SUPPLEMENTAL DATA

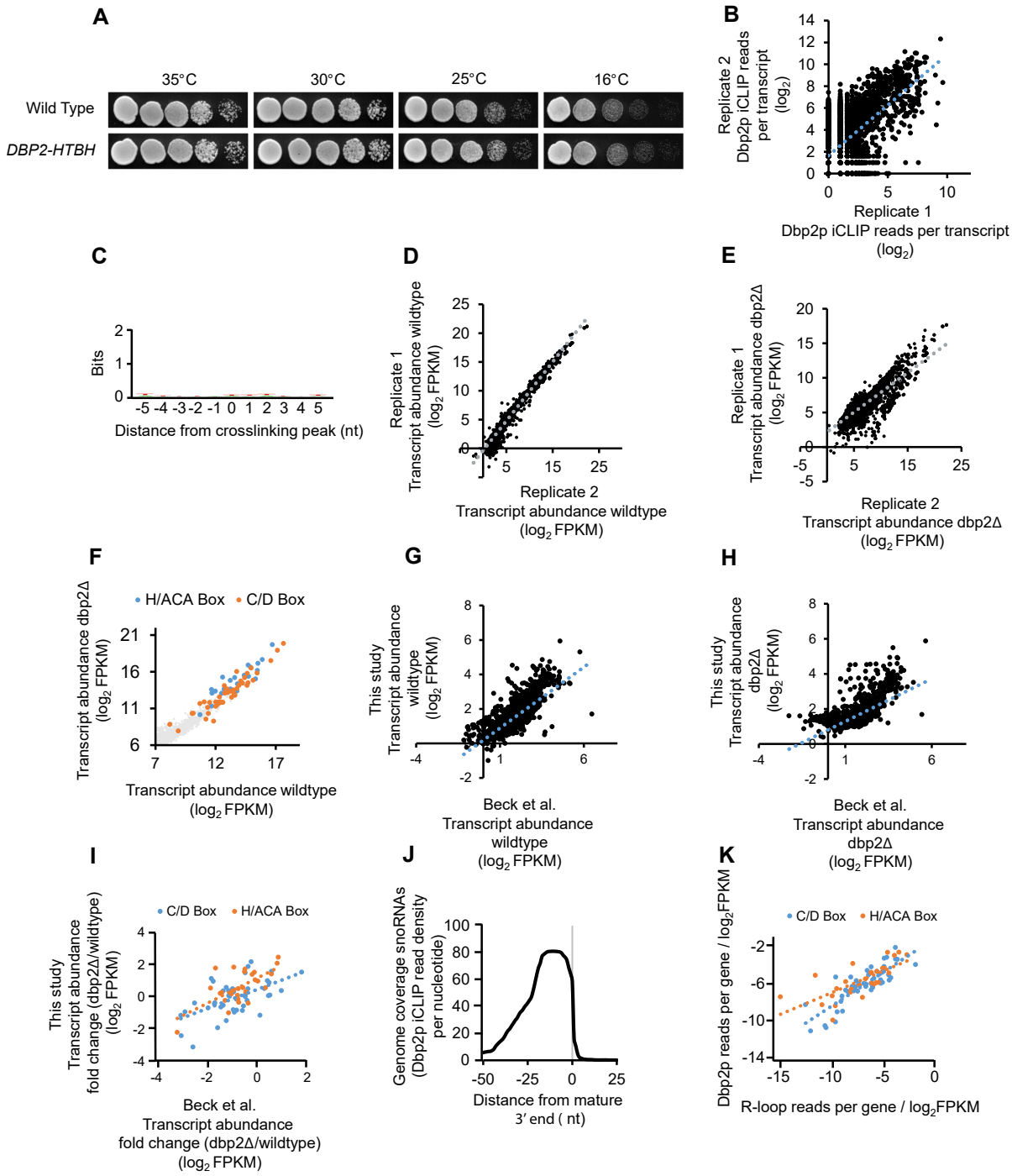
The DEAD-box protein Dbp2p is linked to non-coding RNAs, the helicase Sen1p, and R-loops

Frank A. Tedeschi, Sara C. Cloutier, Elizabeth J. Tran & Eckhard Jankowsky

---

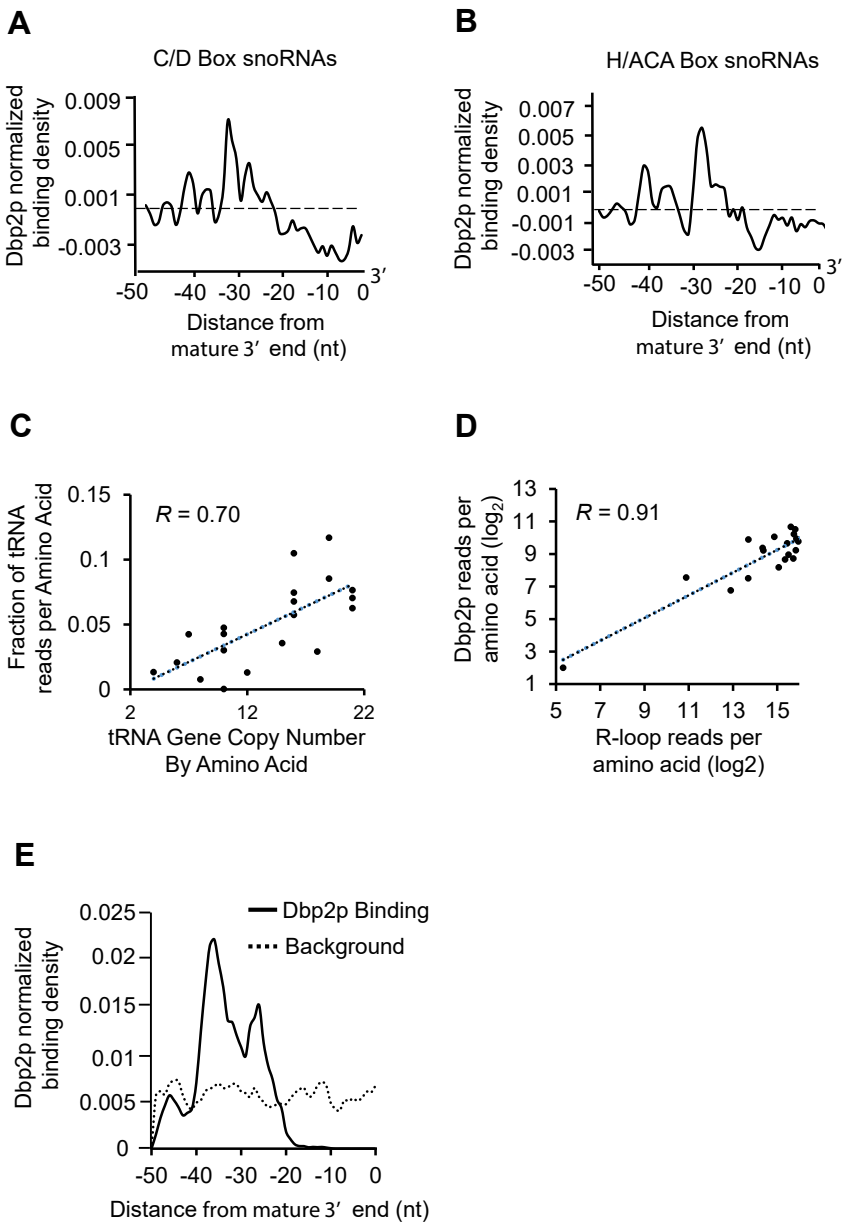
Supplemental Figures 1 - 3

Supplemental Tables T1-T4

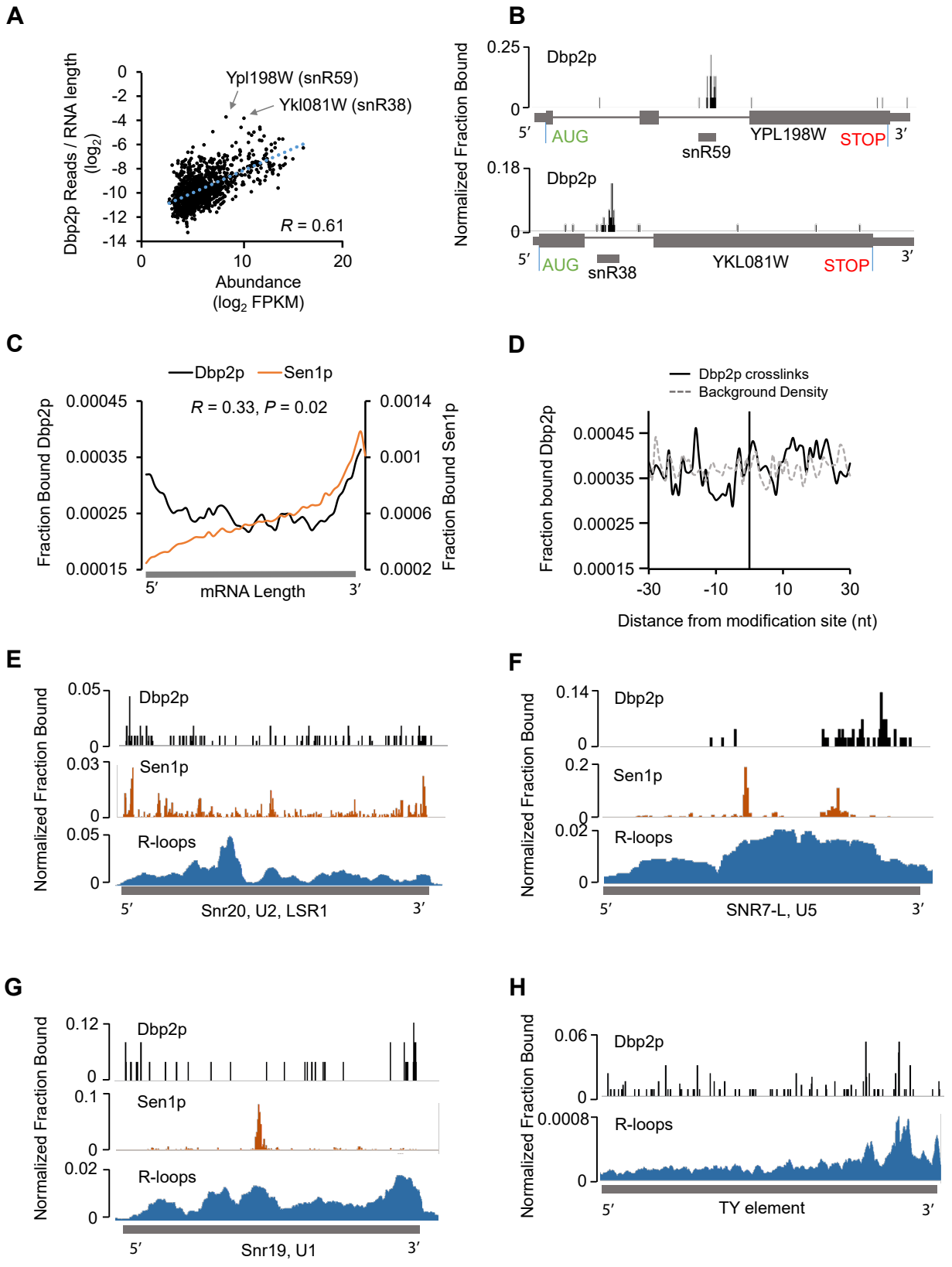


Supplementary Figure S1. ***(A)*** Serial dilution of wild-type and *DBP2-HTBH* strains at different temperatures. No significant differences between the growth of both strains was detected. ***(B)*** Correlation between two biological iCLIP replicates for Dbp2p-HTBH, plotted as the $\log_2$iCLIP reads per transcript. Pearson correlation coefficient (R) between the datasets: R = 0.65, n = 5,321. ***(C)*** Absence of sequence signature for all Dbp2p binding sites. Binding site peaks were determined using Piranha Peakfinder software and used as input to generate the sequence logo (https://weblogo.berkeley.edu/logo.cgi). ***(D)*** Correlation between two biological RNA-seq replicates for the wildtype strain, plotted as the $\log_2$FPKM per transcript (R = 0.98, n = 6,605). ***(E)*** Correlation between two biological RNA-seq replicates for *dbp2*Δ

plotted as log_2FPKM per transcript (*R* = 0.87, n = 6,635). ***(F)*** Correlation of transcript abundance for C/D-box (orange) and H/ACA-box (blue) snoRNA between *dbp2*Δ and wildtype strains (log_2FPKM). For comparison, mRNA transcripts are shown in grey (n = 4,389). ***(G)*** Correlation between transcript abundances (log_2FPKM) of previously published wildtype RNA-seq data (Beck et al. 2014) and RNA-seq generated by this study (*R* = 0.79, n = 1,938). ***(H)***, Correlation between transcript abundances (log_2FPKM) of previously published *dbp2*Δ RNA-seq data (Beck et al. 2014) and RNA-seq generated by this study (*R* = 0.62, n = 1,955). ***(I)*** Correlation between snoRNA abundances (log_2FPKM) of previously published *dbp2*Δ RNA-seq data (Beck et al. 2014) and RNA-seq generated by this study (C/D Box: *R* = 0.58, n = 44, H/ACA Box: *R* = 0.67, n =28). ***(J)*** Average binding density of Dbp2p per nucleotide surrounding the mature 3' end across all snoRNAs. Genome coverage was calculated using full length reads to emphasize the 3' end of the reads (n = 72). ***(K)*** Correlation between R-loop reads (*rnh1*Δ*rnh201*ΔS1-DRIP-Seq) and Dbp2p iCLIP reads for C/D-box (*R* = 0.66, n = 41, only transcripts with a minimum read count of 3 were considered) and H/ACA-box (*R* = 0.65, n = 26, only transcripts with a minimum read count of 3 were considered) snoRNA transcripts, normalize to RNA length and expression level.



Supplementary Figure S2. ***(A)*** Average binding density of Dbp2p per nucleotide surrounding the mature 3' end for all C/D-box snoRNAs (n = 44). Genome coverage was calculated using the 5' crosslink site. Background signal (dashed lined) was generated by randomizing biding sites locations within snoRNA three times, calculating the mean and subtracting the signal. ***(B)*** Average binding density of Dbp2p per nucleotide surrounding the mature 3' end for all H/ACA-box snoRNAs (n = 28). Data were calculated and are presented as in panel (A). ***(C)*** Correlation between tRNA gene copy number (Helm, 2006) and sequencing reads per tRNA amino acid type, for all isoforms of a given amino acid (n = 283). ***(D)*** Correlation between Dbp2p iCLIP reads and R-loop reads(*rnh1*Δ*rnh201*ΔS1-DRIP-seq reads) per tRNA type (grouped by amino acid identification) (n = 283). ***(E)*** Binding density of Dbp2p per nucleotide surrounding the mature 3'-end of all tRNA transcripts (n = 283). The background signal (dashed lined) was generated by randomizing binding sites locations within tRNAs three times, calculating the mean and subtracting the signal.


A
Dbp2p Reads / RNA length (log2)
Ypl198W (snR59)
Ykl081W (snR38)
R = 0.61
Abundance (log2 FPKM)
B
Normalized Fraction Bound
Dbp2p
AUG
YPL198W
STOP
snR59
YKL081W
snR38
C
Dbp2p
Sen1p
R = 0.33, P = 0.02
Fraction Bound Dbp2p
Fraction Bound Sen1p
mRNA Length
D
Dbp2p crosslinks
Background Density
Fraction bound Dbp2p
Distance from modification site (nt)
E
Normalized Fraction Bound
R-loops
Snr20, U2, LSR1
F
SNR7-L, U5
G
Snr19, U1
H
TY element

Supplementary Figure S3. ***(A)*** Correlation between Dbp2p iCLIP reads per transcript, normalized for RNA length and transcript abundance (RNA-seq of wildtype, n = 1,103). The highlighted transcripts with preferred Dbp2p binding encode snoRNAs in their introns (detailed in panel ***B***). ***(B)*** Dbp2p iCLIP traces for representative examples of mRNAs with snoRNA encoded in introns. ***(C)*** Metagene plot for Dbp2p iCLIP density and Sen1p PAR-CLIP for mRNA transcripts. Data were generated and presented as in **Fig.2C.** ***(D)*** Binding density of Dbp2p (per nucleotide) surrounding rRNA modification (snoRNA binding) sites (n = 110) (Piekna-Przybylska et al. 2008). The background signal (dashed lined) was generated by randomizing binding site locations within tRNAs three times, calculating the mean and subtracting the signal. ***(E)*** Representative traces for Dbp2p iCLIP, Sen1p PAR-CLIP and R-loops (*rnh1*Δ*rnh201*ΔS1-DRIP-seq read density) on the U2 snRNA transcript. Data are presented as described in **Fig.2A,B**. ***(F)*** Representative traces for Dbp2p iCLIP, Sen1p PAR-CLIP and R-loops on the U5 snRNA transcript. Data are presented as described in **Fig.2A,B**. ***(G)*** Representative traces for Dbp2p iCLIP, Sen1p PAR-CLIP and R-loops on the U1 snRNA transcript. Data are presented as described in **Fig.2A,B**. ***(H)*** Representative traces for Dbp2p iCLIP and R-loops on the retrotransposon YDRCTy1-3transcript. Data are presented as described in **Fig.2A,B**. No Sen1p PAR-CLIP data are available for Ty-elements.

Additional Reference:

Helm M. 2006. Post-transcriptional nucleotide modification and alternative folding of RNA. Nucleic Acids Res. **34**: 721-733

Supplemental Table T1: DNA Oligonucleotides

Name	Purpose	DNA Oligonucleotide Sequence 5' -3'
X1	iCLIP, RT primer	/5Phos/**GTAGA**CCGCATCGTCCTCCCTCCCTATAGTGAGTCGTATTA/iSp18/CACTCA/iSp18/CCGCTGGAAGTGACTGACAC
X2	iCLIP, PCR1	AGGGAGGACGATGCGG
X3	iCLIP, PCR1	CCGCTGGAAGTGACTGACAC
X4	iCLIP, PCR2	CAAGCAGAAGACGGCATACGACCGCTGGAAGTGACTGACAC
X5	iCLIP, PCR2	AATGATACGGCGACCACCGACTATGGATACTTAGTCAGGGAGGACGATGCGG
X6	iCLIP, Seq	CTATGGATACTTAGTCAGGGAGGACGATGCGG
X7r	iCLIP, RT primer	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGC/iSp18/CACTCA/iSp18/TTCAGACGTGTGCTCTTCCGATCTATTGATGGTGCCTACAG
X8r	iCLIP, PCR	CAAGCAGAAGACGGCATACGAGAT**CGTCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCG
X9r	iCLIP, PCR	AATGATACGGCGACCACCGAGATCTACAC
X10r	iCLIP, 5' Adaptor	/5rApp/CTGTAGGCACCATCAAT/3ddC/
X11r	iCLIP, seq	ACACTCTTTCCCTACACGACGCTCTTCCGATCT

HIS R	Tag PCR	GTG GTGGTG GCC GAT CTT GAT TAG ACC TTG
HIS F	Tag PCR	CAC CACCAC TAA TGA GGC GCG CCA CTT C
DBP2-HTBH F	Recombination	GGT TAC GGC GGT AAC AGG CAG AGA GAT GGT GGC TGG GGT AAC AGA GGT CGT TCA AAC TAT CGG ATC CCC GGG TTA ATT AA
DBP2-HTBH R	Recombination	GAA ATA AGG TGC AGT CAA CTT ATA TAA TTA TTATTA ATA GAG ATG AAT GAA TTG AAT CAC TTT GAC TGA ATT CGA GCT CGT TTA AAC

Supplemental Table T2: RNA Oligonucleotides

Name	**Purpose**	**RNA Oligonucleotide Sequence 5' -3'**
X12	iCLIP, 5' Adaptor	/5Phos/GUGUCAGUCACUUCCAGCGG/3Puro/

Supplemental Table T3: Yeast Strains

Name	**Genotype**	**Source**
DBP2-HTBH (BTY140)	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 DBP2-HTBH:kanMX6*	This study
DBP2Δ (BTY115)	MATa dbp2::KanMx ura3Δ0 leu2Δ0 his3Δ0 TRP1 met-lys?	Cloutier et al, JBC 2012; 287, 26155-26166.

Supplemental Table T4: Plasmids

Name	Description	Source
pFA6a-HTB	pFA6a-HTB-kanMX6	Tagwerker et al, Yeast 2006; 23:623-632
pFA6a-HTBH (BTP16)	pFA6a-HTBH-kanMX6	This study

Barcodes in **bold**

Modifications:

/5Phos/ 5'-Phosphorylation

/3Puro/ 3'-Puromycin

/iSp18/ Int Spacer 18

/5rApp/ 5' Adenylation

/3ddC/ 3' Dideoxy-C