RNA Tagging identified *in vivo* protein-RNA interactions.

**a** Schematic of the RT-PCR assay for targeted RNA Tagging. The RT primers and PCR primer sets used in panels b and c are shown. PCR primer set 1 was two gene-specific primers and primer set 2 used a gene-specific forward primer and the U-select RT primer as the reverse PCR primer. **b,c** PUF3-PUP tagged *HSP10* (b) and *COX17* (c) mRNAs. In each panel, gel slices were run on the same gel and were separated here for clarity. RT and PCR primers used in each column are indicated. "-RT" lanes (no reverse transcriptase)
monitored genomic DNA contamination, which was minimal. "dT" lanes used the oligo(dT) primer, and illustrate that polyadenylated mRNA was present in all samples. "U-select" lanes used the U-select primer, which detects RNAs with U-tags. The control strain (BY4742) lacked an RNA Tagging chimera. PUF3-PUP is the active RNA Tagging chimera and PUF3-PUPmut is a catalytically inactive chimera, which harbors active site mutations in the PUP (Asp185Ala, Asp187Ala). d) Representative Sanger sequencing results of tagged HSP10 mRNA. The PCR product from the U-select (U-select) lane of the PUF3-PUP sample in panel b was cloned and individual colonies were sequenced. Black text indicates genomically encoded HSP10 3' UTR sequence, bold blue text indicates non-genomically encoded adenosines (the poly(A) tail), and bold red text indicates non-genomically encoded thymidines, which represent the 3' U-tag added by PUF3-PUP. e) PHD1 mutant alleles. The two PUF-binding elements in PHD1 mRNA were disrupted via UGU to ACA substitutions in the endogenous PHD1 locus. Active or inactive (DD185/187AA) versions of PUP-2 were fused to the endogenous copy of PUF5 (PUF5-PUP and PUF5-PUPmutant, respectively) in the wild-type and mutant PHD1 strains. f) PUF5-PUP requires its binding elements to tag PHD1 mRNA. Lanes are as in panels b and c.
Supplementary Figure 2

Illumina sequencing accurately detected U-tags of multiple lengths.

**a)** Synthetic libraries with various length U-tags, shown here as the reverse complement for clarity. The indicated libraries were paired-end sequenced on an Illumina HiSeq2500. The purple sequence represents the Illumina 5' adapter, the blue sequence represents a poly(A) tail of 12 nucleotides, the red sequence represents U-tags of multiple lengths, and the black sequence represents the U-select RT primer. The starred (*) position in the U0 library was further analyzed in panel b. **b)** Accuracy of identifying Tagged RNAs by a single non-templated uridine. The percent nucleotide composition of position 13 in Read 2 of the U0 library, which corresponds to the starred (*) position in panel a, was calculated and plotted \( (n = 310,745) \). The actual bases detected by sequencing were reverse complemented here for clarity.

---

Nature Methods: doi:10.1038/nmeth.3651
Supplementary Figure 3

Comparison of Puf3p RNA Tagging results and RNA abundance.

a) The mean number of Tagged RNAs detected for Puf3p targets was correlated with the mean length of their U-tag ($\rho = 0.5$, $P = 0$, $n = 476$). Spearman’s correlation coefficient ($\rho$) and associated $P$-value ($P$) are indicated. TRPM, Tagged RNAs Per Million uniquely mapped reads. b) The mean number of Tagged RNAs (TRPM) detected for Puf3p targets was uncorrelated with their mean abundance (Spearman correlation, $P > 0.1$). FPKM, fragments per kilobase of exon per million reads mapped. c) The mean length of the U-tag on Puf3p targets was weakly correlated with their mean abundance (FPKM) ($\rho = -0.37$, $P = 0$, $n = 476$). Spearman’s correlation coefficient ($\rho$) and associated $P$-value ($P$) are indicated.
Supplementary Figure 4

Comparison of Puf3p-binding elements identified with multiple methods.

Proportional Venn diagram of Puf3p targets identified using RNA Tagging, RIP-chip\textsuperscript{15}, and PAR-CLIP\textsuperscript{25}. The numbers indicate the number targets in each area of the plot. Position-weight matrices (plotted in bits) of the Puf3p-binding elements (PBEs) found in each group of targets are indicated. PBEs were derived as follows: PBEs with grey stars, MEME analysis of all RNA Tagging targets; PBEs with grey squares, MEME analysis of all RIP-chip targets; PBEs with grey triangles, PBEs in PAR-CLIP peaks.
**Supplementary Figure 5**

Puf3p target rank was correlated with TRPM and U-tag length but was largely uncorrelated with RNA abundance.

**a)** The mean number of Tagged RNAs (TRPM) detected for Puf3p targets was correlated with their RNA Tagging rank ($\rho = -0.91$, $P = 0$, $n = 476$). Spearman’s correlation coefficients ($\rho$) and associated $P$-values ($P$) are indicated in all panels. TRPM, Tagged RNAs Per Million uniquely mapped reads.

**b)** The RNA Tagging rank of Puf3p targets was correlated with the mean length of their U-tags ($\rho = -0.75$, $P = 0$, $n = 476$).

**c)** RNA Tagging rank of Puf3p targets was largely uncorrelated with their mean RNA abundance ($\rho = 0.16$, $P = 0.0007$, $n = 476$). FPKM, fragments per kilobase of exon per million reads mapped.
Comparison of RNA abundance and the position of binding elements across Puf3p targets.

**a)** Class C targets were the most abundant Puf3p targets. Empirical cumulative distributions of RNA abundance were plotted for all Puf3p targets (left) and the three Puf3p target classes (middle) relative to all mRNAs (all mRNAs, \( n = 6,595 \); Class A, \( n = 92 \); Class B, \( n = 189 \); Class C, \( n = 195 \)). The \( P \)-values from Kolmogorov-Smirnov (KS) tests comparing the different distributions are indicated (right).

**b,c)** Puf3p-binding elements were similarly positioned in the 3’ UTRs of each class of Puf3p targets. The distance from each binding element to the 3’ terminus (b) and the stop codon (c) of the target was calculated and plotted (all targets, \( n = 404 \); Class A, \( n = 90 \); Class B, \( n = 169 \); Class C, \( n = 145 \)) (Tukey whiskers indicated). There were no statistical differences between any of the groups (Fisher-Pitman permutation tests, \( P > 0.1 \)).

**d,e)** The mean number of Tagged RNAs (d) and number of U’s (e) detected for targets were compared to the distance from the PBE to the 3’ terminus for isoforms of 64 Puf3p targets (143 distinct mRNAs) detected by at least 31 reads (24,417 reads total). No significant correlations were observed (Pearson and Spearman correlations, \( P > 0.1 \)).
The number of tagged RNAs and U-tag length were correlated with in vitro binding affinity.

a) Published in vitro binding affinity data of purified Puf3p for the six indicated RNA sequences was obtained and shown here. The median number of Tagged RNAs detected (TRPM) and median U-tag length of Puf3p targets containing six distinct binding elements was calculated and compared to the published in vitro binding affinity ($K_d$) of purified Puf3p for those sequences. Pearson’s ($r$) and Spearman’s ($\rho$) correlation coefficients and the associated $P$-values ($P$) are indicated. TRPM, Tagged RNAs Per Million uniquely mapped reads.
Supplementary Figure 8

Puf3p targets were enriched for mRNAs translated at mitochondria in the absence of cycloheximide.

Published mitochondria-specific ribosome profiling (RP) data in the absence of cycloheximide was mined. Empirical cumulative distributions were plotted for all Puf3p targets (left) and the Puf3p target classes (middle) relative to all mRNAs (all mRNAs, n = 5,609; Class A, n = 92; Class B, n = 188; Class C, n = 193). The P-values from Kolmogorov-Smirnov (KS) tests that compared the different distributions are indicated (right).
### Supplementary Figure 9

mRNAs with known PUF3-dependent half-lives were class A or class B targets.

Summary of published RNA half-lives of the indicated genes in wild-type and *puf3Δ* strains. Puf3p target class, RNA Tagging rank, and Puf3p-binding elements are indicated. "NA" indicates the gene was not identified as a Puf3p target by RNA Tagging.

<table>
<thead>
<tr>
<th>mRNA tested</th>
<th>WT Half-life</th>
<th><em>puf3Δ</em> Half-life</th>
<th>Fold Change (<em>puf3Δ</em> / WT)</th>
<th>Puf3p Target Class</th>
<th>Puf3p Target Rank</th>
<th>PBE</th>
</tr>
</thead>
<tbody>
<tr>
<td>YLL009C</td>
<td>4.0</td>
<td>24.3</td>
<td>6.1</td>
<td>A</td>
<td>5</td>
<td>CUUGUAUAUA</td>
</tr>
<tr>
<td>YML009C</td>
<td>2.9</td>
<td>6.7</td>
<td>2.3</td>
<td>A</td>
<td>11</td>
<td>CCUGUAUAUA</td>
</tr>
<tr>
<td>YKL087C</td>
<td>1.7</td>
<td>4.6</td>
<td>2.7</td>
<td>A</td>
<td>60</td>
<td>CCUGUAUAUA</td>
</tr>
<tr>
<td>YNR037C</td>
<td>4.2</td>
<td>8.5</td>
<td>2.0</td>
<td>A</td>
<td>12</td>
<td>CAUGUAUAUA</td>
</tr>
<tr>
<td>YOR187W</td>
<td>5.5</td>
<td>16.7</td>
<td>3.0</td>
<td>A</td>
<td>25</td>
<td>CGUGUAUAUA</td>
</tr>
<tr>
<td>YNR017W</td>
<td>3.1</td>
<td>8.0</td>
<td>2.6</td>
<td>A</td>
<td>47</td>
<td>CUUGUAUAUA</td>
</tr>
<tr>
<td>YDR347W</td>
<td>3.2</td>
<td>5.1</td>
<td>1.6</td>
<td>B</td>
<td>132</td>
<td>UCUGUAUAUA</td>
</tr>
<tr>
<td>YNL315C</td>
<td>3.8</td>
<td>10.2</td>
<td>2.7</td>
<td>B</td>
<td>183</td>
<td>CCUGUAUAUA</td>
</tr>
<tr>
<td>YHR147C</td>
<td>3.0</td>
<td>4.8</td>
<td>1.6</td>
<td>B</td>
<td>131</td>
<td>CUUGUAUAUA</td>
</tr>
<tr>
<td>YOR158W</td>
<td>1.7</td>
<td>7.2</td>
<td>4.2</td>
<td>B</td>
<td>125</td>
<td>CAUGUAUAUA</td>
</tr>
<tr>
<td>YBL090W</td>
<td>3.1</td>
<td>4.7</td>
<td>1.5</td>
<td>B</td>
<td>202</td>
<td>UUUGUAUAUA</td>
</tr>
<tr>
<td>YDR041W</td>
<td>3.2</td>
<td>3.0</td>
<td>0.9</td>
<td>C</td>
<td>362</td>
<td>CUUGUAUAUA</td>
</tr>
<tr>
<td>YDL069C</td>
<td>&lt;2.0</td>
<td>&lt;2.0</td>
<td>1.0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

** All half-lives taken from Miller, et al. *NAR* (2013)
** "NA" indicates the gene was not identified as a Puf3p target using RNA Tagging.
Supplementary Figure 10

Comparison of Bfr1p RNA Tagging results and RNA abundance.

a) The mean number of Tagged RNAs (TRPM) detected and the mean length of their U-tag were uncorrelated (Spearman correlation, $P > 0.1$). TRPM, Tagged RNAs Per Million uniquely mapped reads. b) The mean number of Tagged RNAs (TRPM) detected for Bfr1p targets was weakly correlated with their mean abundance ($\rho = 0.3, P = 0; n = 1,298$). FPKM, fragments per kilobase of exon per million reads mapped. Spearman's correlation coefficient ($\rho$) and associated $P$-value ($P$) are indicated. c) The mean length of the U-tag on Bfr1p targets was largely uncorrelated with their mean abundance ($\rho = -0.12, P = 0; n = 1,298$). Spearman's correlation coefficient ($\rho$) and associated $P$-value ($P$) are indicated.

Nature Methods: doi:10.1038/nmeth.3651
Supplementary Figure 11

Bfr1p targets identified by both RNA Tagging and RIP-chip were enriched for membrane-associated functions.

Proportional Venn diagram of Bfr1p targets identified using RNA Tagging and RIP-chip\textsuperscript{29}. GO analyses were performed on the three groups and enrichments for representative terms from Biological Process and Cellular Component ontologies are indicated (see Supplementary Data 2 for complete lists).
Supplementary Figure 12

Bfr1p target rank was correlated with TRPM and was very weakly correlated with U-tag length and RNA abundance.

a) The mean number of Tagged RNAs (TRPM) detected for Bfr1p targets was correlated with their RNA Tagging rank ($\rho = -0.87$, $P = 0$; $n = 1,298$). Spearman’s correlation coefficients ($\rho$) and associated $P$-values ($P$) are indicated in all panels. TRPM, Tagged RNAs Per Million uniquely mapped reads. b) The RNA Tagging rank of Bfr1p targets was weakly correlated with the mean length of their U-tags ($\rho = -0.35$, $P = 0$; $n = 1,298$). c) RNA Tagging rank of Bfr1p targets was weakly correlated with their mean RNA abundance ($\rho = -0.28$, $P = 0$; $n = 1,298$). FPKM, fragments per kilobase of exon per million reads mapped.
Supplementary Figure 13

Bfr1p target class was correlated with protein localization to the ER.

The fraction of each class of Bfr1p targets that are localized to the cytoplasm, endoplasmic reticulum (ER), nucleus, mitochondria, and nucleolus, obtained from the yeast GFP database\textsuperscript{32}, was plotted. Classes A-C of Bfr1p targets were highly enriched for ER-localized proteins (hypergeometric tests, $P < 1 \times 10^{-16}$), and the enrichment progressively decreased from Class A to D targets. No other significant enrichments were observed (hypergeometric tests, $P > 0.01$).

\textsuperscript{32} Nature Methods: doi:10.1038/nmeth.3651
Supplementary Figure 14

Bfr1p targets were highly enriched for abundant mRNAs.

Empirical cumulative distributions of RNA abundance (FPKM) were plotted for all Bfr1p targets (left) and the Bfr1p target classes (middle) relative to all mRNAs (all mRNAs, \( n = 6,595 \); Class A, \( n = 174 \); Class B, \( n = 297 \); Class C, \( n = 564 \); Class D, \( n = 261 \)). The \( P \)-values from Kolmogorov-Smirnov (KS) tests comparing the different distributions are indicated (right). Class A Bfr1p targets were most enriched for abundant RNAs and the enrichment progressively decreased to Class C and D targets. FPKM, fragments per kilobase of exon per million reads mapped.

Nature Methods: doi:10.1038/nmeth.3651
Supplementary Figure 15

**Bfr1p bound abundant, ER-translated mRNAs.**

**a)** Plot of the fraction of ER-translated mRNAs (> 2-fold enrichment, \( n = 736 \)), obtained from a published ER-specific ribosome profiling experiment\(^{42}\) (log\(_2\)(ubc6.7mchx) enrichment), that were tagged by Bfr1p (422 mRNAs). **b)** Plots of the RNA abundance (FPKM) of the indicated groups of mRNAs (Tukey whiskers indicated). Of the mRNAs specifically translated at the ER, those tagged by Bfr1p were significantly more abundant than those not tagged by Bfr1p (Fisher-Pitman permutation test, \( P < 10^{-6} \)). FPKM, fragments per kilobase of exon per million reads mapped.