB, Boeke JD (1991) Isolation and characterization of the gene encoding yeast debranching enzyme. Cell 65:483–492

Cheng Z, Menees TM (2004) RNA branching and debranching in the yeast retrovirus-like element Ty1. Science 303:240–243

Cloonan N, Forrest AR, Kolle G, Gardiner BB, Faulkner GJ, Brown MK, Taylor DF, Steptoe AL, Wani S, Bethel G, Robertson AJ, Perkins AC, Bruce SJ, Lee CC, Ranade SS, Peckham HE, Manning JM, McKernan KJ, Grimmond SM (2008) Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat Methods 5(7):613–619

Conklin JF, Goldman A, Lopez AJ (2005) Stabilization and analysis of intron lariats in vivo. Methods 37(4):368–375

Coombes CE, Boeke JD (2005) An evaluation of detection methods for large lariat RNAs. RNA 11(3):323–331

Daniels DL, Michels WJ Jr, Pyle AM (1996) Two competing pathways for self-splicing by group II introns: a quantitative analysis of in vitro reaction rates and products. J Mol Biol 256(1):31–49

Engleman A (2010) Reverse transcription and integration. In: Kurth R, Bannert N (eds) Retroviruses: molecular biology, genomics, and pathogenesis. Caister Academic Press, Norfolk, pp 129–159

Falaleeva MV, Chetverina HV, Ugarov VI, Uzlova EA, Chetverin AB (2008) Factors influencing RNA degradation by *Thermus thermophilus* polynucleotide phosphorylase. FEBS J 275(9):2214–2226

Filipowicz W, Pogacic V (2002) Biogenesis of small nucleolar ribonucleoproteins. Curr Opin Cell Biol 14(3):319–327

Gallwitz D, Sures I (1980) Structure of a split yeast gene: complete nucleotide sequence of the actin gene in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 77(5):2546–2550

Gao K, Masuda A, Matsuura T, Ohno K (2008) Human branch point consensus sequence is yUnAy. Nucleic Acids Res 36(7):2257–2267

Garcia-Martinez J, Aranda A, Perez-Ortin JE (2004) Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms. Mol Cell 15(2):303–313

Goff SP (2007) Retroviridae: the retroviruses and their replication. In: Knipe DM, Howley PM (eds) Fields virology. Lipincott, Williams, and Wilkins, Philadelphia, pp 1999–2069

Gray M, Kupiec M, Honigberg SM (2004) Site-specific genomic (SSG) and random domain-localized (RDL) mutagenesis in yeast. BMC Biotechnol 4:7

Griffith JL, Coleman LE, Raymond AS, Goodson SG, Pittard WS, Tsui C, Devine SE (2003) Functional genomics reveals relationships between the retrovirus-like Ty1 element and its host *Saccharomyces cerevisiae*. Genetics 164(3):867–879

Guarneros G, Portier C (1990) Different specificities of ribonuclease II and polynucleotide phosphorylase in 3′mRNA decay. Biochimie 72(11):771–777

Hallegger M, Llorian M, Smith CW (2010) Alternative splicing: global insights. FEBS J 277(4):856–866

Hill JE, Myers AM, Koerner TJ, Tzagoloff A (1986) Yeast/*E. coli* shuttle vectors with multiple unique restriction sites. Yeast 2(3):163–167

Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA (1998) Dissecting the regulatory circuitry of a eukaryotic genome. Cell 95(5):717–728

Irwin B, Aye M, Baldi P, Beliakova-Bethell N, Cheng H, Dou Y, Liou W, Sandmeyer S (2005) Retroviruses and yeast retrotransposons use overlapping sets of host genes. Genome Res 15(5):641–654

Juneau K, Palm C, Miranda M, Davis RW (2007) High-density yeast-tiling array reveals previously undiscovered introns and extensive regulation of meiotic splicing. Proc Natl Acad Sci USA 104(5):1522–1527

Kaiser C, Michaelis S, Mitchell A (1994) Methods in yeast genetics. CSHL Press, Cold Spring Harbor

Karst SM, Rutz M-L, Menees TM (2000) The yeast retrotransposons Ty1 and Ty3 required the RNA lariat debranching enzyme, Dbr1p, for efficient accumulation of reverse transcripts. Bio-Chem Biophys Res Comm 268:112–117

Kataoka N, Fujita M, Ohno M (2009) Functional association of the Microprocessor complex with the spliceosome. Mol Cell Biol 29(12):3243–3254

Khalid MF, Damha MJ, Shuman S, Schwer B (2005) Structure-function analysis of yeast RNA debranching enzyme (Dbr1), a manganese-dependent phosphodiesterase. Nucleic Acids Res 33(19):6349–6360

Kim YK, Kim VN (2007) Processing of intronic microRNAs. EMBO J 26(3):775–783

Kiss T, Fayet E, Jady BE, Richard P, Weber M (2006) Biogenesis and intranuclear trafficking of human box C/D and H/ACA RNPs. Cold Spring Harb Symp Quant Biol 71:407–417

Lestrade L, Weber MJ (2006) snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. Nucleic Acids Res 34(database issue):D158–162

Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. Cell 133(3):523–536

Loeb JD, Kerentseva TA, Pan T, Sepulveda-Becerra M, Liu H (1999) *Saccharomyces cerevisiae* G1 cyclins are differentially involved in invasive and pseudohyphal growth independent of the filamentation mitogen-activated protein kinase pathway. Genetics 153(4):1535–1546

Lopez PJ, Seraphin B (2000) YIDB: the Yeast Intron DataBase. Nucleic Acids Res 28(1):85–86

Martin A, Schneider S, Schwer B (2002) Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. J Biol Chem 277(20):17743–17750

McLaren RS, Newbury SF, Dance GS, Causton HC, Higgins CF (1991) mRNA degradation by processive 3'–5' exoribonucleases in vitro and the implications for prokaryotic mRNA decay in vivo. J Mol Biol 221(1):81–95

Miller C, Schwalb B, Maier K, Schulz D, Dumcke S, Zacher B, Mayer A, Sydow J, Marcinowski L, Dolken L, Martin DE, Tresch A, Cramer P (2011) Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. Mol Syst Biol 7:458

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5(7):621–628

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. Science 320(5881):1344–1349

Nam K, Hudson RH, Chapman KB, Ganeshan K, Damha MJ, Boeke JD (1994) Yeast lariat debranching enzyme. Substrate and sequence specificity. J Biol Chem 269(32):20613–20621

Ng R, Abelson J (1980) Isolation and sequence of the gene for actin in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 77(7):3912–3916

Nilsen TW, Graveley BR (2010) Expansion of the eukaryotic proteome by alternative splicing. Nature 463(7280):457–463

Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC (2007) The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. Cell 130(1):89–100

Ooi SL, Samarsky DA, Fournier MJ, Boeke JD (1998) Intronic snoRNA biosynthesis in *Saccharomyces cerevisiae* depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. RNA 4(9):1096–1110

Ooi SL, Dann C 3rd, Nam K, Leahy DJ, Damha MJ, Boeke JD (2001) RNA lariat debranching enzyme. Methods Enzymol 342:233–248

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 40(12):1413–1415

Pastuszak AW, Joachimiak MP, Blanchette M, Rio DC, Brenner SE, Frankel AD (2010) An SF1 affinity model to identify branch point sequences in human introns. Nucleic Acids Res 39(6):2344–2356

Pelechano V, Perez-Ortin JE (2010) There is a steady-state transcriptome in exponentially growing yeast cells. Yeast 27(7):413–422

Pratico ED, Silverman SK (2007) Ty1 reverse transcriptase does not read through the proposed 2', 5'-branched retrotransposition intermediate in vitro. RNA 13(9):1528–1536

Preker PJ, Guthrie C (2006) Autoregulation of the mRNA export factor Yra1p requires inefficient splicing of its pre-mRNA. RNA 12(6):994–1006

Preker PJ, Kim KS, Guthrie C (2002) Expression of the essential mRNA export factor Yra1p is autoregulated by a splicing-dependent mechanism. RNA 8(8):969–980

Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A (2004) Identification of mammalian microRNA host genes and transcription units. Genome Res 14(10A):1902–1910

Rodriguez-Navarro S, Strasser K, Hurt E (2002) An intron in the *YRA1* gene is required to control Yra1 protein expression and mRNA export in yeast. EMBO Rep 3(5):438–442

Ruby JG, Jan CH, Bartel DP (2007) Intronic microRNA precursors that bypass Drosha processing. Nature 448(7149):83–86

Salem LA, Boucher CL, Menees TM (2003) Relationship between RNA lariat debranching and yeast Ty1 element retrotransposition. J Virol 77:12795–12806

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 3(6):1101–1108

Sikorski RS, Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. Genetics 122(1):19–27

Smale ST (2009). Nuclear run-on assay. Cold Spring Harb Protoc 2009(11): pdb prot5329

Smith DJ, Query CC, Konarska MM (2008) "Nought may endure but mutability": spliceosome dynamics and the regulation of splicing. Mol Cell 30(6):657–666

Spingola M, Grate L, Haussler D, Ares M Jr (1999) Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. RNA 5(2):221–234

Storici F, Lewis LK, Resnick MA (2001) In vivo site-directed mutagenesis using oligonucleotides. Nat Biotechnol 19(8):773–776

Sultan M, Schulz MH, Richard H, Magen A, Klingenhoff A, Scherf M, Seifert M, Borodina T, Soldatov A, Parkhomchuk D, Schmidt D, O'Keeffe S, Haas S, Vingron M, Lehrach H, Yaspo ML

(2008) A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. Science 321(5891):956–960

Suzuki H, Zuo Y, Wang J, Zhang MQ, Malhotra A, Mayeda A (2006) Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. Nucleic Acids Res 34(8):e63

Tang GQ, Maxwell ES (2008) *Xenopus* microRNA genes are predominantly located within introns and are differentially expressed in adult frog tissues via post-transcriptional regulation. Genome Res 18(1):104–112

Vincent HA, Deutscher MP (2006) Substrate recognition and catalysis by the exoribonuclease RNase R. J Biol Chem 281(40):29769–29775

Vogel J, Borner T (2002) Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. EMBO J 21(14): 3794–3803

Vogel J, Hess WR, Borner T (1997) Precise branch point mapping and quantification of splicing intermediates. Nucleic Acids Res 25(10):2030–2031

Wahl MC, Will CL, Luhrmann R (2009) The spliceosome: design principles of a dynamic RNP machine. Cell 136(4):701–718

Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, Penkett CJ, Rogers J, Bahler J (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature 453(7199):1239–1243

Ye Y, De Leon J, Yokoyama N, Naidu Y, Camerini D (2005) DBR1 siRNA inhibition of HIV-1 replication. Retrovirology 2:63

Zhang Z, Hesselberth JR, Fields S (2007) Genome-wide identification of spliced introns using a tiling microarray. Genome Res 17(4):503–509