



## OPEN

## Coordinating Expression of RNA Binding Proteins with Their mRNA Targets

Huifeng Jiang<sup>1,2</sup>, Lin Xu<sup>2</sup>, Zhe Wang<sup>2</sup>, Jack Keene<sup>3</sup> & Zhenglong Gu<sup>2</sup>

SUBJECT AREAS:

GENE EXPRESSION

EVOLUTIONARY GENETICS

Received  
8 August 2014Accepted  
4 November 2014Published  
24 November 2014Correspondence and  
requests for materials  
should be addressed to  
H.J. (jiang\_hf@tib.cas.  
cn) or Z.G. (zg27@  
cornell.edu)

<sup>1</sup>Key Laboratory of Systems Microbial Biotechnology, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China, <sup>2</sup>Division of Nutritional Sciences, Cornell University, Ithaca, New York 14853, USA, <sup>3</sup>Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27701, USA.

Post-transcriptional regulation by RNA binding proteins (RBPs) plays prominent roles in a variety of biological processes. In this study, by analyzing the global regulatory relationship between RBPs and their target mRNAs in yeast, we discovered that most RBP genes are co-regulated with their target genes, but the RBPs tend to dampen expression variation among their target mRNAs. We further examined a well-studied RBP gene, *PUF3*, and found that the protein decreases the variation of its target mRNAs by differentially affecting their decay. We also constructed a mathematical model to explain the relationship between RBPs and the expression of their target genes. Our results provided new insights into the functional importance of RBPs in coordinating the expression of their target genes.

Coordinating the regulation of functionally related genes by reducing their expression variation in a cell is crucial for the survival of organisms with limited resource in nature<sup>1–4</sup>. In prokaryotes, expression of functionally related genes can be coordinated by forming operons to reduce their expression fluctuation<sup>5</sup>. It was proposed that posttranscriptional regulons in eukaryotes may play a similar role in coordinating the expression of their target genes during posttranscriptional regulation<sup>6–8</sup>. Gene expression is controlled at multiple steps. Post-transcriptional regulation is mediated by small RNAs (e.g. microRNAs) or RNA binding proteins (RBPs) which usually bind to elements in the 3' UTR and orchestrate the fate of their targeted mRNAs. The systems properties and evolution of microRNA posttranscriptional regulons have been well studied<sup>9–12</sup>. A theoretical model for microRNA predicts that they can both modify the mean of their target genes expression level and reduce variance of expression for their target genes<sup>8</sup>. While the functional importance of post-transcriptional regulons becomes increasingly clear, their system properties and evolution remain under-investigated.

RBPs and their regulation have been shown to be important in splicing, stability, translocation and translation of their target mRNAs<sup>3,13</sup>. Experimental studies indicated that the function of RBPs on gene expression is complicated and sometimes will be opposite in different growth conditions<sup>14–17</sup>. Many high-throughput approaches to study RBPs have been recently developed<sup>18–20</sup>. In humans, numerous diseases have been linked to defects in RBP functions<sup>16,21–24</sup>. With many examples of RBPs being identified, it becomes feasible to test whether these post-transcriptional regulons can coordinate the expression of their target genes at the genome level. Budding yeast is an ideal model organism to address this issue, in part, because of the absence of microRNA system in this lineage during evolution<sup>25</sup>. The global identification of target mRNAs for multiple RBPs<sup>26,27</sup> in budding yeast also offers a unique opportunity to examine the regulatory relationships between RBPs and their target mRNAs<sup>28,29</sup>.

In this study, we firstly examined the regulatory relationship between RBPs and their target genes in budding yeast using RNA-seq data under different conditions. The relationships between gene expression of each RBP and the mean and variation of its target genes were investigated. We also used a well-studied RBP gene, *PUF3*, as an example to investigate the impact of RBP on the expression of their target genes. Our results indicate that most RBPs are co-regulated with their target genes. Furthermore, they play an important role in coordinating expression variation of their target genes.

## Results

**RBP buffering the expression of target genes.** To investigate the general relationship between RBPs and their target genes at transcriptional regulation level, we examined the expression correlation coefficient between RBPs and their targets. We selected 33 RBPs each of which has more than 10 target RNAs in *S. cerevisiae*<sup>27</sup>. The



expression of the studied RBPs and their target mRNAs was assessed using currently available RNA-Seq data from 148 expression profiles under multiple conditions in yeast. We calculated the correlation coefficients between the expression level of each RBP with the average expression of all of its mRNA target genes in each condition. The same calculation was also conducted between the level of RBP expression and CD (CD: Coefficient of Deviation, the standard deviation of all target genes expression divided by their mean) of its mRNA target genes.

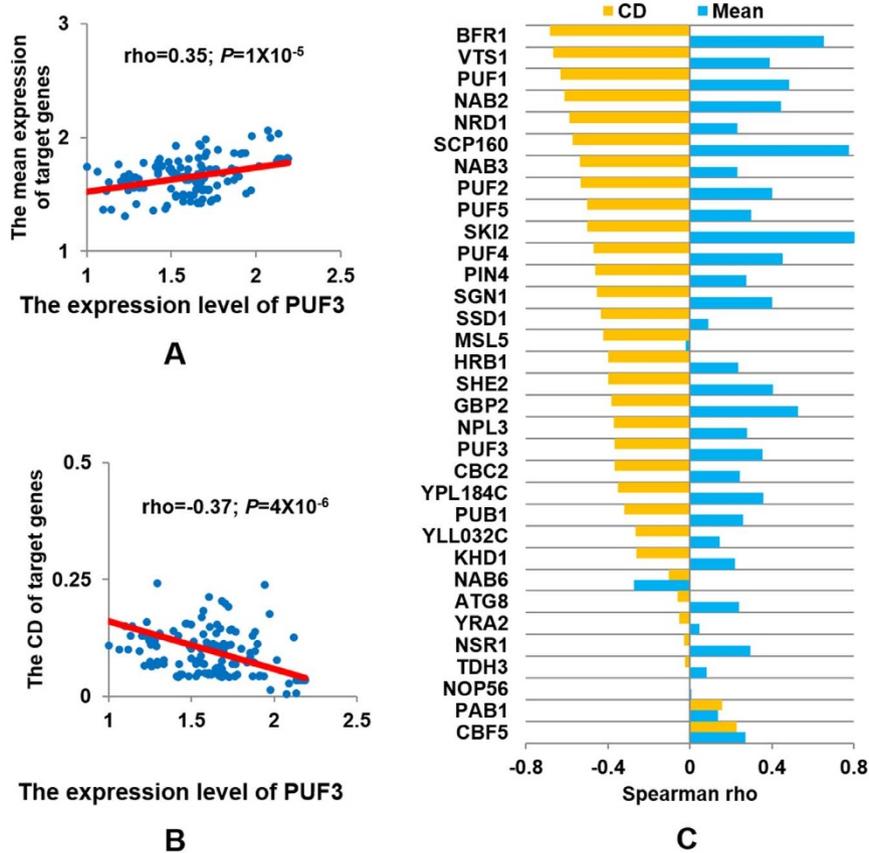
Taking *PUF3*, a well-studied RBP gene, as an example, we found that expression of this RBP is positively correlated with the average expression of its mRNA targets (Fig. 1A). In contrast, the CD value of the expression of Puf3p target genes progressively decreases as its expression level increases (Fig. 1B). We further investigated the relationship between all RBPs and the expression and variation of their target genes in various conditions. As shown in Fig. 1C and Supplementary Table 1, strikingly, a similar pattern was observed for most of the 33 RBPs. As CD is the ratio of the standard deviation (SD) of gene expression among targets to their mean expression, to delineate the impact of the mean expression level of the mRNA targets on the negative relationship we observed, we investigated the relationship between the expression of these RBPs and the SD of the expression of their target genes. As shown in Supplementary Fig. 1, the results remained overall similar.

**Puf3p reducing the expression variation of its target genes.** We used *PUF3* as an example to further investigate the buffering of RBP on the expression of their target genes. The functions of Puf3p and the evolution of this post-transcriptional regulon have been well-

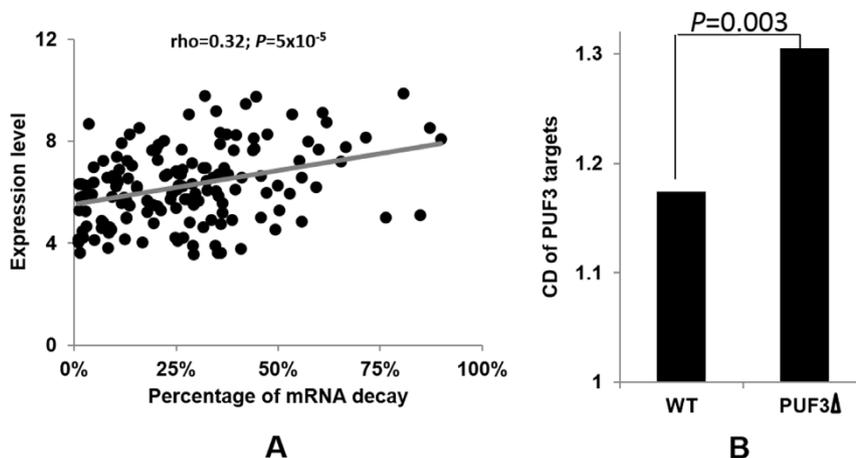
established<sup>15,26,30,31</sup>. As Puf3p was shown to be involved in its target mRNA degradation, we investigated the impact of deleting *PUF3* gene on the expression of its targeted mRNA. The expression of the *Puf3p* target genes in the *PUF3* deletion strain was assumed to approximate their expression before *Puf3p* performs its function. The expression of these target genes in the WT strain was assumed to represent their gene expression upon Puf3p function. The genome-wide expression of both strains was profiled using RNA-Seq (Supplementary Figure 2). The degradation efficiency of each Puf3p target gene was defined as the percentage of gene expression level reduced from the *PUF3* deletion mutant to the WT strain. Interestingly, we found that the percentage of mRNA degradation increases with the level of the targeted gene expression (Fig. 2A).

The significantly positive correlation between the degradation efficiency and the original expression level of the target genes indicated that Puf3p may have a selective effect on its target mRNAs based on their level of expression. If this hypothesis is correct, we inferred that this biased mRNA degradation based on the level of gene expression could decrease the gene expression variation among all the Puf3p targeted genes. To test this, we calculated the coefficient of deviation for gene expression among all the Puf3p target genes in the WT parental strain and the *PUF3Δ* strain. As shown in Figure 2B, our results confirmed our prediction and showed that deletion of the *PUF3* gene indeed leads to an increased variation of gene expression among its target genes. (Fig. 2B, F-test *P*-value = 0.003).

**Coordination of gene expression by RBPs and transcription factors (TFs).** The observed gene expression results from a joint interaction between generation and degradation of mRNAs. To



**Figure 1 | The regulatory relationship for the studied RBPs and their target genes.** (A) Positive correlation between *PUF3* gene expression and the mean expression of all its target genes (long<sub>10</sub>RPKM for each gene is used). (B) Negative correlation between *PUF3* expression and the expression variation (CD, Coefficient of Deviation) of all its target genes. (C) The correlation coefficients (spearman rho) between the gene expression of 33 RBPs and the mean (and CD) of the expression of their target genes. RBPs are ordered based on the spearman rho for the CD values (P-values, for the regression analysis, most of which are statistically significant, were listed in Supplementary Table 1).

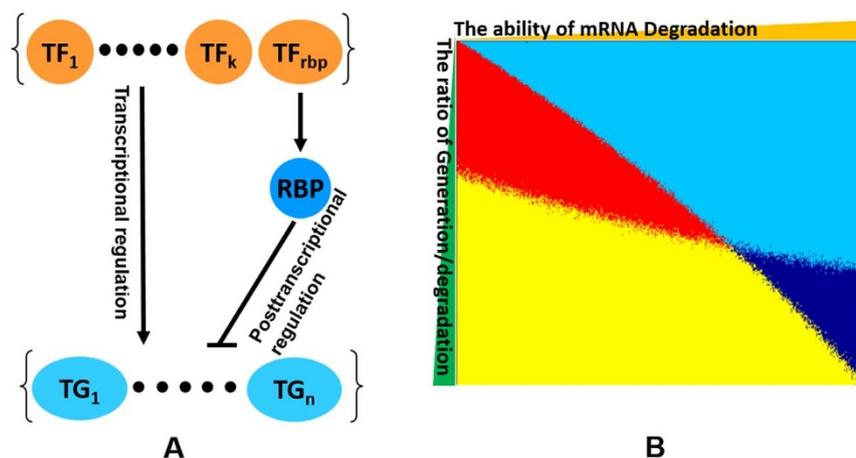


**Figure 2 | Reducing expression variation of target genes by Puf3p.** (A) Biased degradation of Puf3p mRNA targets. The X axis is the percentage of mRNA reduction from the *PUF3Δ* to the WT parental strain, and the Y axis denotes the expression level of Puf3p target genes in the *PUF3Δ* strain (long<sub>2</sub>RPKM for each gene is used). (B) CD values in the WT parental strain and the *PUF3Δ* strain among the Puf3p target genes. F-test was used to test the CD differences.

explain our observation that most RBPs are positively correlated with the average, but negatively correlated with the variation, of the expression of their target genes, we constructed a simple mathematical model to simulate the behavior of RBPs and their targets during gene expression. As shown in Fig. 3A, TFs increase the expression of the target genes and RBPs degrade the transcript of the target genes. According to the relative importance of mRNA generation by TFs and mRNA degradation by RBPs, we could observe four different relationships between RBPs and the expression of their target genes (Fig. 3B). More than 85% studied RBPs are co-expressed with their targets and at the same time reduce the expression variation of the target genes (the yellow region). This relationship between RBPs and their targets occurs when the generation of mRNA by TFs is stronger than the transcript degradation by RBPs, even though the mRNAs of targeted gene are degraded by RBPs. This mode of gene regulation by RBPs is similar to the incoherent feed-forward loop (iFFL) proposed for the miRNA regulation networks where iFFL is frequently used to reduce expression variation of the miRNA target genes<sup>1,8,32</sup>.

## Discussion

In summary, we discovered that most RBPs tend to be co-regulated with their targeted mRNAs under various conditions, but the up-regulation of RBPs can reduce the expression variation among their targeted mRNAs. According to the combinatorial RNA regulon model, RBPs can be involved in multiple overlapping biological processes<sup>3</sup>. Our results revealed that besides these well-demonstrated functions, many RBPs appear to have an ability to reduce expression variation among their targeted mRNA. Gene expression profiling with a genetic knockout of the yeast *PUF3* gene directly support this conclusion. DNA operons were proposed to enable prokaryotes to synchronize the levels of different components of the same functional pathways, thus reducing the variation that is associated with the production of the functionally related proteins<sup>5</sup>. Upon disappearance of DNA operons in most eukaryotic organisms during evolution, the suppression of expression variation among the targeted mRNAs by RBPs may represent a unique mode to achieve similar functional advantages. We further demonstrated that RBPs could conduct similar functions as miRNAs in other organisms in reducing



**Figure 3 | The competition and cooperation between RBPs and TFs.** (A) The model of RBPs regulation system was listed.  $TF_{rbp}$  activates the expression of RBP and  $TF_1$  to  $TF_k$  activate the expression of target genes from  $TG_1$  to  $TG_n$ . RBP represses the expression of all target genes. (B) The X axis (the parameter  $j$ ) is the impact of mRNA degradation by RBPs and the Y axis (the parameter  $m$ ) is the ratio of mRNA generation by TFs and degradation by RBPs. Four colors indicate the relationships between RBPs and the expression of their target genes. The red means a positive correlation with both average and CD of the targets; the yellow means a positive correlation with average and a negative correlation with CD of the targets; the light blue means a negative correlation with the average and a positive correlation with the CD of the targets and the dark blue means a negative correlation with both average and CD of targets.



expression variation. The molecular mechanisms underlying our observation and its evolutionary significance warrant further study, which can also be employed to create coherent biological circuits in genetic engineering<sup>33,34</sup>.

## Methods

**RNA-Seq data and the target genes of RBPs.** The RBP genes and their targets were downloaded from Dr. Patrick Brown's laboratory<sup>27</sup>. The cutoff to define a target gene for each RBP is  $q$ -value  $< 0.001$ . Finally 33 RBPs with more than 10 target genes were used in this study. The Fastq files of RNA-Seq data for *S. cerevisiae* were downloaded from the DNAnexus website (<http://sra.dnexus.com/>) and NCBI, 148 experimental conditions were listed (Supplementary Table 2). All Fastq files were mapped onto the *S. cerevisiae* genome by SOAP2 and the expression level of each gene was estimated by RPKM (Reads Per Kilobase per Million reads)<sup>35</sup>. After log transformation, the mean, SD and CD of expression for target genes of each RBP were calculated. The correlation coefficients between the expression of RBPs and the mean, SD & CD of their target genes were calculated using `cor.test` function in R<sup>36</sup>. We also conducted similar analysis using microarray data in different conditions and our conclusion in Fig. 1 still hold true (data not shown).

**Yeast growth condition and RNA Seq experiments.** The *PUF3* gene deletion mutant strain (*PUF3Δ*) was from our previous study<sup>31</sup>. Both the *PUF3* gene deletion strain and the wild type BY4741 (*MATa, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0*) strain were cultured overnight in YPD media. 0.01 OD cells were transferred into fresh YPEG media until 1.0 OD. Total RNA was extracted by the standard Trizol protocol. mRNA was then purified using oligo-dT DynaBeads. cDNA sequencing library was constructed according to the protocol described by Wang et al.<sup>37</sup>. After sequencing, the reads were also mapped into *S. cerevisiae* genome by SOAP2 with no more than two mismatches. RPKM was used to represent the expression level of each gene. The high-throughput sequencing data has been deposited in NCBI Gene Expression Omnibus under the accession number: GSE55419.

**Constructing the gene regulation model.** In this model, we assumed that there was a biological system with one transcription factor (TF) and one RNA binding protein (RBP) to regulate 100 target genes (TGs) at the same time. Based on previous studies<sup>38,39</sup>, a well-established mathematical model that was used to study transcriptional programs of a gene is:

$$\frac{d[TG]_i}{dt} = \frac{A_i}{\left(\frac{T_i}{TF}\right)^{n_i} + 1} - j \times [RBP] \times [TG]_i - k_i \times [TG]_i$$

$[TG]_i$ ,  $[TF]$  and  $[RBP]$  represent the expression level of transcriptional target gene  $i$  ( $i = 1, 2, \dots, 100$ ), transcriptional factor and RNA binding protein at time  $t$ , respectively.  $A_i$  and  $T_i$  represent the maximal expression level and the activation coefficient for transcriptional target gene  $i$ , respectively. In addition,  $n_i$  and  $k_i$  are Hill coefficient governing steepness of transcriptional activation term and the decay rate of transcriptional target gene  $i$ , respectively.  $j$  represents the rate for RBP to break down the TGs, which is dependent on the properties of unique RBP. Different RBPs might have distinct  $j$  values. The initial expression level for transcriptional target gene  $i$  is defined as  $[TG]_{i0}$ .

Note that, different transcriptional target gene  $i$  might have distinct  $[TG]_{i0}$ ,  $A_i$ ,  $T_i$ ,  $n_i$ , and  $k_i$  even when the same TF and RBP regulate them. To compare TFs and RBPs, we assume  $m = [TF]/[RBP]$ . As a result,  $[TF]$  in the above equation could be substituted by  $m \times [RBP]$ . In addition,  $m$  represents an important parameter to measure the relative importance between TF in producing transcript and RBP in degrading transcripts. Based on a previous experimental estimation for the above parameters in yeast *S. cerevisiae*<sup>40</sup> and other studies (reviewed in<sup>39</sup>), we assumed the parameter spaces for  $[TG]_{i0}$ ,  $A_i$ ,  $T_i$ ,  $n_i$ , and  $k_i$  to be (0, 1), (1, 2), (0, 4), (1, 4) and (0, 1), respectively. Note that,  $[TG]_{i0}$  and  $A_i$  are defined as (0, 1) and (1, 2), which is to avoid the unrealistic situation that  $[TG]_{i0}$  could be larger than  $A_i$  in the random simulation process. Similarly, the parameter space for expression level of RBP ( $[RBP]$ ) is (0, 2), for which 0 and 2 represent no expression and maximum expression level, respectively. To simply the simulation, we assumed the parameter spaces for  $m$  and  $j$  to be (0, 2) and (0, 1), respectively. As a result, since  $m = [TF]/[RBP]$ , (0, 2) as  $m$ 's parameter space would allow us to simulate both situations that  $[TF]$  is higher or lower than  $[RBP]$ . Based on all of these reasonable assumptions, we randomly produced a series of parameters for each gene of the 100 TGs. The average expression levels and CD for the 100 TGs under one  $[RBP]$  value are calculated. Under each combination of  $m$  and  $j$  values, this process is repeated 10,000 times with 1,000 different  $[RBP]$  values to calculate the correlation coefficients between  $[RBP]$  and average expression levels (or CD) of 100 TGs.

- Hornstein, E. & Shomron, N. Canalization of development by microRNAs. *Nat. Genet.* **38**, S20–4 (2006).
- Cui, Q. et al. MicroRNA regulation and interspecific variation of gene expression. *Trends Genet.* **23**, 372–5 (2007).
- Keene, J. D. RNA regulons: coordination of post-transcriptional events. *Nat. Rev. Genet.* **8**, 533–43 (2007).

- Li, X. et al. A microRNA imparts robustness against environmental fluctuation during development. *Cell* **137**, 273–82 (2009).
- Sneppen, K. et al. Economy of operon formation: cotranscription minimizes shortfall in protein complexes. *MBio* **1**, e00177–10 (2010).
- Keene, J. D. & Tenenbaum, S. A. Eukaryotic mRNPs may represent posttranscriptional operons. *Mol. Cell* **9**, 1161–7 (2002).
- Halbeisen, R. E. & Gerber, A. P. Stress-Dependent Coordination of Transcriptome and Translatome in Yeast. *PLoS Biol.* **7**, e105 (2009).
- Wu, C. L., Shen, Y. & Tang, T. Evolution under canalization and the dual roles of microRNAs: a hypothesis. *Genome Res.* **19**, 734–43 (2009).
- Lu, J. et al. The birth and death of microRNA genes in Drosophila. *Nat. Genet.* **40**, 351–5 (2008).
- Krol, J., Loedige, I. & Filipowicz, W. The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* **11**, 597–610 (2010).
- Carthew, R. W. & Sontheimer, E. J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**, 642–55 (2009).
- Berezikov, E. Evolution of microRNA diversity and regulation in animals. *Nat. Rev. Genet.* **12**, 846–60 (2011).
- Halbeisen, R. E. et al. Post-transcriptional gene regulation: from genome-wide studies to principles. *Cell. Mol. Life Sci.* **65**, 798–813 (2008).
- Foat, B. C. et al. Profiling condition-specific, genome-wide regulation of mRNA stability in yeast. *Proc. Natl. Acad. Sci. USA* **102**, 17675–80 (2005).
- Saint-Georges, Y. et al. Yeast mitochondrial biogenesis: a role for the PUF RNA-binding protein Puf3p in mRNA localization. *PLoS One* **3**, e2293 (2008).
- Klein, M. E. et al. RNA-binding protein Sam68 controls synapse number and local beta-actin mRNA metabolism in dendrites. *Proc. Natl. Acad. Sci. USA* **110**, 3125–30 (2013).
- Zhao, W. et al. Massively parallel functional annotation of 3' untranslated regions. *Nat. Biotechnol.* **32**, 387–91 (2014).
- Zhang, C. & Darnell, R. B. Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat. Biotechnol.* **29**, 607–14 (2011).
- Tome, J. M. et al. Comprehensive analysis of RNA-protein interactions by high-throughput sequencing-RNA affinity profiling. *Nat. Methods* **11**, 683–8 (2014).
- Hafner, M. et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–41 (2010).
- Kechavarzi, B. & Janga, S. C. Dissecting the expression landscape of RNA-binding proteins in human cancers. *Genome Biol.* **15**, R14 (2014).
- Medrano, J. V. et al. Divergent RNA-binding proteins, DAZL and VASA, induce meiotic progression in human germ cells derived in vitro. *Stem Cells* **30**, 441–51 (2012).
- Lukong, K. E. et al. RNA-binding proteins in human genetic disease. *Trends Genet.* **24**, 416–25 (2008).
- Cooper, T. A., Wan, L. & Dreyfuss, G. RNA and disease. *Cell* **136**, 777–93 (2009).
- Drinnenberg, I. A. et al. RNAi in budding yeast. *Science* **326**, 544–50 (2009).
- Gerber, A. P., Herschlag, D. & Brown, P. O. Extensive association of functionally and cytologically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol.* **2**, E79 (2004).
- Hogan, D. J. et al. Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. *PLoS Biol.* **6**, e255 (2008).
- Joshi, A., Beck, Y. & Michoel, T. Post-transcriptional regulatory networks play a key role in noise reduction that is conserved from micro-organisms to mammals. *FEBS J.* **279**, 3501–12 (2012).
- Joshi, A., Van de Peer, Y. & Michoel, T. Structural and functional organization of RNA regulons in the post-transcriptional regulatory network of yeast. *Nucleic Acids Res.* **39**, 9108–17 (2011).
- Jiang, H. et al. Rewiring of posttranscriptional RNA regulons: Puf4p in fungi as an example. *Mol. Biol. Evol.* **29**, 2169–76 (2012).
- Jiang, H., Guan, W. & Gu, Z. Tinkering evolution of post-transcriptional RNA regulons: puf3p in fungi as an example. *PLoS Genet.* **6**, e1001030 (2010).
- Yu, X. et al. Analysis of regulatory network topology reveals functionally distinct classes of microRNAs. *Nucleic Acids Res.* **36**, 6494–503 (2008).
- Murphy, K. F. et al. Tuning and controlling gene expression noise in synthetic gene networks. *Nucleic Acids Res.* **38**, 2712–26 (2010).
- Wang, Z. & Zhang, J. Impact of gene expression noise on organismal fitness and the efficacy of natural selection. *Proc. Natl. Acad. Sci. USA* **108**, E67–76 (2011).
- Li, R. et al. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966–7 (2009).
- Best, D. J. & Roberts, D. E. Algorithm AS 89: The Upper Tail Probabilities of Spearman's Rho. *J. Roy. Stat. Soc. C-App.* **24**, 377–379 (1975).
- Wang, L. et al. A low-cost library construction protocol and data analysis pipeline for Illumina-based strand-specific multiplex RNA-seq. *PLoS One* **6**, e26426 (2011).
- Sasik, R. et al. Extracting transcriptional events from temporal gene expression patterns during Dictyostelium development. *Bioinformatics* **18**, 61–6 (2002).
- Alon, U. *An introduction to systems biology: design principles of biological circuits.* (Chapman & Hall/CRC, Boca Raton, Florida, USA, 2007).
- Xu, L. & Gu, Z. Asynchronous regulation of backup circuits as a strategy for survival in the fluctuating environments. *Int. J. Syst. Synth. Biol.* **1**, 12 (2010).

## Acknowledgments

We thank Jian Lu, Kaixiong Ye, Zhihui Yan, Lina Lu and Brandon Barker for discussions and comments. This work was supported by a 973 Program (2011CBA00805) and a Tianjin



Sci-tech Support Program (11ZCKFSY08000) to H.J. and NSF (MCB-1243588) and NIH (1R01AI085286) to Z.G.

### Author contributions

H.J., J.K. and Z.G. designed experiments. H.J. and L.X. performed data analysis and Z.W. performed the experiments. H.J. and Z.G. wrote the paper. All authors reviewed the manuscript.

### Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Jiang, H., Xu, L., Wang, Z., Keene, J. & Gu, Z. Coordinating Expression of RNA Binding Proteins with Their mRNA Targets. *Sci. Rep.* 4, 7175; DOI:10.1038/srep07175 (2014).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>