Supplementary methods

Construction of yeast strains and standard molecular biology analyses.

Yeast strains used in this paper are listed in Supplementary Table 1. Gene deletions, tagging and insertion of the *GAL1* promoter (P_{Gal}) were performed with standard procedures (Longtine et al, 1998; Rigaut et al, 1999). Plasmids used in this work are shown in Supplementary Table 2. All plasmids are derived from pCM190(Gari et al, 1997) and were constructed by homologous recombination in yeast. The *CUP1* based transcription interference reporter, P_{Tet}-*HSP104*$_{1-1000}$-P_{Gal}-*CUP1* (pDL367) was constructed by cloning the P_{Gal} promoter (positions -542 to –1 relative to the start codon) followed by the *CUP1* ORF immediately downstream of the doxycycline repressible promoter P_{Tet} in pCM190. A stuffer fragment derived from the coding region of the *HSP104* gene (position 1-1000 relative to the start codon) was introduced between the two promoters and the random pool was cloned by recombination between positions +126 and +754 of the *HSP104* sequence after digestion with PflM1 and Tth111 restriction enzymes. P_{Tet}- *LEU2* –HSP104$_{465-1000}$-P_{Gal}-*CUP1* (pDL519) was constructed by replacing the 1-465 fragment of *HSP104* in pDL367 with the *LEU2* gene from *Candida glabrata*. The putative terminators and their mutant derivatives were inserted after the stop codon of *CgLEU2*.

RNAs were analyzed by 1.2% agarose or 5% polyacrylamide gels depending on the resolution range required. Hybridizations were performed using a commercial buffer (Ultrahyb, Ambion) and radiolabeled probes were prepared by random priming of short double-stranded fragments (Magaprime kit, GE Healthcare) or phosphorylation of oligonucleotides. All oligonucleotides used in this work are listed in Supplementary Table 3.

Proteins purification.

Details on cloning, expression and purification of the Nab3$_{321-415}$ and Nab3$_{191-565}$ – Nrd1$_{1-548}$ heterodimer constructs have been described previously (Carroll et al, 2007; Hobor et al, 2011).

Cloning, expression and purification of Nab3 αhxRRM (283-415). The coding sequence corresponding to the N-terminally extended RRM of the Nab3 gene from *Saccharomyces cerevisiae* (847-1245) was amplified by polymerase chain reaction (PCR), and cloned into a pET22b expression vector (Novagen) via NdeI and XhoI restriction sites. The resulting C-terminal 6x histidine-tagged construct was verified by DNA sequencing. The protein was overexpressed in *E.coli* BL21-Codon Plus (DE3)-RIPL (Stratagene), transformed with the pET22b-αhxRRM Nab3 construct at 37 °C in M9 minimal medium, supplemented with 50

mg/l Ampicilin. For isotope labeling, the medium was supplemented with $^{15}NH_4Cl$. The cells were grown at 37°C to OD_{600} ~0.7 and induced with 1mM isopropyl β-D-thiogalactoside (IPTG). Cells were harvested by centrifugation (6000 rpm, 7 minutes), resuspended in lysis buffer (LB) (50 mM sodium phosphate, 300 mM NaCl, 10 mM β–mercaptoethanol, pH 8) supplemented with 20% sucrose, and disrupted by sonication. The cell debris were cleared by centrifugation (14000 rpm, 60 minutes). Soluble lysate was loaded on Ni-NTA column (QIAGEN), equilibrated with the LB, washed with LB supplemented with 5 mM imidazole, and eluted with imidazole gradient (50 – 500 mM) of LB. The protein was subsequently loaded on a Superdex 75 gel filtration column (GE Healthcare), equilibrated with the LB. The purified protein was 99% pure, as judged by Coomassie-stained SDS-PAGE. For the fluorescence anisotropy and NMR measurements the pure protein was concentrated up to 3 mM in 50 mM sodium phosphate (pH 8.0), containing 300 mM NaCl, and 10 mM β–mercaptoethanol. Prior to measurement, the protein was tested for the residual RNase activity using RNaseAlert Lab Test (Ambion).

References

Carroll KL, Ghirlando R, Ames JM, Corden JL (2007) Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. *Rna* **13**(3)**:** 361-373

Gari E, Piedrafita L, Aldea M, Herrero E (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in Saccharomyces cerevisiae. *Yeast* **13**(9)**:** 837-848

Hobor F, Pergoli R, Kubicek K, Hrossova D, Bacikova V, Zimmermann M, Pasulka J, Hofr C, Vanacova S, Stefl R (2011) Recognition of transcription termination signal by the nuclear polyadenylated RNA-binding (NAB) 3 protein. *J Biol Chem* **286**(5)**:** 3645-3657

Longtine MS, McKenzie A, 3rd, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* **14**(10)**:** 953-961

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**(10)**:** 1030-1032

Supplementary figures legends

Supplementary Figure 1. A. Schematics of the selection system. Expression of *CUP1*, under control of the P_{Gal} promoter confers copper resistant growth to yeast. Transcription from the upstream P_{Tet} promoter inhibits expression of *CUP1* unless a functional terminator (black box) is present upstream of P_{Gal}. **B.** Growth on copper containing plates of cells expressing a reporter containing a fragment of the *HSP104* gene (lacking a terminator) or the *NEL025c* sequence, containing NNS-dependent termination signals. **C.** Northern blot analysis of *CUP1* mRNA expression in constructs containing or lacking a terminator. A short, functional *CUP1* mRNA is only produced when the construct contains a terminator or when P_{Tet} is repressed by the addition of doxycycline (+dox). Note that the longer RNA transcribed from P_{Tet} does not produce a functional Cup1p. The images shown in the figure were obtained from the same gel (see source files). **D.** Northern blot analysis of transcripts produced from P_{Tet} in the absence of doxycycline. Note that the short transcripts produced in the presence of the *NEL025c* sequence are unstable and only revealed in the absence of the exosome component Rrp6p. The two arrows indicate read-through transcripts terminating at the *CUP1* terminator and at cryptic signals within P_{Gal}.

Supplementary Figure 2. A. Scheme of the selection procedure employed. A pool of random sequences was obtained by annealing and extending two oligonucleotides containing 60 random nucleotides. The pool was cloned by recombination directly into a Δ*cup1* yeast strain. After selection by growth on copper plates, the clones were pooled and the inserts were amplified by PCR for a second round of selection. Inserts from the second round were barcoded and sequenced together with inserts from the starting pool. Two classes of terminators were identified according to their likely dependency on the NNS complex. **B.** Examples of sequences from the naïve (non-selected) and selected pool of terminators belonging to class 1 of terminators (NNS-dependent).

Supplementary Figure 3. A. Analysis of the nucleotides flanking the known Nrd1 and Nab3 sites. Z-scores dispersion plot analysis as in Figure 2a for all pentanucleotides containing either one of the 4-bases motifs TCTT, GTAA or GTAG as indicated. The relative position of the pentanucleotides of a given class indicates the level of preference for a given flanking nucleotide. For instance the statistical significance of the overabundance of TCTTG is much higher than that of the three

other possibilities (TCTTA/C/T) indicating that a G is preferred 3' of TCTT. **B.** P-values for the most significant motifs of 4 and 5 nucleotides, calculated by RSAT separately for the two regions of the artificial terminators. Oligonucleotides containing Nrd1 or Nab3 sites are highlighted in orange or green respectively.

Supplementary Figure 4. Northern blot analysis of termination induced by clone 78 or by its A-region alone. The sequence of part A is indicated on the top together with a scheme of the construct. The Nrd1p and Nab3p binding sites as well as the AU-rich motif revealed by the statistical analysis are highlighted.

Supplementary Figure 5. Copper growth assay using clone 78 and its mutant derivatives. Sequences of the clones are shown in Figures 3a and 4a except for the 78-INV-μAU mutant in which the AU-rich motif was modified in the context of 78-INV (B).

Supplementary Figure 6. Northern blot analysis of RNAs produced from clones 78 (**A**) and 78-INV (**B**) during Nrd1p and Nab3p metabolic depletion. Note that read-through at terminator 78 does not increase upon longer time points of Nrd1p depletion. *NEL025c* is detected (**A,** lower panel**)** as control that is strongly sensitive to Nrd1p.

Supplementary Figure 7. The AU-rich termination motif is not recognized by Hrp1p or Rna15p. Northern blot analysis of transcription termination induced by clone 78, containing the AU-rich motif, upon metabolic depletion of Hrp1p (left panel) or heat inactivation of an *rna15-3* mutant (right panel). Experiments were performed in a Δ*rrp6* strain to visualize the unstable termination products. Expression of the *SUA7* mRNA was assessed to control for the efficient inactivation of Hrp1p and Rna15p functions.

Supplementary Figure 8. Poly(A) or poly(U) can functionally replace the AU-rich element. **A.** Sequence of clone 78 and its mutant variants 78-A_8 and 78-U_8. **B.** Analysis of transcription termination by northern blot (left) and copper growth assays (right) for the constructs indicated. Read-through transcripts are indicated by black arrows. A construct harboring 1Kb of *HSP104* coding sequence (no termination signals) was used as a negative control.

Supplementary Figure 9. 1H-15N HSQC spectra of Nab3p RRM (321-415) alone (in blue) and in the presence of 1 eq of 5'-UCUU-3' (in red) and 1 eq of 5'-UCUUG-3'(in black) at 303K.

Supplementary Figure 10. Termination induced by artificial terminators in the 3'-end of a gene is not dependent on the NNS complex. Northern blot analysis of clones 78 and 78-INV inserted either in a promoter proximal (as a positive control) or a distal position as indicated (5' *vs* 3'). P_{Gal}-*NAB3*, Δ*rrp6* cells were grown in galactose or in glucose for 14 hrs to deplete Nab3p.

SUPPLEMENTARY TABLES

Supplementary Table 1. Yeast strains used in this study.

Name	Number	Genotype	Source
WT	W303	*ura3-1, ade2-1, his3-11,5, trp1-1, leu2-3,112, can1-100*	(Thomas & Rothstein, 1989)
BMA	DLY671	*as W303, ∆trp1, Mat a*	F. Lacroute
∆*cup1*	DLY841	*as W303, cup1::KAN*	This study
∆*rrp6*	DLY814	*as BMA, rrp6::KAN*	F. Lacroute
P_{Gal}-*NRD1,* ∆*rrp6*	DLY885	*as W303, HIS::Pgal::NRD1, rrp6::KAN*	(Thiebaut et al, 2006)
P_{Gal}-*NAB3,* ∆*rrp6*	DLY891	*as W303, HIS::Pgal::NAB3, rrp6::KAN*	(Thiebaut et al, 2006)
P_{Gal}-*SEN1,* ∆*rrp6*	DLY1676	*as BMA, TRP1::Pgal::TAP::SEN1, rrp6::KAN*	This study
P_{Gal}-*HRP1,* ∆*rrp6*	DLY2019	*as BMA, HIS::Pgal::NAB3, rrp6::URA3*	This study
rna15-2	DLY128	as W303, *rna15-2*	F. Lacroute
rna15-2, ∆*rrp6*	DLY172	as W303, *rna15-2, rrp6::KAN*	(Libri et al, 2002)
rna14-3, ∆*rrp6*	DLY171	as W303, *rna14-3, rrp6::KAN*	(Libri et al, 2002)

Supplementary Table 2. Plasmids used in this study.

Name	Description	Vector	Reference
pDL301	P_{Tet}-P_{Gal}-*CUP1*	pCM190	This study
pDL367	P_{Tet}-HSP104$_{1-1000}$-PGal-*CUP1*	pCM190	This study
pDL386	P_{Tet}- *NEL025c* -P_{Gal}-*CUP1*	pCM190	This study
pDL400	P_{Tet}- HSP104$_{1-400}$- P_{Gal}-*CUP1*	pCM190	This study
pDL403	P_{Tet}- 78-μAU P_{Gal}-*CUP1*	pCM190	This study
pDL404	P_{Tet}- 78-μG μAU-P_{Gal}-*CUP1*	pCM190	This study
pDL405	P_{Tet}-78-μNab3$_1$-P_{Gal}-*CUP1*	pCM190	This study
pDL406	P_{Tet}-78-μNrd1-P_{Gal}-*CUP1*	pCM190	This study
pDL407	P_{Tet}-78-μNrd1-μG-P_{Gal}-*CUP1*	pCM190	This study
pDL464	P_{Tet}-78-wt-P_{Gal}-*CUP1*	pCM190	This study
pDL500	P_{Tet}-78-μG-P_{Gal}-*CUP1*	pCM190	This study
pDL501	P_{Tet}-78-INV-P_{Gal}-*CUP1*	pCM190	This study
pDL502	P_{Tet}-78-INV-μNrd1-P_{Gal}-*CUP1*	pCM190	This study
pDL503	P_{Tet}-78-μNab3$_2$-P_{Gal}-*CUP1*	pCM190	This study
pDL507	P_{Tet}- *LEU2* –78-wt-P_{Gal}-*CUP1*	pCM190	This study
pDL519	P_{Tet}- *LEU2* –HSP104$_{465-1000}$-P_{Gal}-*CUP1*	pCM190	This study
pDL520	P_{Tet}-78-A_8-P_{Gal}-*CUP1*	pCM190	This study
pDL521	P_{Tet}-78-U_8-P_{Gal}-*CUP1*	pCM190	This study
pDL522	P_{Tet}-78-INV -μNab3-P_{Gal}-*CUP1*	pCM190	This study
pDL526	P_{Tet}- *LEU2* –78-μG -P_{Gal}-*CUP1*	pCM190	This study

pDL527	P_{Tet}- *LEU2* –78-INV -P_{Gal}-*CUP1*	pCM190	This study
pDL528	P_{Tet}- *LEU2* –78- INV-μNrd1-P_{Gal}-*CUP1*	pCM190	This study
pDL529	P_{Tet}- *LEU2* – 78-INV-μNab3$_1$-P_{Gal}-*CUP1*	pCM190	This study
pDL540	P_{Tet}- *LEU2* –78-μAU-P_{Gal}-*CUP1*	pCM190	This study
pDL541	P_{Tet}- *LEU2* –78-μNab3$_1$-P_{Gal}-*CUP1*	pCM190	This study

Supplementary Table 3. Oligonucleotides used in this work.

Name	Sequence (5'-3')	Information/use
DL1698	CTTCATTGAAACG-N_{60}-GATACCGGTCACCGCATTCG	Fwd primer for the generation of a pool of random sequences. Contains part A of pool
DL1665	GTTAATGCGG-N_{60}-CGAATGCGGTGACCGGTATC	Rev primer for the generation of a pool of random sequences. Contains part B of pool
DL1702	TTCGGATCATCAACATCCACAATTACAACC TATACATATTCTAGCTGCCTTCATTGAAACG	Fwd primer to amplify the pool of random sequences and clone in pDL367 by recombination.
DL1666	GAATCTTTCTTCGAAATCACCTTTGTATTTA GCACCTGCGGTTAATGCGG	Rev primer to amplify the pool of random sequences and clone in pDL367 by recombination
DL2402	TGATGCAGTAGCGAAAGCTGTCAGGGAAC TATTAGCTTAGCTTCATTGAAACG	Fwd primer to clone wt and mutant 78 derivatives in pDL519
DL2403	TTATTTCTTCTTCACGGCCGATGACAGGGT CAAGTTTACCGTTAATGCGG	Rev primer to clone wt and mutant 78 derivatives in pDL519
DL190	TTGAGCCAACGTCAAAATCGTTAGAGCCCT TTCTGTAAATTGCGTTTGGTCGTTCAT	Probe oligonucleotide annealing to HSP104 (+57)
DL1360	ATCCCCCGAATTGATCCGG	Probe oligonucleotide annealing to P_{Tet} in pCM190.
DL377	ATGTTCCCAGGTATTGCCGA	Fwd primer to generate an *ACT1* probe
DL378	ACACTTGTGGTGAACGATAG	Rev primer to generate an *ACT1* probe
DL474	GCAAAGATCTGTATGAAAGG	Rev primer to generate a *NEL025c* probe
DL478	CCTGTTGACATTGCAGACAA	Fwd primer to generate a *NEL025c* probe
DL750	GGTCATGAGTGCCAATGCCA	Rev primer to generate a *CUP1* probe
DL751	TTTCCCAGAGCAGCATGACT	Fwd primer to generate a *CUP1* probe

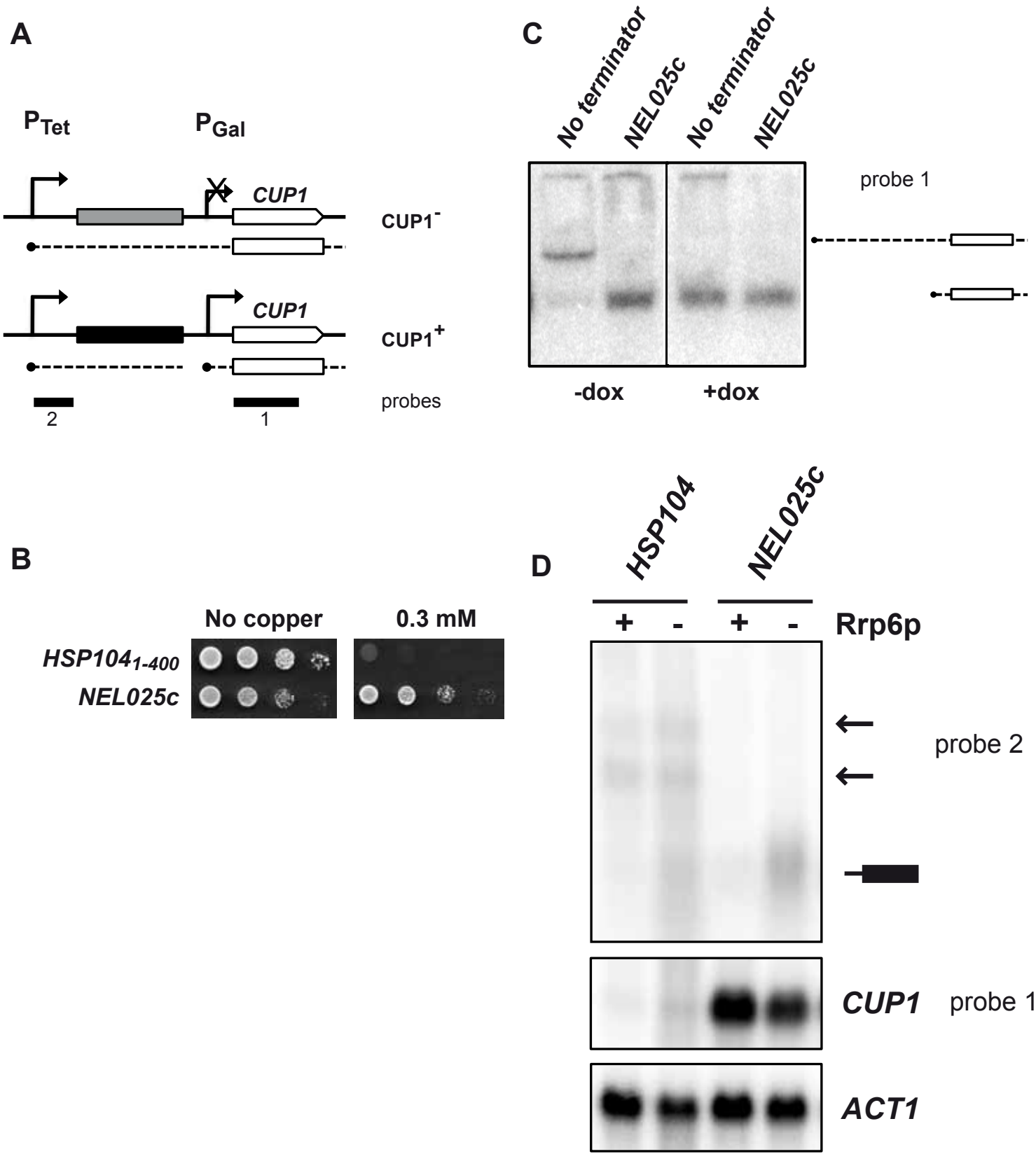
DL265	AACACAATCTCGGACGAATC	Probe oligonucleotide annealing to U4
DL270	AAAAAGTCTCTTCCCGTCCA	Probe oligonucleotide annealing to U2
DL2621	GATGATCACAACGGTGATGACC	Fwd primer to generate a *SUA7* probe
DL2622	GAAGTAGTGACTTGCATCGG	Rev primer to generate a *SUA7* probe

References

Libri D, Dower K, Boulay J, Thomsen R, Rosbash M, Jensen TH (2002) Interactions between mRNA export commitment, 3'-end quality control, and nuclear degradation. *Mol Cell Biol* **22**(23)**:** 8254-8266

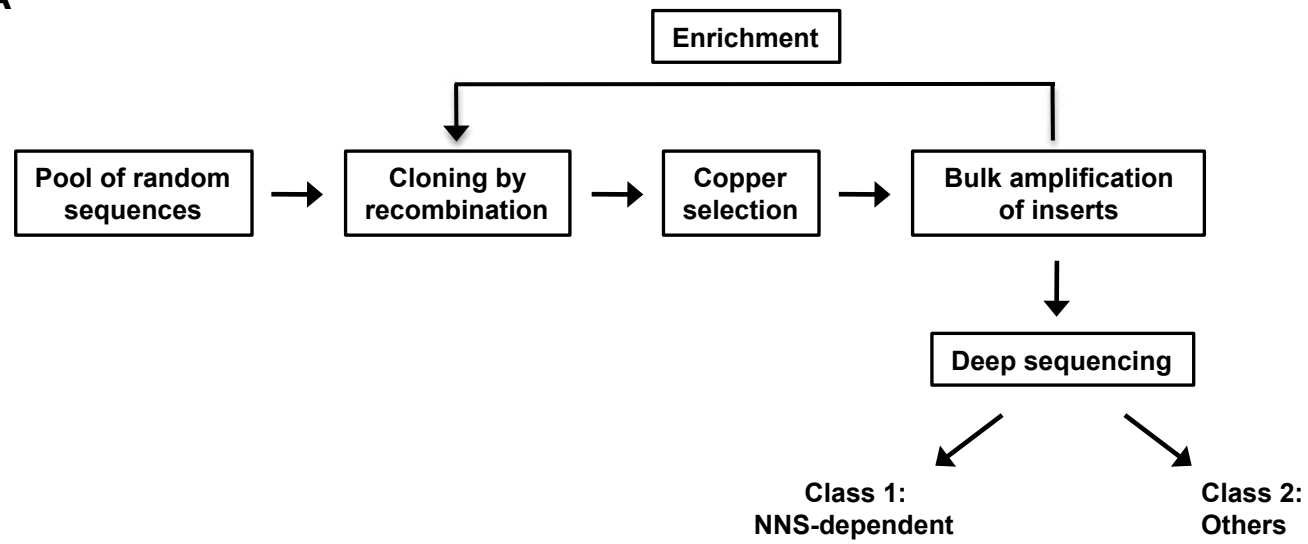
Thiebaut M, Kisseleva-Romanova E, Rougemaille M, Boulay J, Libri D (2006) Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance. *Mol Cell* **23**(6)**:** 853-864

Thomas BJ, Rothstein R (1989) Elevated recombination rates in transcriptionally active DNA. *Cell* **56**(4)**:** 619-630

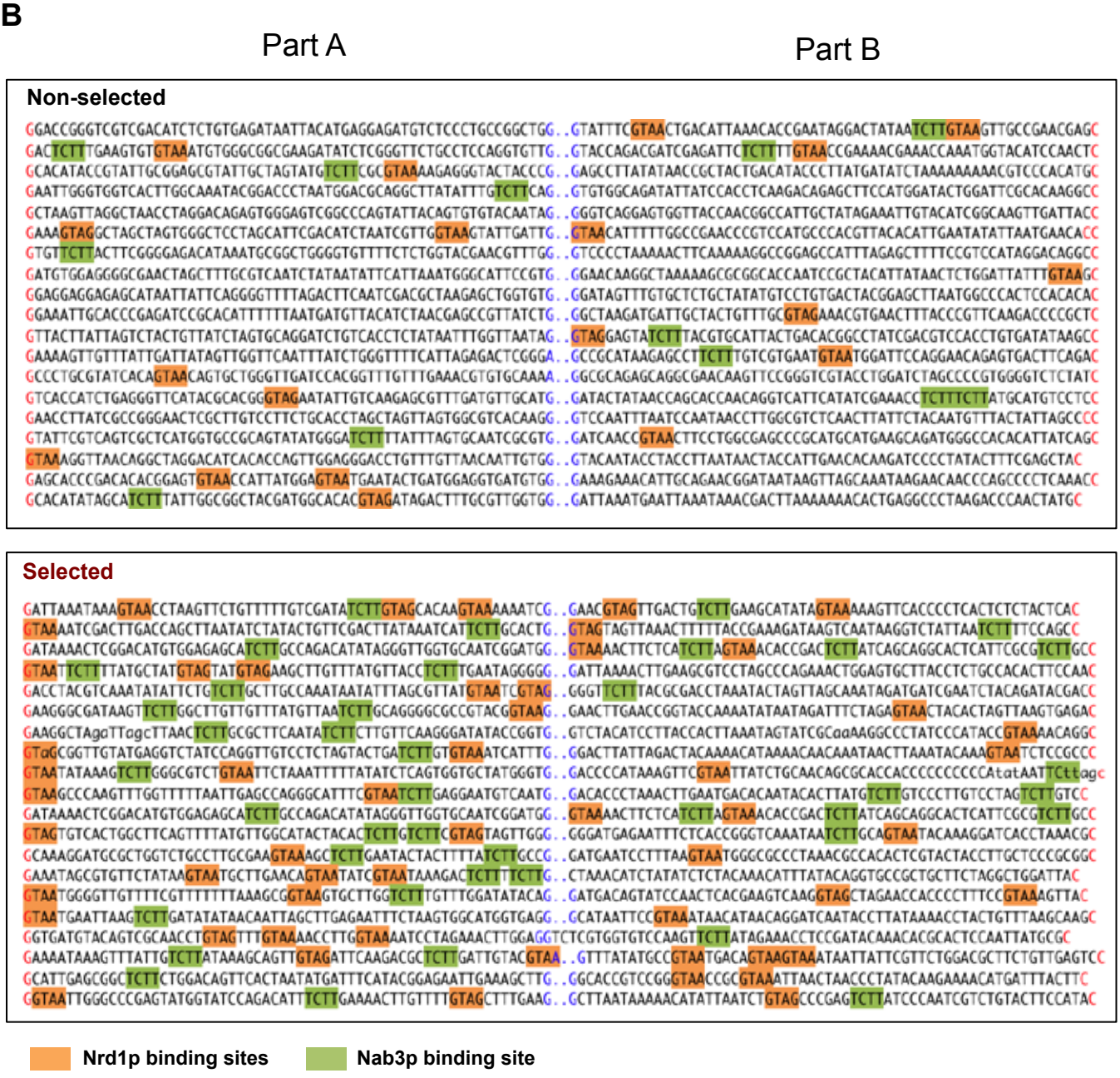


Porrua et al., Supplementary Fig. 1

A

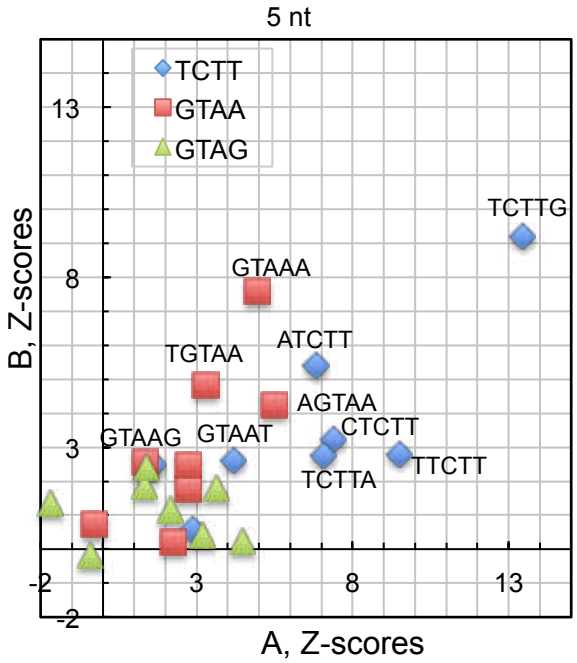


B



Porrua et al., Supplementary Fig. 2

A

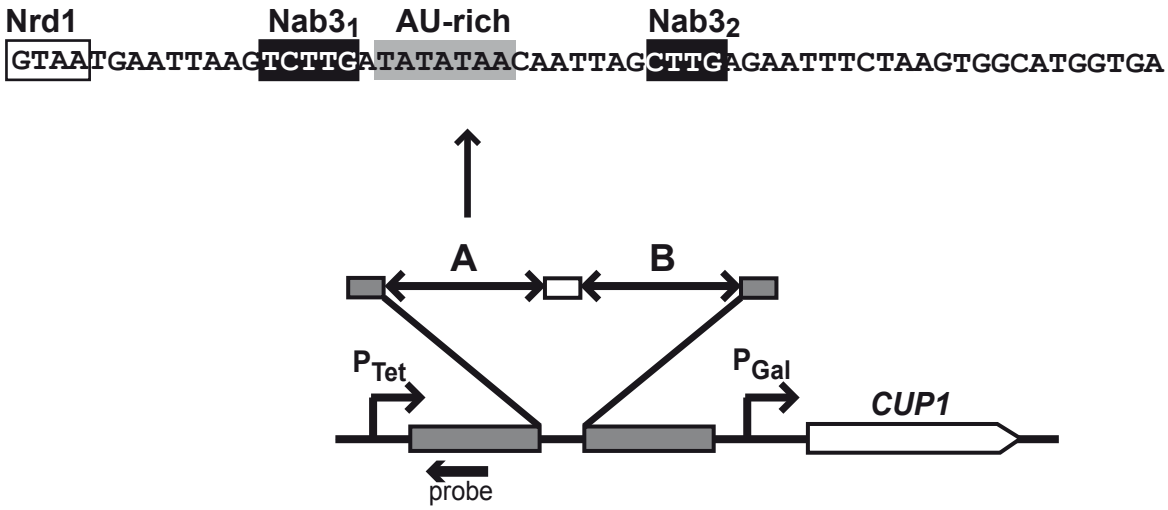


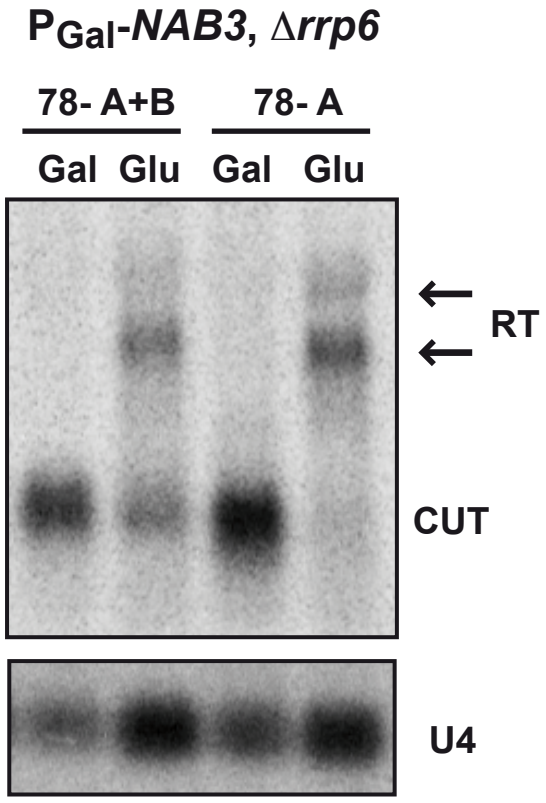
B

A		B	
sequence	P-value	sequence	P-value
TCTT	2.30E-31	TAAA	4.80E-18
CTTG	1.70E-21	AAAT	1.30E-11
TAAA	1.70E-14	TCTT	2.60E-11
TTTA	2.80E-09	GTAA	3.20E-09
TTCT	5.90E-09	AAAA	6.60E-08
TTAT	1.10E-08	AAAG	1.80E-06
AAAT	1.20E-08	ATAA	4.00E-06
AGTA	2.40E-07	AATA	4.30E-06
GTAA	3.00E-07	CTTG	5.20E-06
ATCT	9.70E-07	TGTA	9.30E-06

A		B	
sequence	P-value	sequence	P-value
TCTTG	4.40E-28	TCTTG	2.90E-14
TTCTT	7.40E-16	TAAAA	3.30E-12
TTTAT	4.40E-11	GTAAA	2.90E-11
CTTGT	1.90E-10	AAATA	9.20E-10
CTCTT	3.80E-10	TAAAT	1.00E-07
TCTTA	7.70E-10	AAAAT	4.40E-07
ATCTT	2.00E-09	ATCTT	8.50E-07
TAAAT	1.00E-08	ATAAA	2.30E-06
CTTGC	3.80E-08	AAAGT	4.00E-06
TAAAA	8.60E-08	TGTAA	6.90E-06

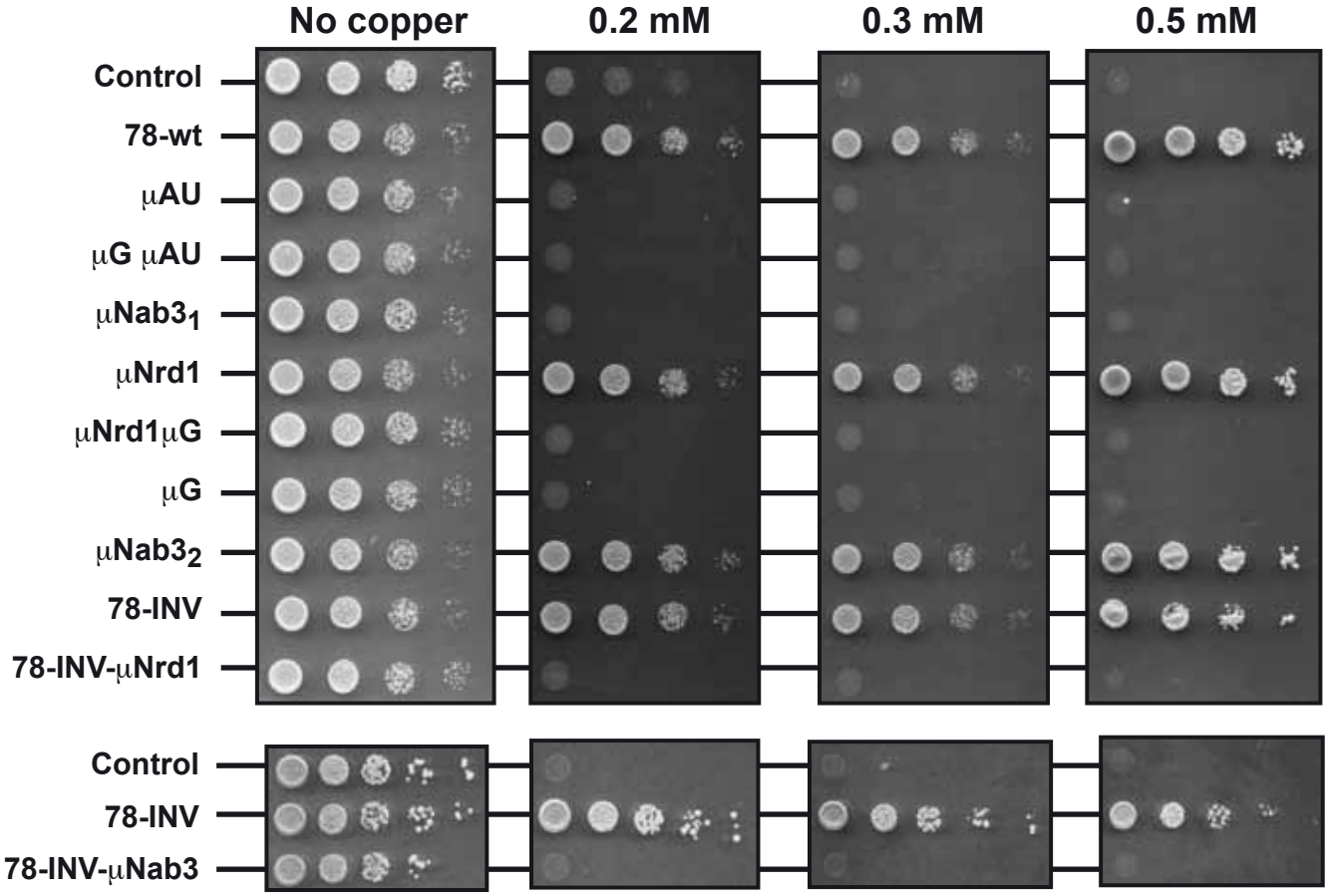
Porrua et al., Supplementary Fig. 3



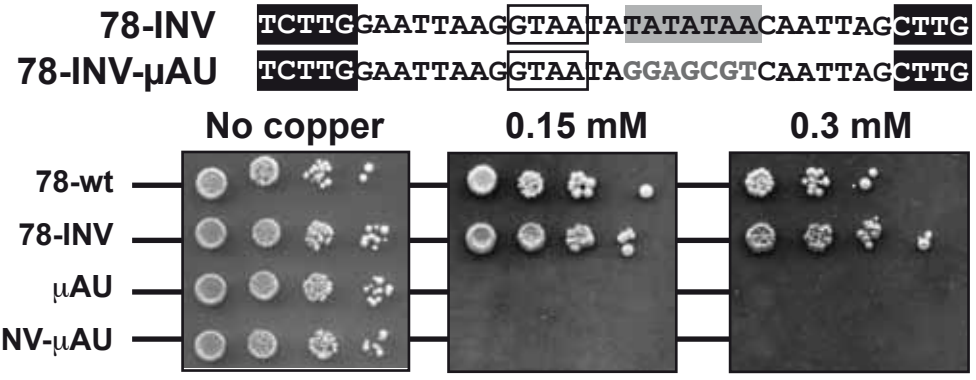


Porrua et al., Supplementary Fig. 4

A

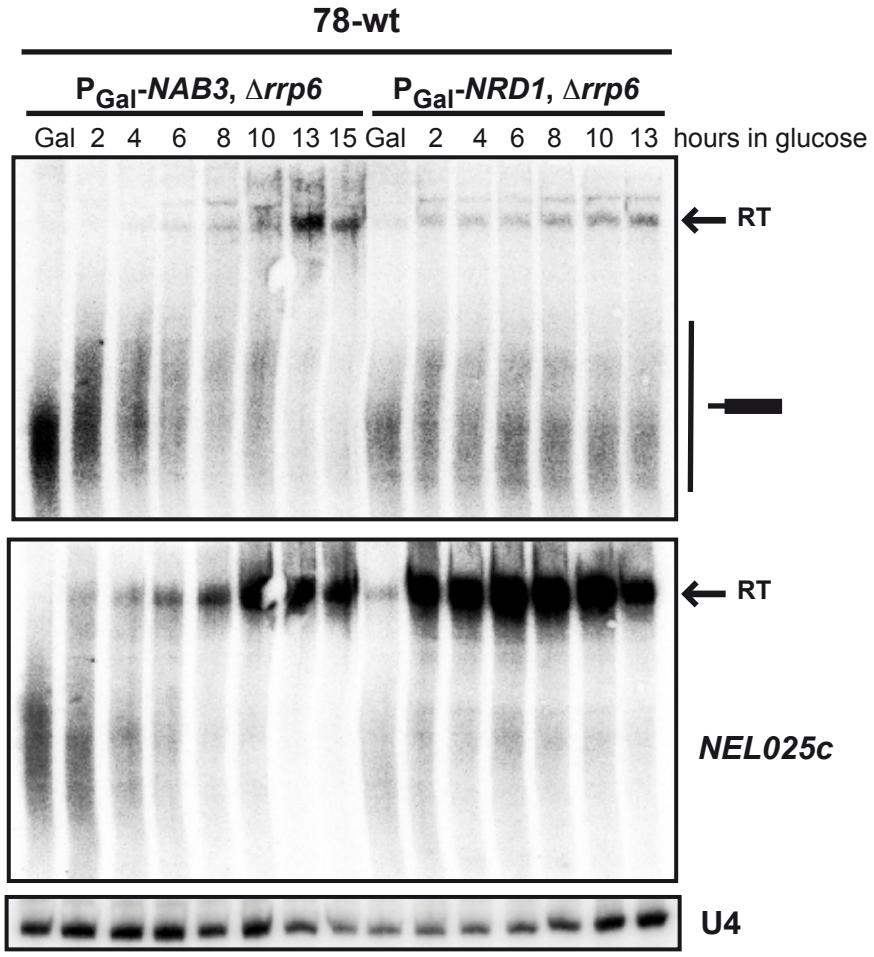


B

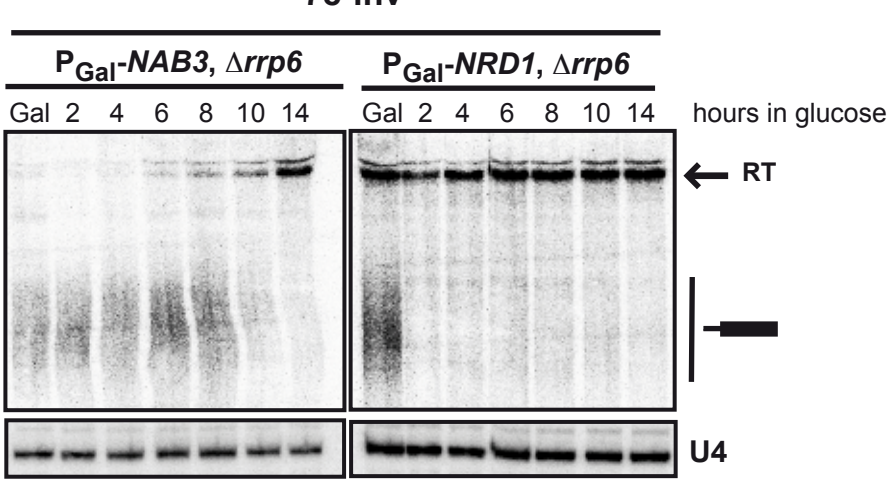


Porrua et al., Supplementary Fig. 5

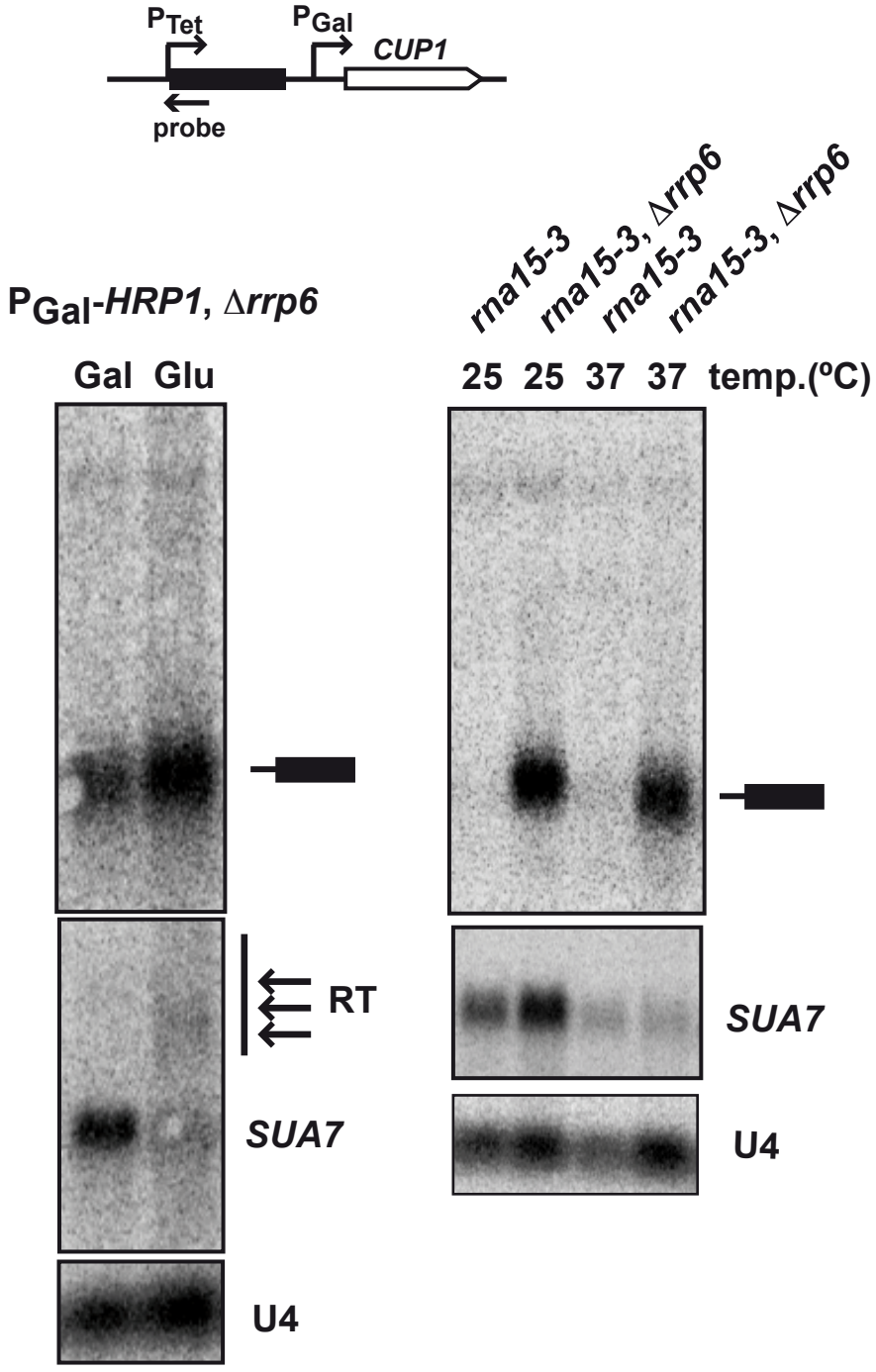
A



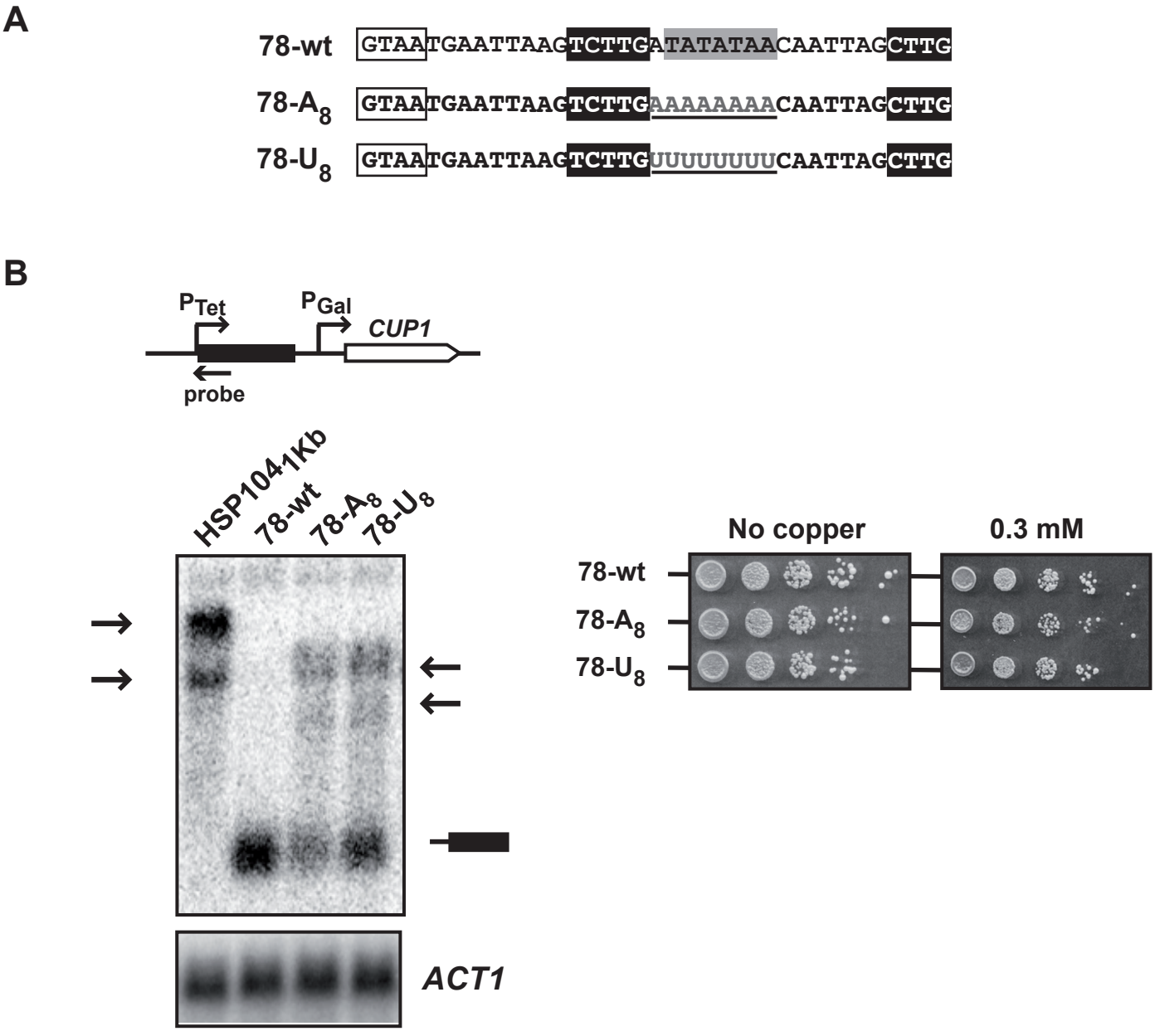
B



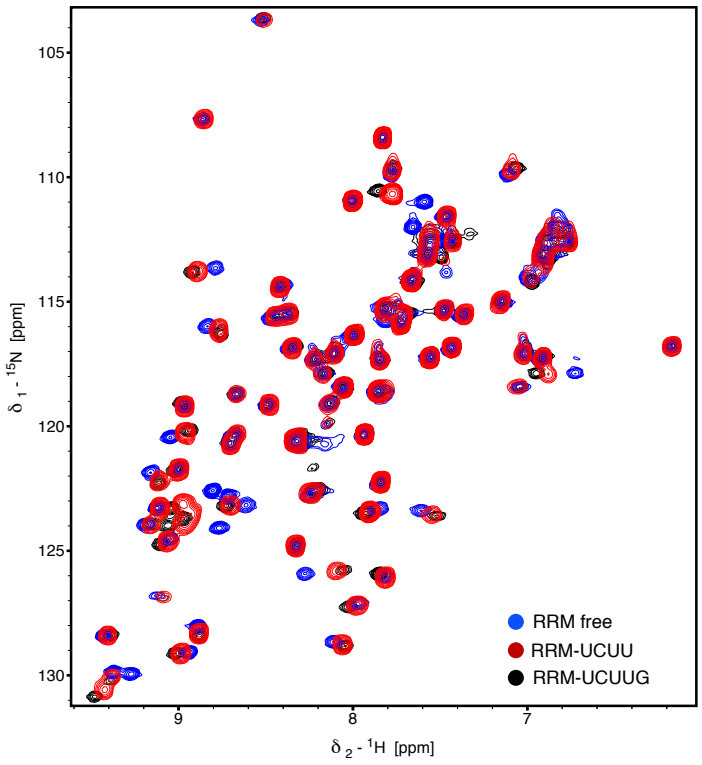
Porrua et al., Supplementary Fig. 6



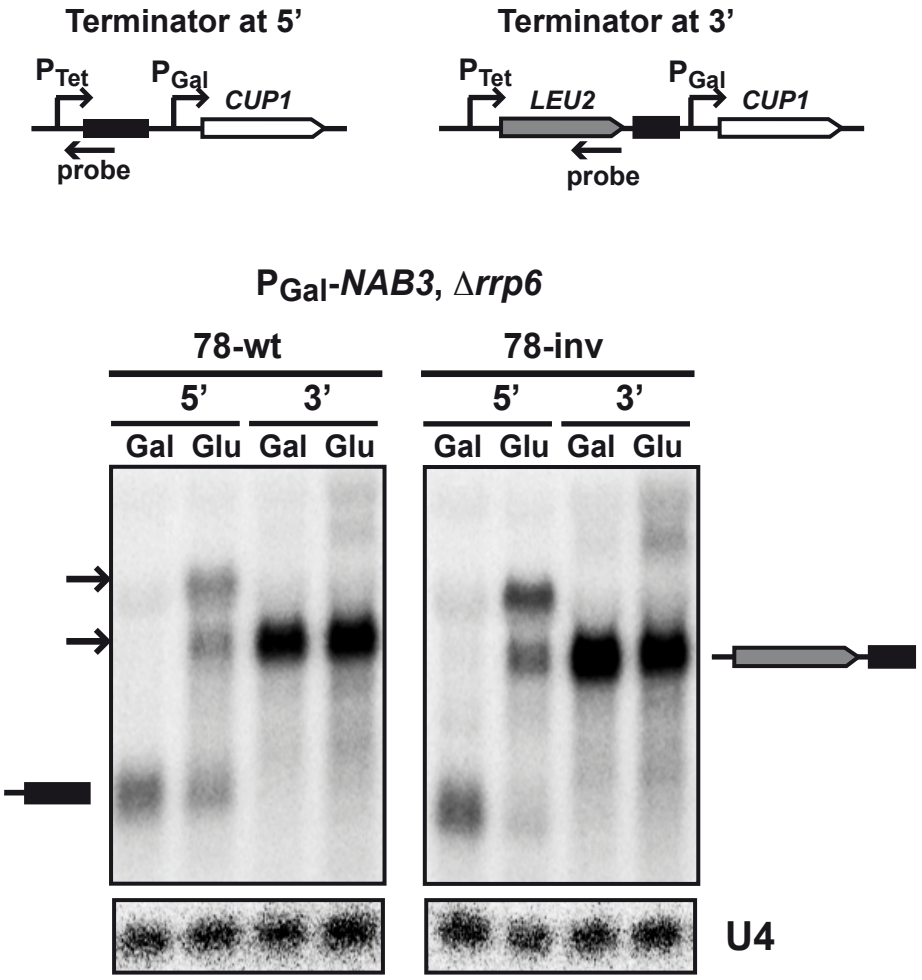
Porrua et al., Supplementary Fig. 7



Porrua et al., Supplementary Fig. 8



Porrua et al., Supplementary Fig. 9



Porrua et al., Supplementary Fig. 10