Supplemental Results and Discussion

mRNA Stability

As mRNA is thought to be extensively regulated after transcription (Dreyfuss et al, 2002; Garneau et al, 2007; Hogan et al, 2008) and recent results have shown that mRNA decay and translation are coupled (Harel-Sharvit et al, 2010), we explored the relationship between our protein expression measurements and previous mRNA half-life measurements. We restricted our analysis to data collected from *S. cerevisiae* because mRNA stability measurements have, to date, not been obtained for *S. bayanus*. If our expression measurements and mRNA half-life ratios between the species were positively correlated, then mRNAs with long half-lives measured in *S. cerevisiae* should on average also have higher protein expression in *S. cerevisiae* under the following assumption*: S. bayanus* mRNA half-lives differ significantly from *S. cerevisiae* half-lives, but they are similarly distributed. Under this assumption, an mRNA with a half-life in the tail of the distribution of measurements in *S. cerevisiae* should on average have a less extreme half-life in *S. bayanus*. When we directly correlated our interspecies protein expression ratios and our protein ASE (pASE) ratios with mRNA half-life measured in *S. cerevisiae* (Shalem et al, 2008), we found no correlation with our protein ASE measurements (Fig. S4a) and a modest correlation to our interspecies protein measurements (Fig. S4b) as well as our estimates of the *trans*-acting component of protein expression divergence (Fig. S4c). Interestingly, we found similar pattern (Fig. S4d, S4e, S4f), although less pronounced, using mRNA expression divergence estimates from an independent study comparing *S. cerevisiae* to *S. paradoxus* (Tirosh et al, 2009). Under the limited assumptions of our analysis, our results point to a possible link between mRNA decay and mRNA and protein expression divergence. While the correlations we observe are difficult to interpret without the corresponding mRNA measurements in *S. bayanus*, they warrant further study of the coupling of mRNA decay to mRNA and protein expression divergence between *S. bayanus* and *S. cerevisiae* and other yeast species.

RNA binding proteins

In addition to mRNA half-life, we examined the relationship between our protein ASE measurements and targets of RNA binding proteins also measured only in *S. cerevisiae* (Hogan et al, 2008). To find associations between RNA binding proteins and the protein expression divergence of these targets, we scanned the targets of 46 RNA binding proteins for any differences in the magnitude and directional estimates of interspecies protein expression divergence, protein ASE, and protein *trans* divergence relative to non-target proteins at a p-value cutoff of 0.01. We found that targets of PUF4, a potential destabilizer of ribosome biogenesis factors (Foat et al, 2005), were more highly expressed in the parental *S. bayanus* strain than in *S. cerevisiae* parental strain (Fig. S5a. top) and a significant positive bias in the corresponding *trans* component of this divergence indicated divergence in a *trans* factor, possibly PUF4 itself, contributed to the observed protein expression divergence (Fig. S5a, bottom). We found a similar pattern with the few targets of RNA binding protein PUF3 that we measured (Fig. S5b). PUF3 has been shown to destabilize its targets, primarily mitochondrially targeted proteins, when *S. cervisiae* is grown in glucose (Foat et al, 2005). We also found that the magnitude *cis* and magnitude *trans* divergence of targets of BFR1 to be significantly lower than non-targets (Fig. 5c). Little is known about BFR1, a component of polyribosome associated mRNP complexes (Lang et al, 2001), and our observation warrants further study of this protein's role in mRNA turnover and translation. Last, we found that the magnitude *trans* divergence of targets of RNA binding protein PUB1, a gene with a demonstrated role in stabilizing transcripts containing upstream ORFs (Ruiz-Echevarrìa & Peltz, 2000), to be significantly less than non-targets. We additionally scanned our raw protein expression measurements for shared peptides (e.g. Table S1) and our direct measurements (Table S2) between PUB1 orthologs and found that the median interspecies protein expression level was 31.2% higher in *S. cerevisiae* than in *S. bayanus*. Our results suggest that PUB1 activity may have diverged between these species. We also asked whether the targets of RNA binding proteins were associated estimates of the post-transcriptional component of *cis*-effects on protein expression divergence using mRNA ASE measurements from a previous study (Bullard et al, 2010). We scanned through targets of RNA binding proteins for significant associations between a magnitude and directional estimate of the post-transcriptional component of *cis*-acting regulatory variation, estimated as $\log_2$(posttranscriptional ASE) = $\log_2$(protein ASE) - $\log_2$(mRNA ASE), the off diagonal

component of a scatter plot of mRNA ASE and protein ASE measurements. We found that magnitude of this quantity was significantly less for targets of BFR1 and Scp160 than non-targets of these proteins (Fig. S6). Interestingly, BFR1 and Scp160 are known to interact, and they share significant overlap in their targets (Hogan et al, 2008; Lang et al, 2001). This result may reflect the distinct functional roles of these proteins, as the association with Scp160 was not found using protein expression measurements alone. This additional result further warrants study of the role of these proteins in transcription and translation. While these associations between protein expression divergence and RNA binding proteins may be confounded by divergence in other mechanisms between these yeast species and limited by the absence of corresponding measurements in *S. bayanus*, our analysis illustrates the potential of our method for measuring protein ASE to provide insights into post-transcriptional regulatory differences between species.

Differential translational efficiency

Differential translational efficiency is thought to play a role in phenotypic divergence between species as synonymous codon usage is under extensive selection. Translational efficiency can be predicted accurately in yeast based on the codon adaptation of a gene to the tRNA pool (Man & Pilpel, 2007). We asked whether a simple correlation exists between the differential translational efficiency of orthologous genes between *S. cerevisiae* and *S. bayanus* and our measures of *cis* and *trans* components of protein expression divergence. We found no correlation to mRNA ASE (Spearman's correlation -0.042), protein ASE (0.081), and the post-transcriptional component of protein ASE as estimated by $\log_2$(posttranscriptional ASE) = $\log_2$(protein ASE) - $\log_2$(mRNA ASE) (Spearman's correlation 0.072). However, we found very weak correlations to interspecies protein expression divergence (Spearman's correlation 0.15) and the *trans* component of protein expression divergence (Spearman's correlation 0.12). While little can be concluded by such weak correlations, this result raises the possibility that divergence in the translational machinery might impact protein expression divergence between these yeast species.

uAUG containing transcripts

Recent studies have suggested that uAUGs, start codons in the 5'-untranslated region (5'UTR) of a transcript, act as widespread regulators of mRNA and protein expression levels (Calvo et al, 2009; Yun et al, 2012), we asked whether uAUG-containing transcripts differed in their sensitivity to *cis*-acting regulatory divergence than their non-uAUG containing counterparts. To identify uAUG containing transcripts, we used the 5'UTR lengths from a previous RNA-seq study in yeast (Nagalakshmi et al, 2008). For transcripts where 5'UTR length was not measured by RNA-seq, we used the median 5'UTR length of 100bp. To avoid any confounding signal from misannotated translation start sites in *S. bayanus*, we considered only uAUG containing transcripts in *S. cerevisiae* for which we had both mRNA ASE and protein ASE measurements. Consistent with previous observations that uAUGs are highly conserved across species (Yun et al, 2012), 20 out of 28 of these uAUG containing transcripts also had a uAUG in the corresponding *S. bayanus* transcript. We quantified sensitivity to *cis*-acting regulatory divergence using the magnitude of the $\log_2$ protein ASE and mRNA ASE ratios. We found that uAUG-containing transcripts had a slightly higher sensitivity to *cis*-acting regulatory divergence at mRNA level ($p = 0.021$, Wilcoxon rank sum) and at that the protein level ($p = 0.013$; Wilcoxon rank sum). When we combined sensitivities to *cis*-acting regulatory variation, measured by $|\log_2(\text{protein ASE})| + |\log_2(\text{mRNA ASE})|$, we found a highly significant difference ($p = 0.00091$; Wilcoxon rank sum; Fig. S7). We found no such signal in the magnitude of $\log_2$ interspecies protein expression ratios or in $|\log_2(\textit{trans})|$ the magnitude of protein expression divergence attributable to *trans* regulatory variation. Our result suggests that uAUG-containing transcripts may have greater sensitivity to *cis*-acting regulatory divergence and this sensitivity is jointly reflected in mRNA ASE and protein ASE measurements. Our result warrants further study of the connection between uAUGs and *cis*-acting regulatory divergence.

Supplementary Table Captions

Table S1

Peptides identified and quantified in each replicate at an FDR of 1%. Protein orthologs covered by these peptide quantifications for shared peptides, *S. cerevisiae* variant peptides, and *S. bayanus* variant peptides.

Table S2

High-confidence protein allele-specific expression (pASE) measurements *S. cerevisiae*/*S. bayanus* interspecies hybrid and the corresponding interspecies expression ratios for the parental strains.

Table S3

Interspecies expression measurements computed by directly comparing a heavy labeled *S. cerevisiae* protein sample to an unlabeled *S. bayanus* sample.

Table S4

Protein allele-specific expression (pASE) ratios derived from the synthetic hybrid sample used in our control experiment. Corresponding interspecies expression ratios derived from internal standards comparing *S. cerevisiae* grown in minimal media with a glucose carbon source to *S. bayanus* grown in rich media with an acetate carbon source (YPA).

Supplementary Figures

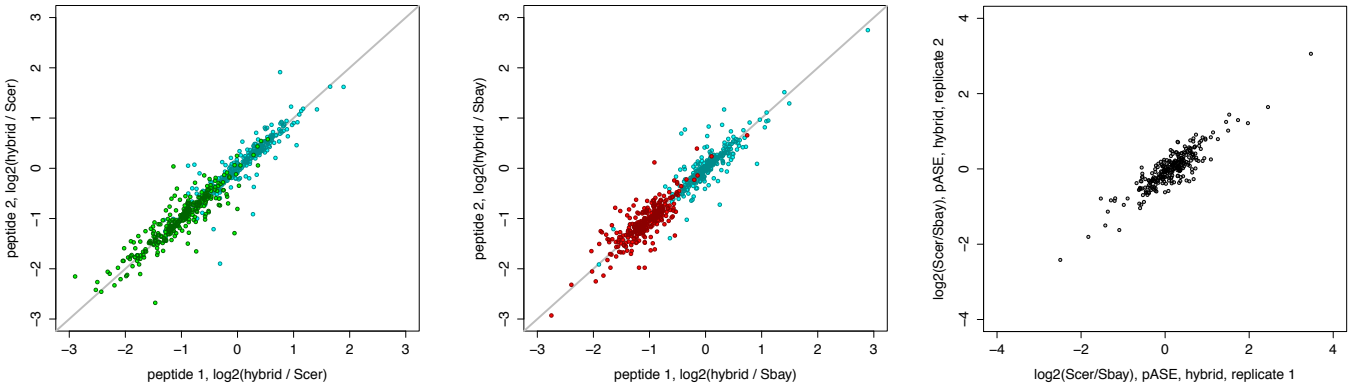


Figure S1

Accuracy and reproducibility of protein allele-specific expression (pASE) ratios. (a) The $\log_2$ peptide ion intensity ratios measured between two, randomly-assigned groups of quantified and distinct shared peptides (cyan) and *S. cerevisiae* (Scer) variant peptides (green) for the hybrid vs. Scer sample. Each point corresponds to 307 proteins for which pASE measurements were derived and two or more distinct shared peptides and two or more variant peptides were quantified in a second technical replicate. (b) The same plot for the hybrid vs. *S. bayanus* (Sbay) comparison for Sbay variant peptides (red) and shared peptides (cyan). (c) Scatter plot of $\log_2$ pASE ratios for 398 proteins that were quantified in both technical replicates.

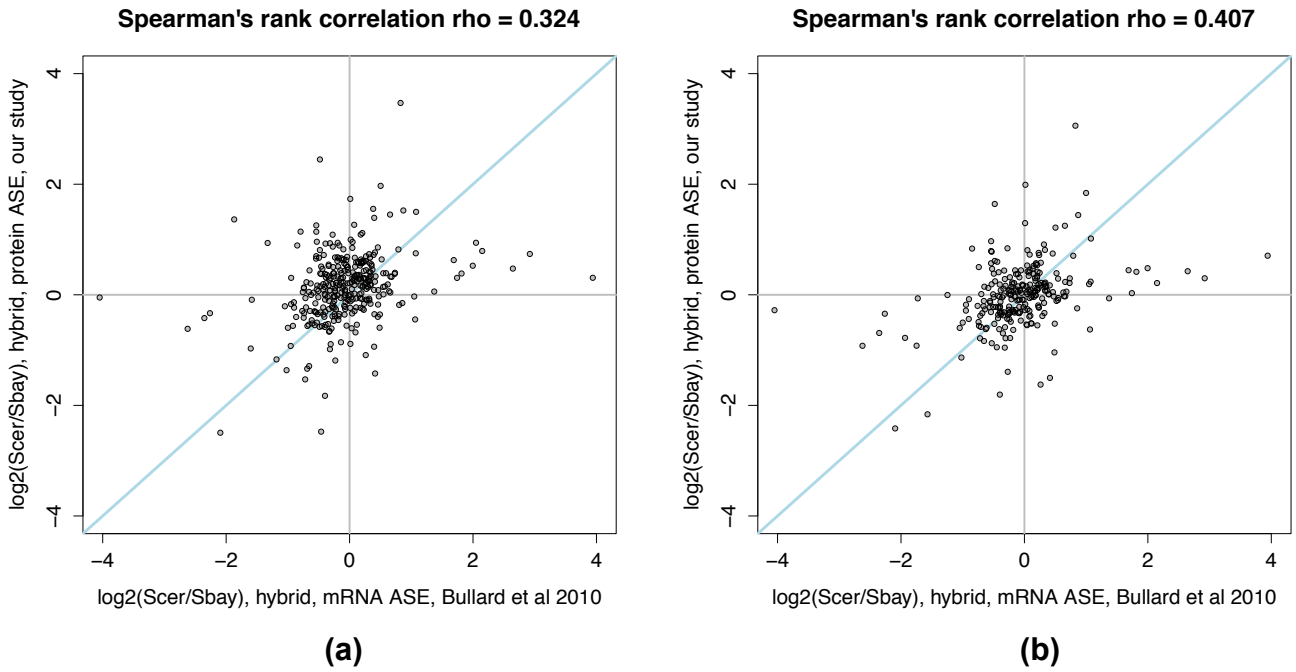


Figure S2

Correlation between mRNA allele-specific expression (ASE) and protein ASE for each individual replicate. (a) Scatterplot of 336 mRNA ASE and protein ASE measurements from replicate 1. The Spearman's correlation of this data is 0.324 and the Pearson's correlation is 0.306. (b) Scatterplot of 261 mRNA ASE and protein ASE measurements from replicate 2. Here, the Spearman's correlation is 0.407 and the Pearson's correlation is 0.354.

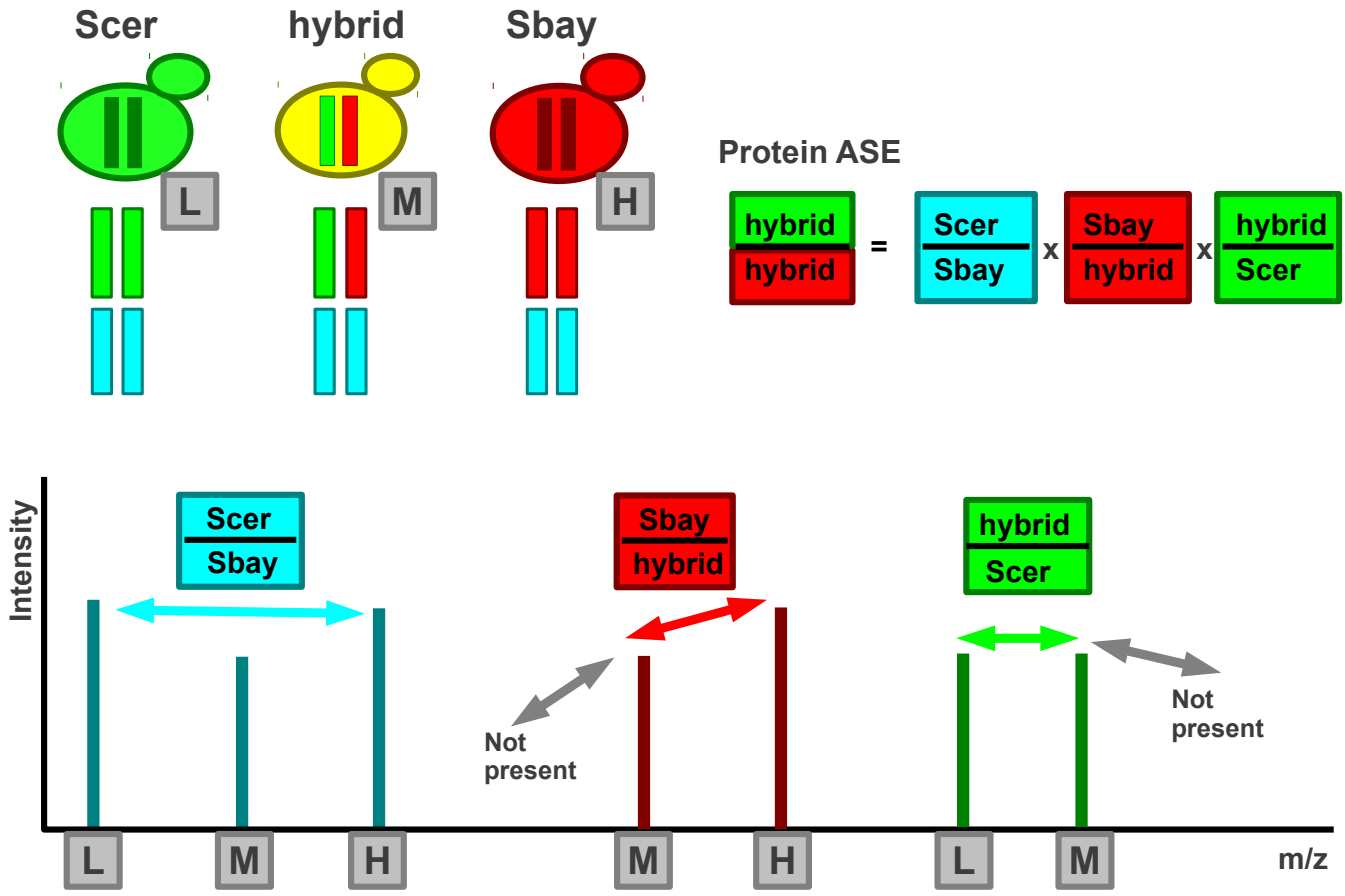


Figure S3

An alternative quantitative proteomics strategy for measuring ratios required for calculation of protein allele-specific expression (pASE). Here, three isotope labels are used. Metabolic incorporation strategies such as stable isotope labeling in cell culture (SILAC) allow for three labels. However, unlike ^{15}N-labeling, which was used in this study, SILAC generally requires auxotrophic strains, more expensive reagents, and limits the proteases that can be used. In this illustration, *S. cerevisiae* (Scer, green) is unlabeled (L, light), the hybrid (yellow) is medium labeled (M, medium), and *S. bayanus* (Sbay, red) is heavy labeled (H, heavy). An example mass spectrum for two protein alleles with shared (cyan) and variant peptides (red and green) is shown. Shared peptides produce light/medium/heavy triplets in mass spectra. Sbay variant peptides generate medium/heavy doublets and Scer variant peptides generate light/medium doublets. The peak heights, or chromatographic peak areas for an LC-MS run, are used to derive the necessary ratios to compute a protein ASE ratio.

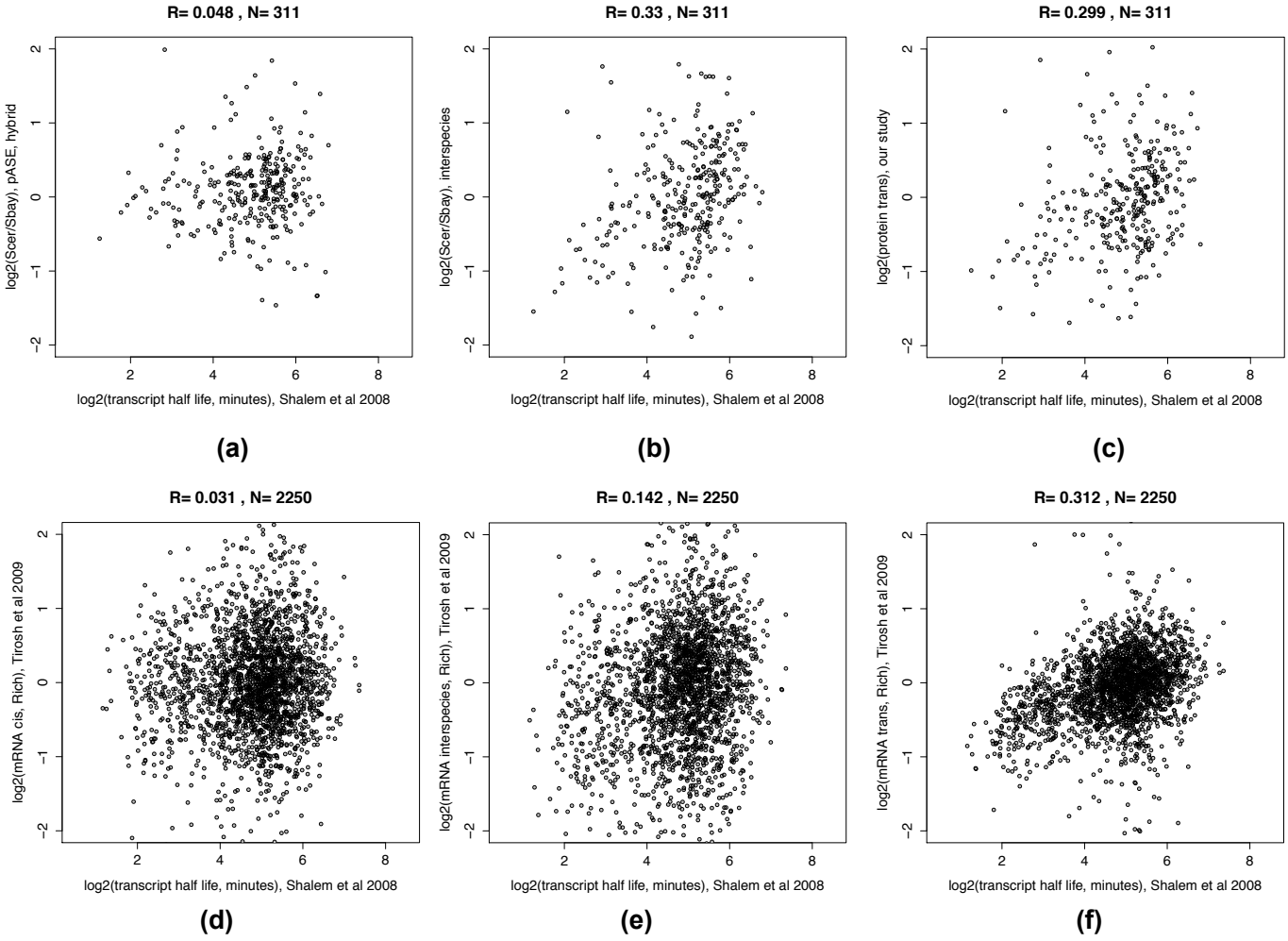


Figure S4

RNA decay half-lives collected in *S. cerevisiae* correlated to protein and mRNA expression divergence. Scatterplots of previous $\log_2$ half-life measurements in *S. cerevisiae* (Shalem et al, 2008) plotted with our measurements of (a) protein ASE from an interspecies hybrid between *S. cerevisiae* and *S. bayanus*; (b) interspecies expression divergence between the parental strains; and (c) the *trans* component of protein expression divergence. Estimates of (a) mRNA ASE from an interspecies hybrid *S. cerevisiae and S. paradoxus*; (b) interspecies mRNA divergence from the respective parental strains; and (c) the *trans* component of mRNA divergence (Tirosh et al, 2009) versus previous measurements of mRNA half-life in *S. cerevisiae* (Shalem et al, 2008). Pearson correlations values are shown above each scatterplot.

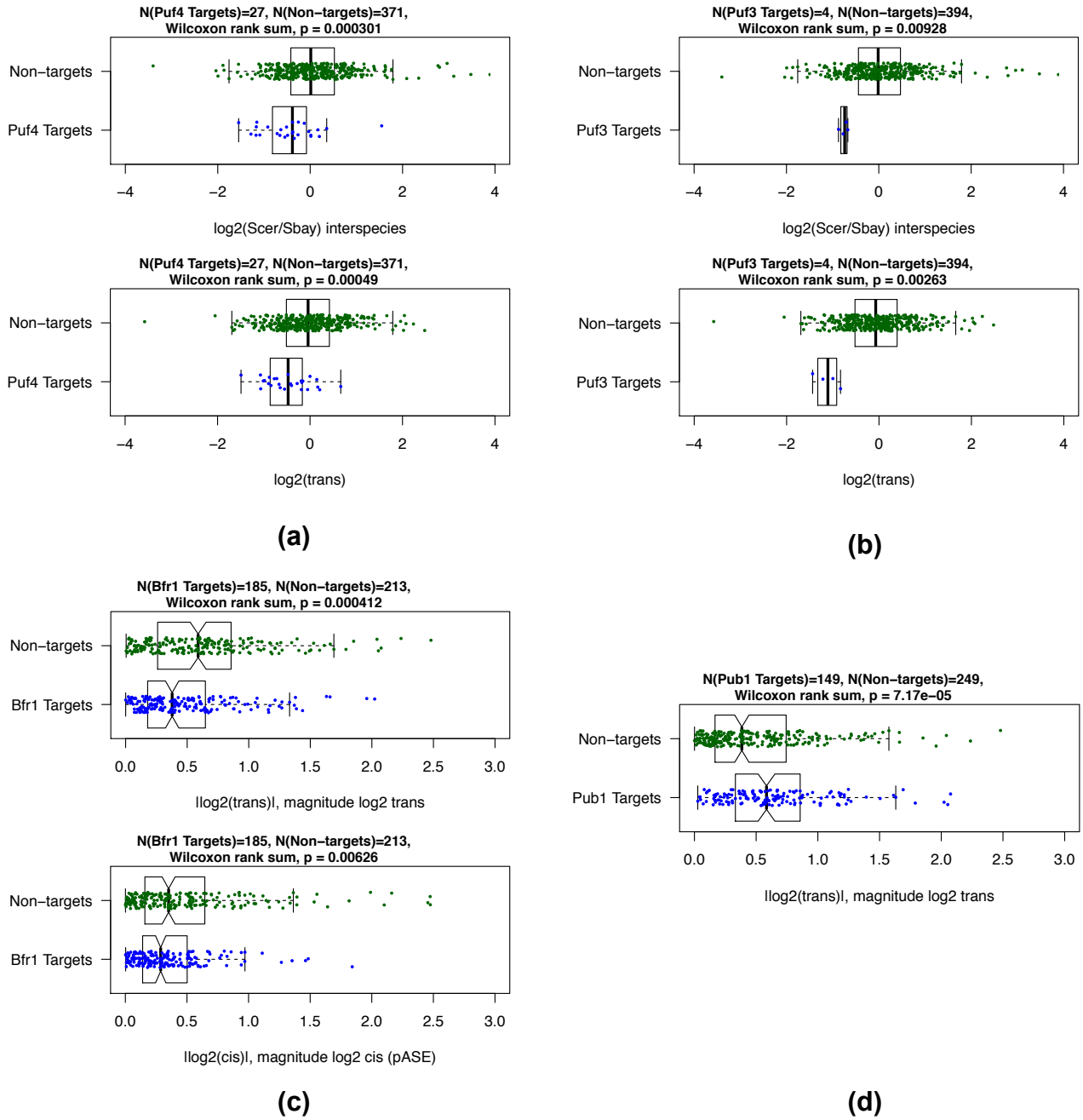
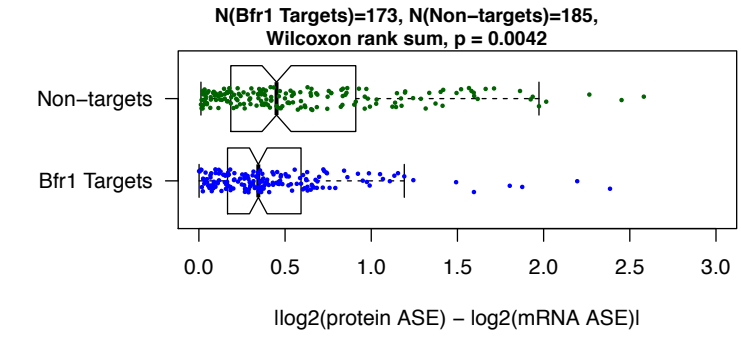
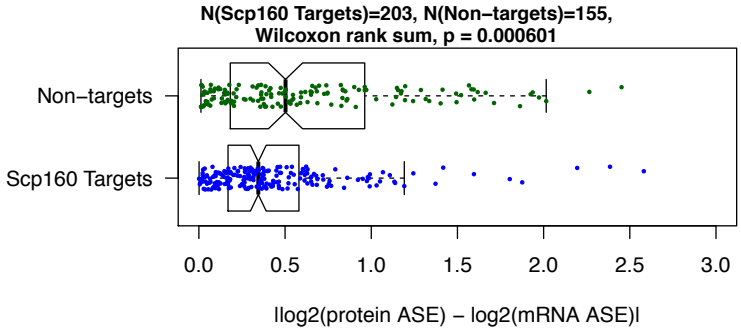


Figure S5
Statistically significant associations between targets and non-targets of RNA binding proteins determined in *S. cerevisiae* (Hogan et al, 2008) and measurements of the protein expression divergence between *S. cerevisiae* and *S. bayanus*. (a) Boxplots of interspecies protein expression measurements and the directional *trans* component of protein expression divergence for targets and non-targets of RNA binding protein PUF4. (b) The same boxplots for RNA binding protein PUF3. (c) Boxplots of significant association between the magnitude of the *trans* component and *cis* component (estimated by protein ASE, pASE measurements in an interspecies hybrid between *S. bayanus* and *S. cerevisiae*) of protein expression divergence and for targets and non-targets of RNA binding protein BFR1. (d) Boxplot of the $|\log_2(\text{trans})|$ magnitude *trans* component of protein expression divergence for targets and non-targets of the RNA binding protein PUB1. Blue points designate measurements of RNA binding protein targets and the corresponding non-target proteins are in green. The center line of the box plots designates the median, ends of boxes designate quartiles, whiskers designate 1.5 times the interquartile range for the respective quartile, and notches, where present, designate the ~95% confidence interval of the median.



(a)



(b)

Figure S6

Statistically significant associations between targets and non-targets of RNA binding proteins determined in *S. cerevisiae* (Hogan et al, 2008) and measurements of the post-transcriptional component of *cis*-acting regulatory variation. mRNA ASE measurements were obtained from a previous study of the same interspecies hybrid (Bullard et al, 2010). Boxplots of the magnitude of these measurements for targets and non-targets of RNA binding proteins (a) BFR1 and (a) Scp160. Blue points designate measurements of RNA binding protein targets and the corresponding non-target proteins are in green. The center line of the box plots designates the median, ends of boxes designate quartiles, whiskers designate 1.5 times the interquartile range for the respective quartile, and notches, where present, designate the ~95% confidence interval of the median.

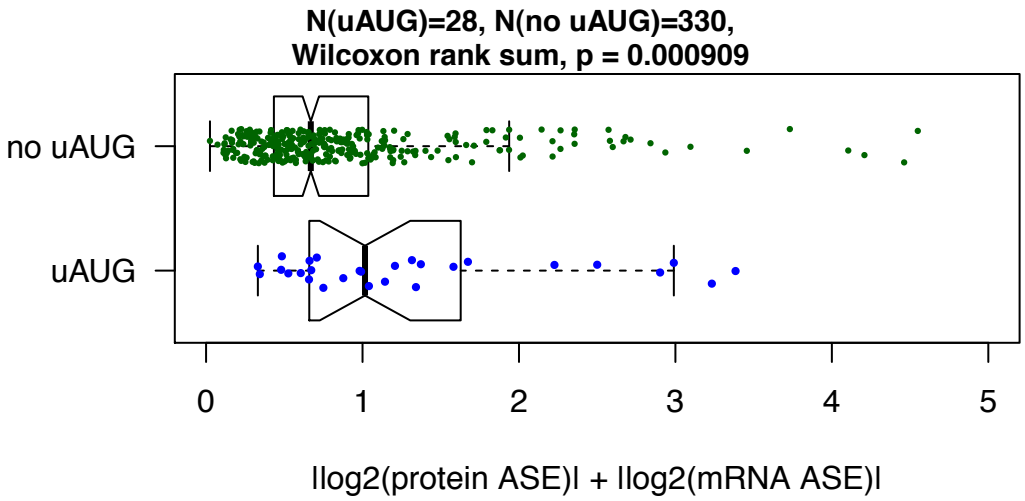


Figure S7

Boxplot of $|\log_2(\text{protein ASE})| + |\log_2(\text{mRNA ASE})|$ the sum of the magnitude mRNA and magnitude protein expression diverge attributable to *cis* regulatory variation for transcripts containing uAUGs, start codons in the 5'-untranslated region (5'UTR), (blue points) and transcripts without these sequence features (green points). The center line of the box plots designates the median, ends of boxes designate quartiles, whiskers designate 1.5 times the interquartile range for the respective quartile, and notches designate the ~95% confidence interval of the median.

References

Bullard JH, Mostovoy Y, Dudoit S, Brem RB (2010) Polygenic and directional regulatory evolution across pathways in Saccharomyces. *Proceedings of the National Academy of Sciences* **107:** 5058-5063

Calvo SE, Pagliarini DJ, Mootha VK (2009) Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci U S A* **106:** 7507-7512

Dreyfuss G, Kim VN, Kataoka N (2002) Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol* **3:** 195-205

Foat BC, Houshmandi SS, Olivas WM, Bussemaker HJ (2005) Profiling condition-specific, genome-wide regulation of mRNA stability in yeast. *Proceedings of the National Academy of Sciences* **102:** 17675-17680

Garneau NL, Wilusz J, Wilusz CJ (2007) The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol* **8:** 113-126

Harel-Sharvit L, Eldad N, Haimovich G, Barkai O, Duek L, Choder M (2010) RNA polymerase II subunits link transcription and mRNA decay to translation. *Cell* **143:** 552-563

Hogan DJ, Riordan DP, Gerber AP, Brown DH, Patrick O (2008) Diverse RNA-Binding Proteins Interact with Functionally Related Sets of RNAs, Suggesting an Extensive Regulatory System. *PLoS Biol* **6:** 255

Lang BD, A L, Black-Brewster HD, L F-KJ (2001) The brefeldin A resistance protein Bfr1p is a component of polyribosome-associated mRNP complexes in yeast. *Nucleic Acids Research* **29:** 2567-2574

Man O, Pilpel Y (2007) Differential translation efficiency of orthologous genes is involved in phenotypic divergence of yeast species. *Nature Genetics* **39:** 415-421

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The Transcriptional Landscape of the Yeast Genome Defined by RNA Sequencing. *Science* **320:** 1344-1349

Ruiz-EchevarrÌa MJ, Peltz SW (2000) The RNA Binding Protein Pub1 Modulates the Stability of Transcripts Containing Upstream Open Reading Frames. *Cell* **101:** 741-751

Shalem O, Dahan O, Levo M, Martinez MR, Furman I, Segal E, Pilpel Y (2008) Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation. *Mol Syst Biol* **4**

Tirosh I, Reikhav S, Levy AA, Barkai N (2009) A Yeast Hybrid Provides Insight into the Evolution of Gene Expression Regulation. *Science* **324:** 659-662

Yun Y, Adesanya TM, Mitra RD (2012) A systematic study of gene expression variation at single-nucleotide resolution reveals widespread regulatory roles for uAUGs. *Genome Res*