Supplementary material

Supplementary Materials and Methods

RNA and protein analysis

Yeast RNA extraction and northern analysis was performed as described (Tollervey, 1987), primer extension as described (Beltrame and Tollervey, 1992). High molecular weight RNA was separated on 1.2% glyoxal gels. Probes for northern and southern blots are described in Supplementary Table 2. Random primed probes were prepared using hexanucleotide mix (Roche) as per manufacturer's instructions, hybridized in "Neverfail" Hybridization Buffer (50% formamide, 5x SSC, 50mM Sodium phosphate pH6.8, 5x Denhardt's solution, 0.1% SDS, 20g/ml sonicated salmon sperm DNA) at 42°, and washed with 0.2x SSC 0.1% SDS at 55°. Probes for *YEL074W*, CEN3 and MAT were prepared using StripEasy (Ambion) and hybridized in UltraHyb (Ambion). RNA probes were prepared as described (Sambrook and Russell, 2001), hybridized in "Neverfail" buffer at 60° and washed with 0.2x SSC 0.1% SDS at 65°. Autoradiographs were acquired using a Storm phosphorimager (GE Healthcare) and quantified using ImageQuant 5.2 (GE Healthcare). Deadenylation of RNA was performed as described (LaCava et al., 2005), poly(A)+ RNA was isolated using a PolyATtract mRNA isolation system (Promega). 3' RACE was performed using SuperScript II (Invitrogen) according to manufacturer's instructions, cDNA synthesis was primed from ADAPT-dT (GACTCGAGTCGACATCGATTTTTTTTTTTTTTT), PCR was performed using ADAPT primer (GACTCGAGTCGACATCG). 5' RACE was performed using a GeneRacer kit (Invitrogen).

Protein isolation and western blotting were carried out by standard methods. Antibodies: goat anti-Act1 (Santa Cruz, 1:200), mouse anti-GFP (BD Biosciences 1:1,000), rabbit anti-goat (Sigma, 1:10,000), donkey anti-mouse (GE Healthcare, 1:10,000).

Pulsed field gel electrophoresis

$1x10^8$ cells were washed with 1ml YWB (10mM Tris pH 7.5, 50mM EDTA), resuspended in 50μl of YWB containing 1mg/ml Lyticase (Sigma). 55μl molten low melting point agarose (1.6% in dH_2O) was added, and the mixture was solidified in a

plug mould (Bio-Rad). Plugs were incubated for 30min at 30 in 500μl YWB containing 1mg/ml lyticase, then overnight at 55° in 500μl PKB (100mM EDTA pH8.0, 1% sodium lauryl sarcosine, 0.2% sodium deoxycholate, 1mg/ml Proteinase K (Roche)). Plugs were washed 2x 1 hour in YWB. 0.8% agarose (pulse field certified, Bio-Rad) gels were run in 1xTBE for 68 hours, 300-900s switch time, 3.0V/cm, 10°C in CHEF II system (Bio-Rad). Gels were stained with ethidium bromide and imaged before soaking sequentially with 0.25M HCl for 20 min, 0.5M NaOH for 20 min, 0.5M Tris HCl pH7.5 / 1.5M NaCl for 30min and transfer for 24 hours in 20xSSC. Blots were probed with a random primed probe to 18S as described for northern blots. PFGE ladders were *S. cerevisiae* (New England Biolabs) and *H. wingii* (Bio-Rad).

Chromatin Immunoprecipitation

ChIP was carried out as described previously (Kotovic et al., 2003). IPs were performed with anti-Myc mAb 9E10 (Santa Cruz), anti-RNA polymerase II mAb 8WG16 (Covance), or non-immune mouse IgG (Sigma) followed by GammaBind Sepharose beads (GE Healthcare).

qPCR was performed with SYBR green JumpStart Taq ReadyMix (Sigma) and a Stratagene MX3005P real-time PCR machine. Reactions volumes were 10μL, 1μL template and 100-300nM of each primer, depending on individual optimization. Cycling parameters were 95°C 5 min, then 40 cycles of 95°C 10 s, 55°C 10 s, 72°C 15 s, measurements were taken after the extension step. Primer sequences are listed in Supplementary Material. Values were determined using $2^{-Ct\ IP}/2^{-Ct\ background}$, where Ct IP is the cycle number for immunoprecipitate and Ct background is the cycle number for non-immune IgG precipitate. Ct values over the rDNA region were normalized to the median Ct of each experiment (method to be described in a forthcoming publication), and the mean of 3 independent experiments is shown. Error bars reflect standard error of the 3 experiments.

Supplementary Tables

Table S1: Yeast strains used in these analyses

BY4741: MAT<u>a</u>; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0

trf4Δ: MAT<u>a</u>; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; trf4Δ::natMX6

trf5Δ: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; trf5Δ::kanMX6

trf4Δ, GAL::trf5: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; trf4Δ::kanMX4; HisMX6-pGAL-3HA::trf5

rrp6Δ: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; rrp6Δ::natMX

air1Δ air2Δ: MATa ; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; air1Δ::kanMX4 ; air2Δ::natMX6

sir2Δ: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; sir2Δ::hygMX6

top1Δ: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; top1Δ::kanMX6

sir2Δ, trf4Δ: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; sir2Δ::hygMX6; trf4Δ::natMX6

sir2Δ, trf5Δ: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; sir2Δ::hygMX6; trf5Δ::kanMX6

top1Δ, trf4Δ: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; top1Δ::kanMX6; trf4Δ::natMX6

top1Δ, trf5Δ: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; top1Δ::natMX6; trf5Δ::kanMX6

mtr3-1: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; mtr3-1-hygMX6

mtr3-1, trf4: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; mtr3-1-hygMX6; trf4Δ::natMX6

pap1-2: MATa *ade2 his3 trp1 ura3 leu2 LEU2::pap1 (pap1-2)*

pap1-2 trf4Δ: as *pap1-2* but with *trf4Δ::natMX6*

SMC1-MYC: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; SMC1-13MYC-HygMX6

SMC1-MYC trf4Δ: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; trf4Δ::natMX6; SMC1-13MYC-HygMX6

SMC1-MYC top1Δ: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; top1Δ::kanMX6; SMC1-13MYC-HygMX6

SMC1-MYC top1Δ, trf4Δ: MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; top1Δ::kanMX6; trf4Δ::natMX6; SMC1-13MYC-HygMX6

nrd1-102: MATα ura3Δ0 his3Δ1 leu2Δ0 met15Δ0 *LYS2 nrd1Δ::KAN* (pJC720 [*LEU2 nrd1-102*])

*nab3-11: MAT*a *ade2 can1-100 his3-11,15 leu2-3,-112 trp1-1 ura3-1 nab3-11*

Table S2. Plasmids used in this work

p-: pNOPPATA1L (Hellmuth et al., 1998)

pTrf4: pNOPPATA1L Trf4 (Vanacova et al., 2005)

pTrf4 DXD: pNOPPATA1L Trf4 DXD (Vanacova et al., 2005)

pMet25p-GFP: The *MET25* ORF was amplified from BMA38 genomic DNA using MET25 F1 and MET25 R1 and cloned into pGFP-C-Fus (Niedenthal et al., 1996) via *Spe*I and *Bam*HI sites.

Table S3. Hybridization probes used in this work

Major RNA species detected are indicated in brackets

403 (*PGK1*): ACCGTTTGGTCTACCCAAGTGAGAAGCCAAGACA

499 (*TSA1*): GGAGTATTCGGAGTCAGTGGAGGCGAAAAGAACT

250 (*SCR1*): ATCCCGGCCGCCTCCATCAC

Act1 (*ACT1*): PCR product of ACT1 F1 and ACT1 R1 labeled by random priming

18S (Chromosome XII): PCR product of 18S F1 and 18S R1 labeled by random priming

NTS1 (IGS1 region): PCR product of NTS1 F1 and NTS1 R1 labeled by random priming

YEL074W (*YEL074W* ORF): PCR product of YEL074W F1 and YEL074W R1 labeled by random priming

MAT (all three MAT loci): PCR product of HMLalpha1 R1 and HMLalpha2 R1 labeled by random priming

CEN3 (centromere chromosome III): PCR product of CEN3 F1 and CEN3 R1 labeled by random priming

NTS1 *Hpa*I (IGS1-R): PCR product of NTS1 F1 and NTS1 R1 cloned in forward orientation into pGEM-T Easy, linearized with *Hpa*I and transcribed with T7 RNA polymerase

NTS1 short (IGS1-R): PCR product of NTS1 R3 and KK NTS1-3 R cloned in forward orientation into pGEM-T Easy, linearized with *Sal*I and transcribed with T7 RNA polymerase

NTS1 R (IGS1-F): PCR product of NTS1 F1 and NTS1 R1 cloned in forward orientation into pGEM-T Easy, linearized with *Hpa*II and transcribed with SP6 RNA polymerase

NTS2 F (IGS2-R): PCR product of NTS2 F1 and NTS2 R1 cloned in forward orientation into pGEM-T Easy, linearized with *Sal*I and transcribed with T7 RNA polymerase

NTS2 R (IGS1-F): PCR product of NTS2 F1 and NTS2 R1 cloned in forward orientation into pGEM-T Easy, linearized with *Nco*I and transcribed with SP6 RNA polymerase

Table S4. ChIP primer sets used in this work

1: KK 25S1 F: ACGCTTACCGAATTCTGCTT
KK 25S1 R: CGTTCATAGCGACATTGCTT

2: KK NTS1-2-F: ACTCATGTTTGCCGCTCTG
KK NTS1-2-R: TGCAAAGATGGGTTGAAAGA

3: KK NTS1-1 F: TACACCCTCGTTTAGTTGCTTCT

KK NTS1-1 R: CGGTATGCGGAGTTGTAAGA

4: KK 5S F: ACCTGCGTTTCCGTTAACT

KK 5S R: AGTTGATCGGACGGGAAAC

5: KK NTS2-2 F: GCCACCATCCATTTGTCTTT

KK NTS2-2 R: TGAAAGTTGGTCGGTAGGTG

6: KK NTS2-1 F: GGGTAACCCAGTTCCTCACTA

KK NTS2-1 R: GCATATATTTCTTGTGTGAGAAAGG

7: KK 35S F: CCACGATGAGACTGTTCAGGT

KK 35S R: GTCGCTAGGTGATCGTCAGA

Table S5. Other oligonucleotides used in this work

SIR2 F1 UP45: AGACACATTCAAACCATTTTTCCCTCATCGGCACATTAAAGCTGG CGGATCCCCGGGTTAATTAA

SIR2 R1 DN45: ATTGATATTAATTTGGCACTTTTAAATTATTAAATTGCCTTCTACG AATTCGAGCTCGTTTAAAC

TOP1 F1 UP45: AAAATCTAAAGGGAGGGCAGAGCTCGAAACTTGAAACGCGTAA AACGGATCCCCGGGTTAATTAA

TOP1 R1 DN45: ATGCGTGAATGTATTTGCTTCTCCCCTATGCTGCGTTTCTTTGC GGAATTCGAGCTCGTTTAAAC

MTR3-1 F1 UP45: ACTGTCGTATCATTCATCAAAAATGGAGAGGAAATTGTAGGT TTTCTAGAAGGATGACGGATCCCCGGGTTAATTAA

MTR3 R1 DN45: TCTTAACCTCTTTCACTTTTCTTTCCGACTAGGATAATCTCGTG TGAATTCGAGCTCGTTTAAAC

SMC1 F1 UP45: AACTCGTCGAAGATCATAACTTTGGACTTGAGCAATTACGCAG AACGGATCCCCGGGTTAATTAA

SMC1 R1 DN45: ATTAGTTATTTGACGGGTTATAGCAGAGGTTGGTTTCATAGATT AGAATTCGAGCTCGTTTAAAC

ACT1 F1: GGTATGTGTAAAGCCGGTTTTGC

ACT1 R1: GGACCACTTTCGTCGTATTCTTG

NTS1 F1: AAGTGGGTACTGGCAGGAGCCGGGG

NTS1 R1: TTCTAGTTTCTTGGCTTCCTATGCT

NTS1 R2: TGTCCCCACTGTTCACTGTTCACTG

NTS2 F1: GATAGTTTAACGGAAACGCAGGTGA

NTS2 R1: GAAGTACCTCCCAACTACTTTTCCT

NTS1 F3: GTGAACAGTGAACAGTGAACAGTG

NTS1 R3: GACCAAATAGTAAATAGTAACTTACATACATTAG

AEH NTS1-3 R: TGTTGTTACGATCTGCTGAG

18S F1: CGGTAATTCCAGCTCCAATAGC

18S R1: GCAATACGCCTGCTTTGAACAC

NTS2 MET25 UP45: ACGATGAAAAAGGTGATTTGTCATTTACAAGAGGTAGGTCG

AAACGATGCAAGGGTTCGAATCC

NTS2 MET25 DN45: TGAAACTACCTCTGCATGCCACCTACCGACCAACTTTCATG

TTCTCGACTCACTATAGGGCGAATTG

MET25 F1: ACTAGTATGCCATCTCATTTCGATACTG

MET25 R1: GGATCCTGGTTTTTGGCCAGCG

CEN3 F1: CAAAAATCACAGTGAGCAACTTG

CEN3 R1: GCTCTTCTATACTACTCCACATTTTG

HMalpha1 R1: ACATCTTCCCAATATCCGTC

HMLalpha2 R1: CATTCTTTCTTCTTTGCCAG

YEL074W F1: ATTCTGTTGGATACTTTAAACATGG

YEL074W R1: GTATTCATACTTCAAGTCAAAGTG

Supplementary References

Arigo, J.T., Eyler, D.E., Carroll, K.L. and Corden, J.L. (2006) Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol Cell*, **23**, 841-851.

Beltrame, M. and Tollervey, D. (1992) Identification and functional analysis of two U3 binding sites on yeast pre-ribosomal RNA. *Embo J*, **11**, 1531-1542.

Christman, M.F., Dietrich, F.S., Levin, N.A., Sadoff, B.U. and Fink, G.R. (1993) The rRNA-encoding DNA array has an altered structure in topoisomerase I mutants of Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A*, **90**, 7637-7641.

Gotta, M., Strahl-Bolsinger, S., Renauld, H., Laroche, T., Kennedy, B.K., Grunstein, M. and Gasser, S.M. (1997) Localization of Sir2p: the nucleolus as a compartment for silent information regulators. *Embo J*, **16**, 3243-3255.

Hellmuth, K., Lau, D.M., Bischoff, F.R., Kunzler, M., Hurt, E. and Simos, G. (1998) Yeast Los1p has properties of an exportin-like nucleocytoplasmic transport factor for tRNA. *Mol Cell Biol*, **18**, 6374-6386.

Huang, J. and Moazed, D. (2003) Association of the RENT complex with nontranscribed and coding regions of rDNA and a regional requirement for the

replication fork block protein Fob1 in rDNA silencing. *Genes Dev*, **17**, 2162-2176.
Kotovic, K.M., Lockshon, D., Boric, L. and Neugebauer, K.M. (2003) Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. *Mol Cell Biol*, **23**, 5768-5779.
LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A. and Tollervey, D. (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell*, **121**, 713-724.
Niedenthal, R.K., Riles, L., Johnston, M. and Hegemann, J.H. (1996) Green fluorescent protein as a marker for gene expression and subcellular localization in budding yeast. *Yeast*, **12**, 773-786.
Sambrook, J. and Russell, D.W. (2001) *Molecular cloning : a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
Tollervey, D. (1987) A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA. *Embo J*, **6**, 4169-4175.
Vanacova, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G. and Keller, W. (2005) A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol*, **3**, e189.
Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.C., Dufour, M.E., Boulay, J., Regnault, B., Devaux, F., Namane, A., Seraphin, B., Libri, D. and Jacquier, A. (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell*, **121**, 725-737.

Supplementary Figures

Supplemental Figure 1. RNase H analysis of IGS1-R

A: Schematic of *IGS1-R* showing poly(A) regions. **B**: RNA from indicated strains grown at 25°C in YPD was incubated with RNase H with or without oligo(dT). Probes are Nts1 *Hpa*I and *TSA1*. Oligo(dT) directed cleavage removes poly(A) tails, but cleavages in IGS1 R also occur at the two genome encoded poly(A) tracts marked in A. The IGS1 R transcript is therefore split into 3 regions, labeled 1-3 in A. The NTS1 *Hpa*I probe detects only two of these (1 and 2). The cleaved products of the IGS1 R transcript are therefore a discrete band derived from region 2 (between the two genome encoded poly(A) tracts) of 270bp, a smear derived from transcripts whose end lies between the two genome encoded poly(A) tracts, and a smear derived from transcripts whose 3' ends occur after both genome encoded poly(A) sites. This complicated fragment pattern does not allow a detailed analysis, but is sufficient to show that smeared products still exist after oligo(dT) cleavage, and that therefore the IGS1-R transcripts do not a have a single well defined 3' end.

Supplemental Figure 2. Effect of *trf5*Δ and *rrp6*Δ on rDNA

A: PFGE analysis of indicated mutants probed with 18S to highlight chromosome XII, and stained with ethidium. **B**: PFGE analysis of *rrp6*Δ mutants probed with 18S to highlight chromosome XII, and stained with ethidium bromide. The different pattern of chromosome XII in this gel compared Sup. Fig. 2A is due to cells being harvested at log rather than stationary phase. This has been shown to alter migration of the rDNA in *top1*Δ strains (Christman et al., 1993). **C**: Northern analysis of transcripts in mutants from A, probed with (top to bottom): NTS2 F, NTS1 R, NTS1 F, *ACT1*.

Supplemental Figure 3. Sir2 binding is unaltered in *rpb1-1*

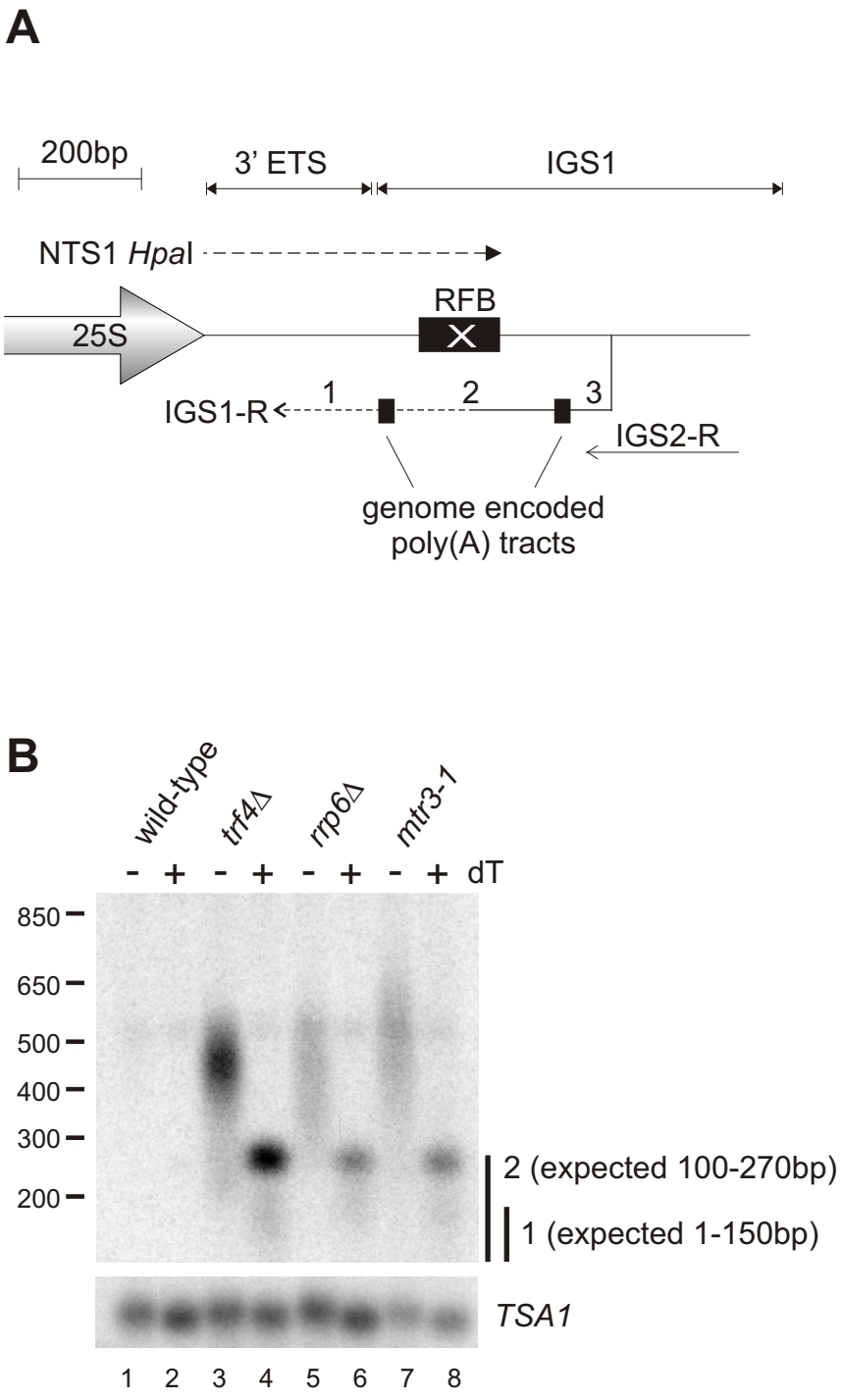
ChIP lysates from Fig. 7A & B were assessed for Sir2 distribution using anti-sir2 (Gotta et al., 1997). Sir2 binds to the rDNA at the IGS1 R locus via Fob1 (Huang and Moazed, 2003), and its binding was little altered in the *rpb1-1* strain at 37°C versus 25°C.

Supplemental Figure 4. Estimate of minimum number of CUTs

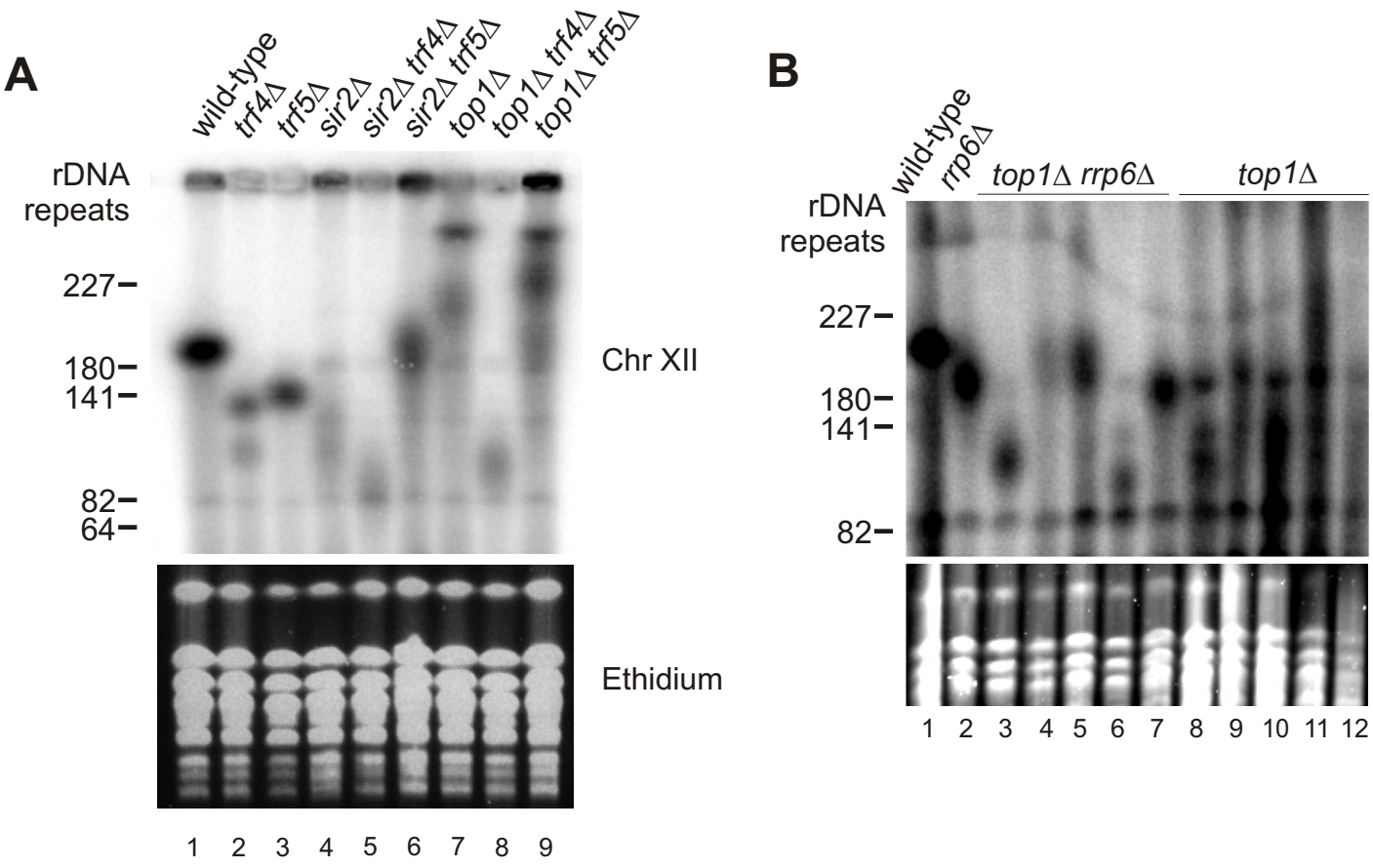
Affymetrix microarrays of three wild-type samples and two *rrp6* samples are described in (Wyers et al., 2005). We hypothesized that probes to genomic regions that show an increased signal in *rrp6* and do not correspond to an annotated gene are likely to represent CUTs. 434 probes show a significant increase ($p<0.05$ by Student's tTest) in *rrp6* and do not correspond to known genes, however, 26 have been excluded as the probes are to adjacent features and may hybridize to the same CUT. We have included probes to dubious and uncharacterized ORFs in our count of potential CUTs, along with features such as pseudogenes and retroviral long terminal repeats. Intergenic regions are not fully covered in these microarrays, which needs to be included in the approximation. There is 6.6Mb intergenic space in the genome (including both strands), and the microarrays contain 2989 strand-specific intergenic probes. If each CUT is estimated to be 1kb long, the intergenic region probes would cover just under half of the intergenic region for detection of CUTs, and therefore there are likely to be double the number of intergenic CUTs than this experiment detects. This gives a final estimate of 598 CUTs in the *S. cerevisiae* genome.

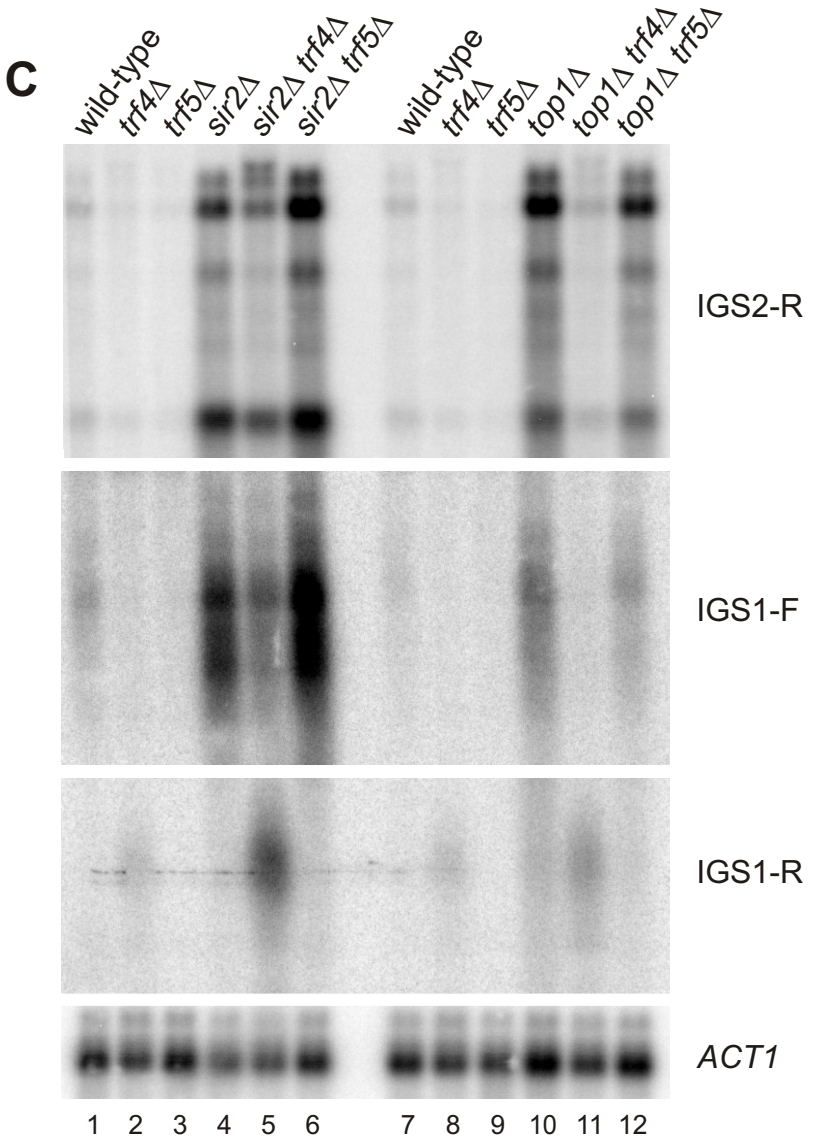
Some of the uncharacterized and dubious ORFs up-regulated in *rrp6* will of course be

genuine, however, CUTs have also been described on the opposite strand to annotated genes (Arigo et al., 2006) and these have not been accounted for in this analysis. We therefore suggest that this value of 598 CUTs represents a realistic minimum in the *S. cerevisiae* genome.

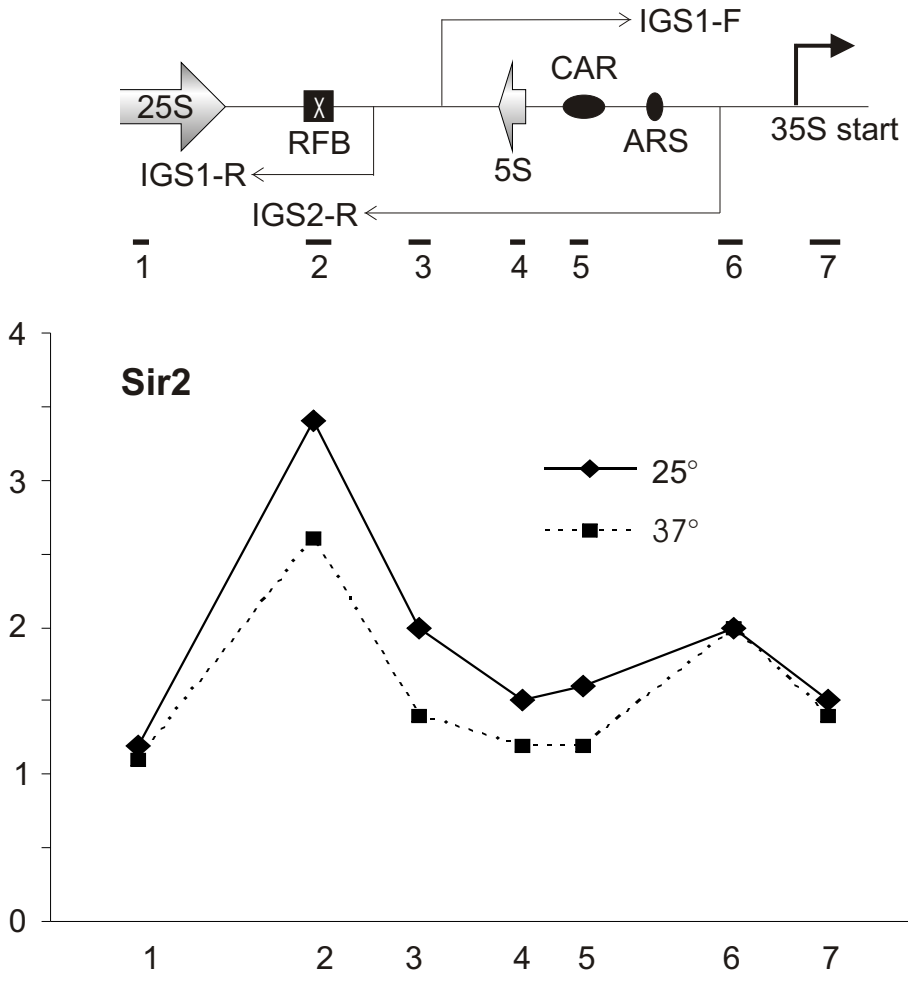


Houseley *et al.* Sup. Fig. 1

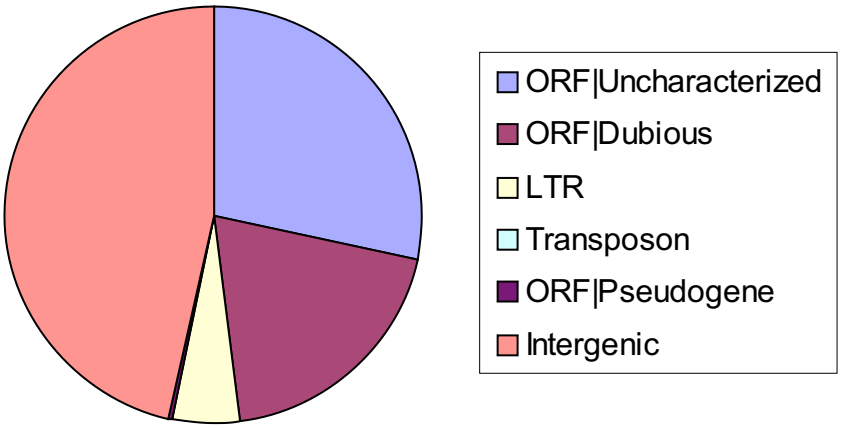




Houseley *et al.* Sup. Fig. 2



Houseley *et al.* Sup. Fig. 3



Feature type	CUTs
ORF|Uncharacterized	116
ORF|Dubious	80
LTR	20
Transposon	1
ORF|Pseudogene	1
Intergenic	190
Total	408

Houseley *et al.* Sup. Fig. 4