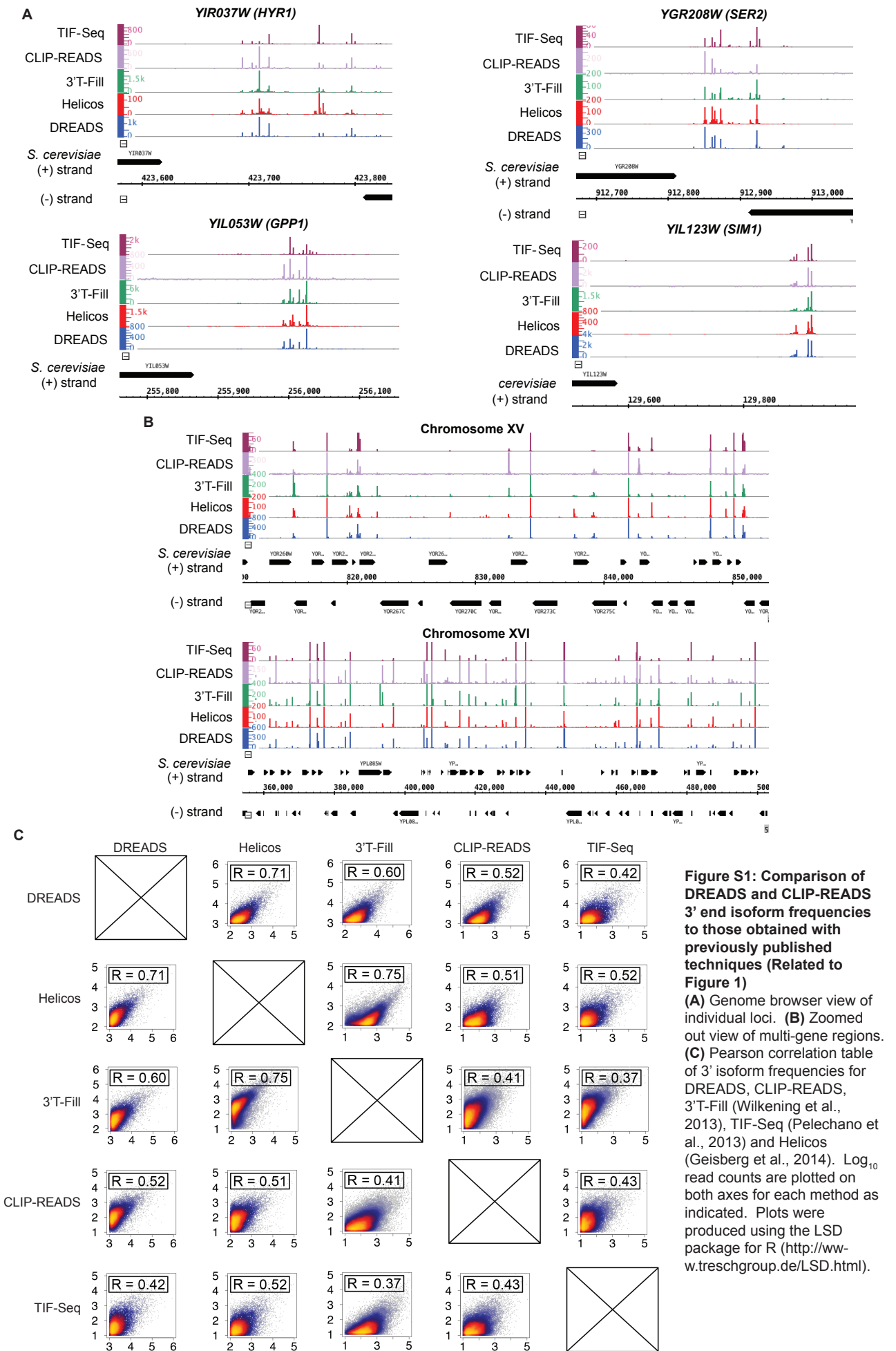


**Molecular Cell, Volume 72**

**Supplemental Information**

**Extensive Structural Differences of Closely  
Related 3' mRNA Isoforms: Links to Pab1 Binding  
and mRNA Stability**

**Zarmik Moqtaderi, Joseph V. Geisberg, and Kevin Struhl**



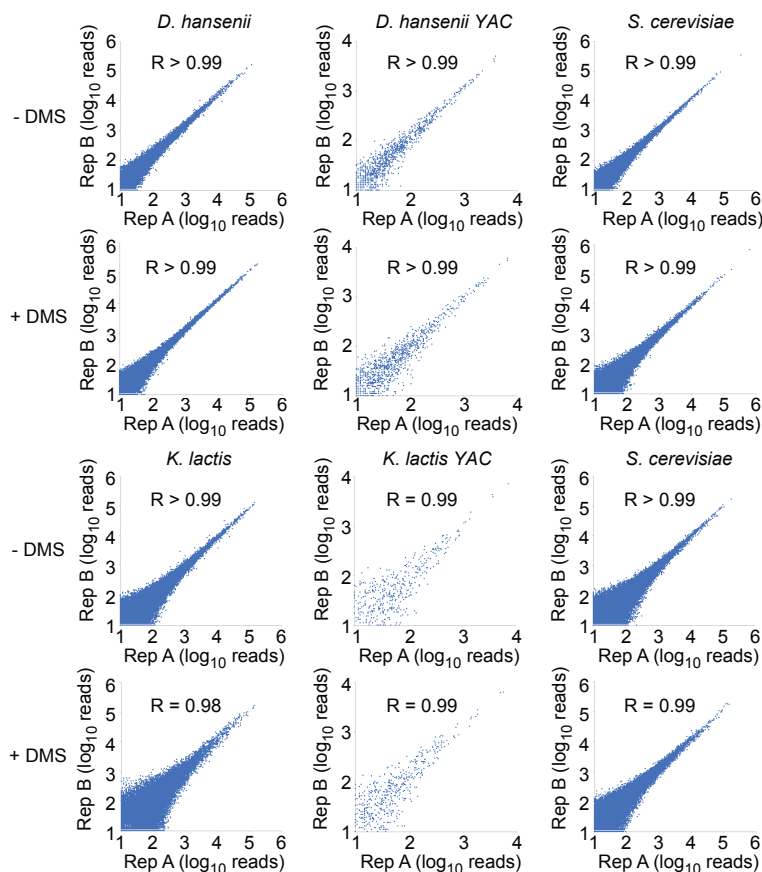
**Figure S1: Comparison of DREADS and CLIP-READS 3' end isoform frequencies to those obtained with previously published techniques (Related to Figure 1)**  
**(A)** Genome browser view of individual loci. **(B)** Zoomed out view of multi-gene regions. **(C)** Pearson correlation table of 3' isoform frequencies for DREADS, CLIP-READS, 3'T-Fill (Wilkening et al., 2013), TIF-Seq (Pelechano et al., 2013) and Helicos (Geisberg et al., 2014). Log<sub>10</sub> read counts are plotted on both axes for each method as indicated. Plots were produced using the LSD package for R (<http://www.treschgroup.de/LSD.html>).



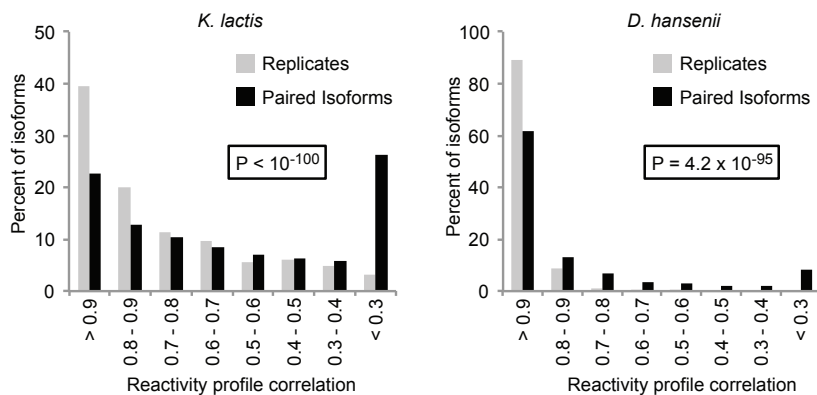
A

	Raw sequence pairs	<i>D. hansenii</i> or <i>K. lactis</i>		<i>S. cerevisiae</i> fully processed reads	
		fully processed reads	Percent	Percent	Percent
<i>D. hansenii</i> rep. A	27,058,154	14,658,184	54.17		
<i>D. hansenii</i> rep. B	36,196,762	20,401,488	56.36		
<i>D. hansenii</i> + DMS rep. A	31,893,005	14,905,516	46.74		
<i>D. hansenii</i> + DMS rep. B	40,802,565	18,878,032	46.27		
JYAC7 ( <i>S. cerevisiae</i> with <i>D. hansenii</i> YAC) rep. A	32,933,878	180,051	0.55	16,837,032	51.12
JYAC7 ( <i>S. cerevisiae</i> with <i>D. hansenii</i> YAC) rep. B	36,427,327	197,532	0.54	18,783,329	51.56
JYAC7 ( <i>S. cerevisiae</i> with <i>D. hansenii</i> YAC) + DMS rep. A	35,220,794	203,959	0.58	16,532,308	46.94
JYAC7 ( <i>S. cerevisiae</i> with <i>D. hansenii</i> YAC) + DMS rep. B	34,574,789	189,795	0.55	15,254,859	44.12
<i>K. lactis</i> rep. A	43,137,283	24,914,146	57.76		
<i>K. lactis</i> rep. B	49,330,899	20,560,482	41.68		
<i>K. lactis</i> + DMS rep. A	35,021,272	17,096,481	48.82		
<i>K. lactis</i> + DMS rep. B	43,757,730	20,185,901	46.13		
JYAC2 ( <i>S. cerevisiae</i> with <i>K. lactis</i> YAC) rep. A	42,712,965	98,452	0.23	20,743,271	48.56
JYAC2 ( <i>S. cerevisiae</i> with <i>K. lactis</i> YAC) rep. B	42,354,740	94,842	0.22	18,092,609	42.72
JYAC2 ( <i>S. cerevisiae</i> with <i>K. lactis</i> YAC) + DMS rep. A	42,858,718	100,051	0.23	18,125,863	42.29
JYAC2 ( <i>S. cerevisiae</i> with <i>K. lactis</i> YAC) + DMS rep. B	44,816,467	134,752	0.30	22,624,815	50.48

B

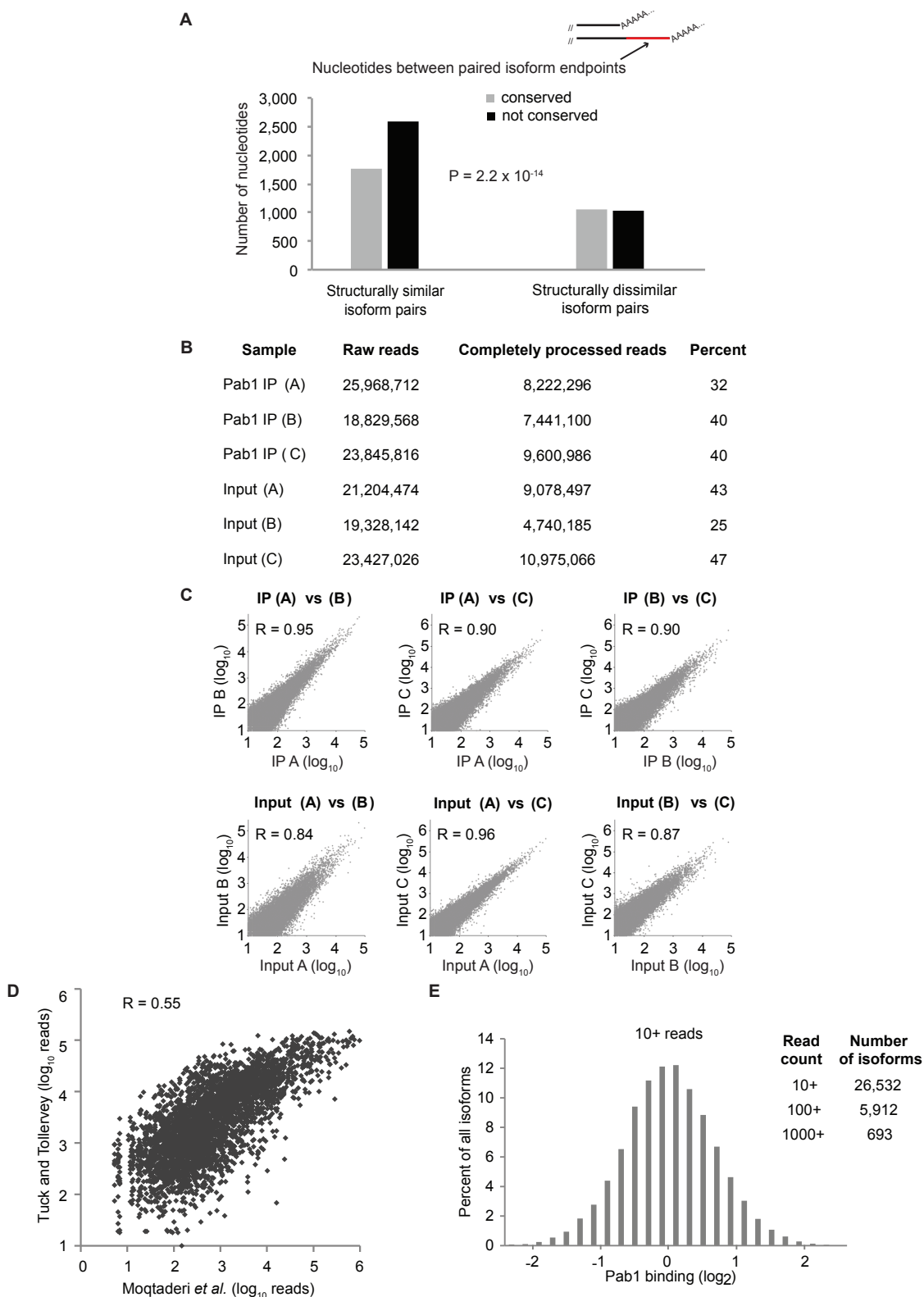


C



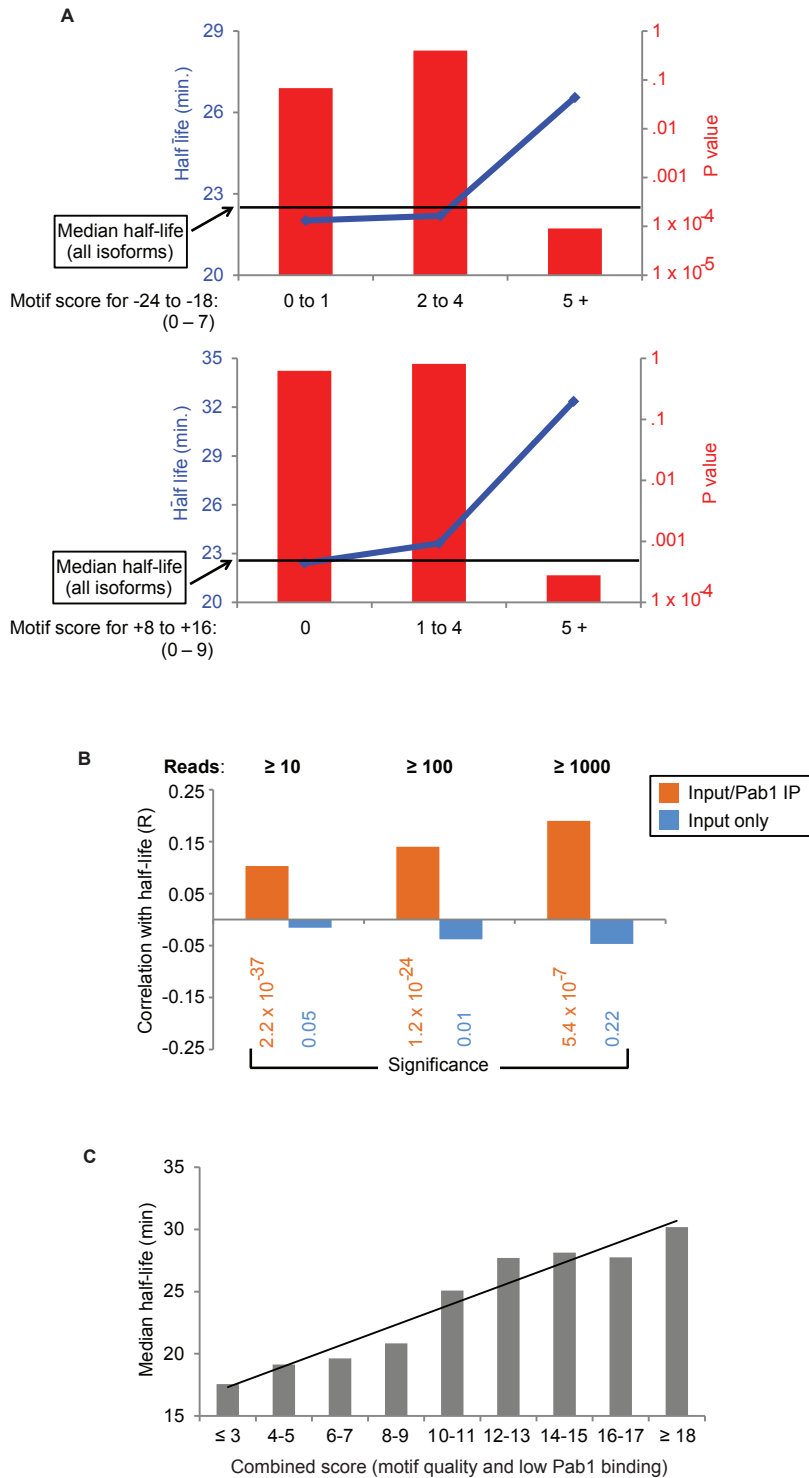
**Figure S3: DREADS sequencing statistics and experimental reproducibility of 3' isoforms in *D. hansenii*, *K. lactis*, and transplanted YACs (Related to Figure 4)**

(A) DREADS sequencing statistics for both DMS treated and untreated samples. (B) Correlation of 3' isoform frequencies in biological replicates of *D. hansenii*, *K. lactis*, *S. cerevisiae* YAC hosts, and the transplanted *D. hansenii* and *K. lactis* YACs. The overall number of unique 3' mRNA isoforms ( $\geq 10$  reads/isoform) ranges from  $\sim 20,000 - 60,000$  isoforms for *D. hansenii*, *K. lactis*, and *S. cerevisiae* YAC host strains to  $\sim 900 - 1,100$  isoforms for *D. hansenii* and *K. lactis* YACs. (C) Genome-wide percentile distribution of Pearson correlation coefficients for isoform reactivity profiles in *K. lactis* and *D. hansenii*. Grey bars: correlations for the same isoform's reactivity profiles in two biological replicates. Black bars: correlation of every isoform's reactivity profile with every other same-gene isoform's reactivity profile.



**Figure S4: Pab1 binding and evolutionary conservation (Related to Figure 5)**

(A) Nucleotides unique to the longer of two paired isoforms exhibit less evolutionary conservation when the two isoforms are structurally similar. All isoform pairs with endpoints spaced  $\leq 20$ nt apart were divided into one of two groups: those with similar folding (left,  $\Delta R < 0.1$ ) and those that are structurally dissimilar (right,  $\Delta R > 0.3$ ). Individual nucleotides within the regions unique to the longer isoforms were classified as either conserved (gray bars; PhastCons  $> 0.67$ ) or non-conserved (black bars; PhastCons  $< 0.33$ ). The proportion of nucleotides that is likely to be conserved is considerably greater in structurally dissimilar isoforms ( $P = 2.2 \times 10^{-14}$ ) than in those pairs that share common structure. (B) Pab1 IP and input sample (CLIP-READS) sequencing statistics. Completely processed reads represent reads that were uniquely aligned and had at least one non-genomically encoded A. (C) Correlation of biological replicates and inputs for Pab1 CLIP-READS. (D) Correlation of Pab1 binding dataset with (Tuck and Tollervey, 2013) dataset. Plotted values represent all ORFs common to both datasets (3,596) that had non-zero Pab1 reads and whose termination codons were spaced at  $\geq 200$  nt apart from other genomic features. (E) Distribution of relative isoform-specific Pab1 binding. Expression-corrected Pab1 binding was calculated for each isoform, and the median binding set to 1.



**Figure S5: Structural motifs linked to mRNA isoform stability (Related to Figure 6)**

(A) Top panel: Predicted unstructured region at -24 to -18 correlates with increased isoform stability. Number of isoforms in groups (L to R) is 2,078, 3,364 and 1,146. Bottom panel: Predicted double-stranded poly(A) region at +8 to +16 correlates with increased isoform stability. Number of isoforms in groups (L to R) is 6,211, 225 and 152. (B) Isoform stability inversely correlates with Pab1 binding. mRNA isoforms containing  $\geq 10$ ,  $\geq 100$  or  $\geq 1000$  Pab1 reads (15,240, 5,301 and 685 isoforms, respectively) exhibit significant correlation to isoform turnover. (C) Predicted unstructured region (-24 to -18), double-stranded poly(A) (+8 to +16) and low Pab1 binding correlate with increased isoform stability.