the FEBS Journal

Post-transcriptional regulatory networks play a key role in noise reduction that is conserved from micro-organisms to mammals

Anagha Joshi[1], Yvonne Beck[2,3] and Tom Michoel[3]

1 Department of Haematology, Cambridge Institute for Medical Research, University of Cambridge, UK
2 Faculty of Biology, University of Freiburg, Germany
3 Freiburg Institute for Advanced Studies, University of Freiburg, Germany

Keywords
network comparison; post-transcriptional regulation; RNA-binding proteins

Correspondence
T. Michoel, Freiburg Institute for Advanced Studies, University of Freiburg, Albertstrasse 19, 79104 Freiburg, Germany
Fax: +49 761 203 97323
Tel: +49 761 203 97346
E-mail: tom.michoel@frias.uni-freiburg.de

(Received 15 December 2011, revised 17 February 2012, accepted 14 March 2012)

doi:10.1111/j.1742-4658.2012.08571.x

RNA-binding proteins (RBPs) are core regulators of mRNA transcript stability and translation in prokaryotes and eukaryotes alike. Genome-wide studies in yeast have shown intriguing relationships between the expression dynamics of RBPs, the structure of post-transcriptional regulatory networks of RBP–mRNA binding interactions and noise reduction in post-transcriptionally regulated expression profiles. In the present study, we assembled and compared the genomic properties of RBPs and integrated transcriptional and post-transcriptional regulatory networks in four species: *Escherichia coli*, yeast, mouse and human. We found that RBPs are consistently regulated to have minimal levels of protein noise, that known noise-buffering network motifs are enriched in the integrated networks and that post-transcriptional feedback loops act as regulators of other regulators. These results support a general model where RBPs are the key regulators of stochastic noise-buffering in numerous downstream cellular processes. The currently available datasets do not allow clarification of whether post-transcriptional regulation by RBPs and by noncoding RNAs plays a similar or distinct role, although we found evidence that specific combinations of transcription factors, RBPs and micro-RNAs jointly regulate known disease pathways in humans, suggesting complementarity rather than redundancy between both modes of post-transcriptional regulation.

Introduction

Transcriptional regulation has been studied in detail for many years through large-scale measurement of transcriptome expression levels by microarrays [1] and, more recently, RNA-seq (i.e. whole transcriptome shotgun sequencing) [2], as well as direct mapping of interactions between transcription factors (TFs) and DNA by chromatin immunoprecipitation combined with microarray technology [3] and also chromatin immunoprecipitation combined with massively parallel DNA sequencing [4]. As the technology for proteome analysis is becoming ever more sophisticated, a view is emerging that an additional, equally pervasive layer of post-transcriptional regulation exists, which acts in concert with transcriptional regulation to control gene expression and translation in a highly coordinated complex system.

One of the first comparisons between transcriptome and proteome levels has shown that $\sim$ 30% of the variance in protein abundance in yeast grown in rich medium cannot be explained by mRNA expression levels

Abbreviations
IGF2BP, insulin-like growth factor 2 mRNA binding protein; miRNA, micro-RNA; ncRNA, noncoding RNA; RBP, RNA-binding protein; TF, transcription factor; VEGF, vascular endothelial growth factor.

[5]. A refined view under six different stress conditions has been obtained using ribosome associations (translatome) rather than true proteome measurements and indicates that severe stress induces a highly correlated transcriptional and translational response, whereas mild stress leads to a noncorrelated response [6]. A direct comparison of functional clusters inferred from transcriptome and translatome data under these six conditions revealed the existence of three classes of proteins: transcriptionally co-regulated proteins cluster together in transcriptome as well as translatome data and represent basic metabolic processes; post-transcriptionally co-regulated proteins cluster together only in translatome data and consist of RNA-binding, ribosomal and protein synthesis proteins; and dually co-regulated proteins have intermediate co-clustering characteristics and are presumably regulated at both levels [7]. Unexpectedly, an almost identical picture emerges when comparing transcriptome and proteome co-regulation in genetically diverse yeast segregants under normal growth conditions [8]. An even more detailed experimental study compared dynamic transcriptome and proteome profiles as yeast cells respond to environmental stress and revealed four distinct classes of post-transcriptionally regulated proteins, depending on the direction and timing of proteome versus transcriptome change and together encompassing more than 40% of the proteome [9].

An important class of post-transcriptionally regulated proteins identified by Lee *et al.* [9] have reduced noise at the proteome level and are enriched for stress-defence proteins, translation factors and ribosomal proteins, consistent with the fact that these proteins cluster together at the proteome but not the transcriptome level [7,8], and that such clusters tend to be expressed with higher mRNA and protein levels and less noise in single-cell data [7]. Within this class of post-transcriptionally regulated proteins with distinct expression dynamics are the RNA-binding proteins (RBPs) [10], which play a crucial role in the post-transcriptional regulation of gene expression itself [11]. RBPs tend to bind functionally related mRNAs and most mRNAs are bound by multiple RBPs, resulting in a complex network of post-transcriptional regulatory interactions [12]. Analysis of this network in conjunction with the transcriptional regulatory network of Harbison *et al.* [3] and the transcriptome and translatome data of Halbeisen and Gerber [6] has shown that RBPs often partner with specific TFs to filter noise from transcriptional expression profiles and that noise-suppressing network motifs such as post-transcriptional feedback loops are associated with post-transcriptionally co-regulated translatome clusters [7].

Although large-scale studies of the effects of post-transcriptional regulation have so far been carried out almost exclusively in the eukaryotic model system *Saccharomyces cerevisiae*, the remarkable degree of consistency between different experimental platforms, between the response to genetic and environmental perturbations, between static and dynamic data, and between expression and interaction data, all suggest that general mechanisms are at work, consistent with RNA regulon theory, which hypothesizes that transacting factors combinatorially regulate multiple mRNAs to achieve functionally coherent translation in the face of stochastic gene transcription [13]. Although still challenging as a result of the limited availability of large-scale datasets, in the present study, we investigated the universality of these conclusions across multiple species, from prokaryotes (*Escherichia coli*) and simple eukaryotes (*S. cerevisiae*) to mammals (mouse and human).

Results

Expression dynamics of RNA-binding proteins

RNA-binding proteins not only are important post-transcriptional regulators, but also are themselves post-transcriptionally regulated. A detailed analysis of the expression dynamics of RBPs in yeast has shown that, at the transcript level, they are highly expressed with a low half-life, whereas, at the protein level, they are highly expressed with low noise [10]. We collected genomic properties for *E. coli*, yeast and mouse (for details, see Materials and methods) and found that RBPs are expressed with a significantly higher mRNA level (all three species), a significantly lower mRNA half-life (yeast and mouse), a significantly higher protein abundance (all three species) and significantly lower protein noise (*E. coli* and yeast, no data available for mouse) (Fig. 1).

Protein noise is generally found to be inversely correlated with protein abundance [14,15], with the caveat that the level of noise in protein abundance is likely to be condition-specific [9]. In mathematical models for stochastic gene expression, it is shown that protein noise is further reduced if the transcript half-life is significantly shorter than the protein half-life as a result of a time-averaging effect [16]. In multiple organisms across wide evolutionary distances, RBPs thus appear to be tightly regulated at the post-transcriptional level to have minimal noise in their protein levels. This is consistent with a view that RBPs play a central role in the noise management of a cell.

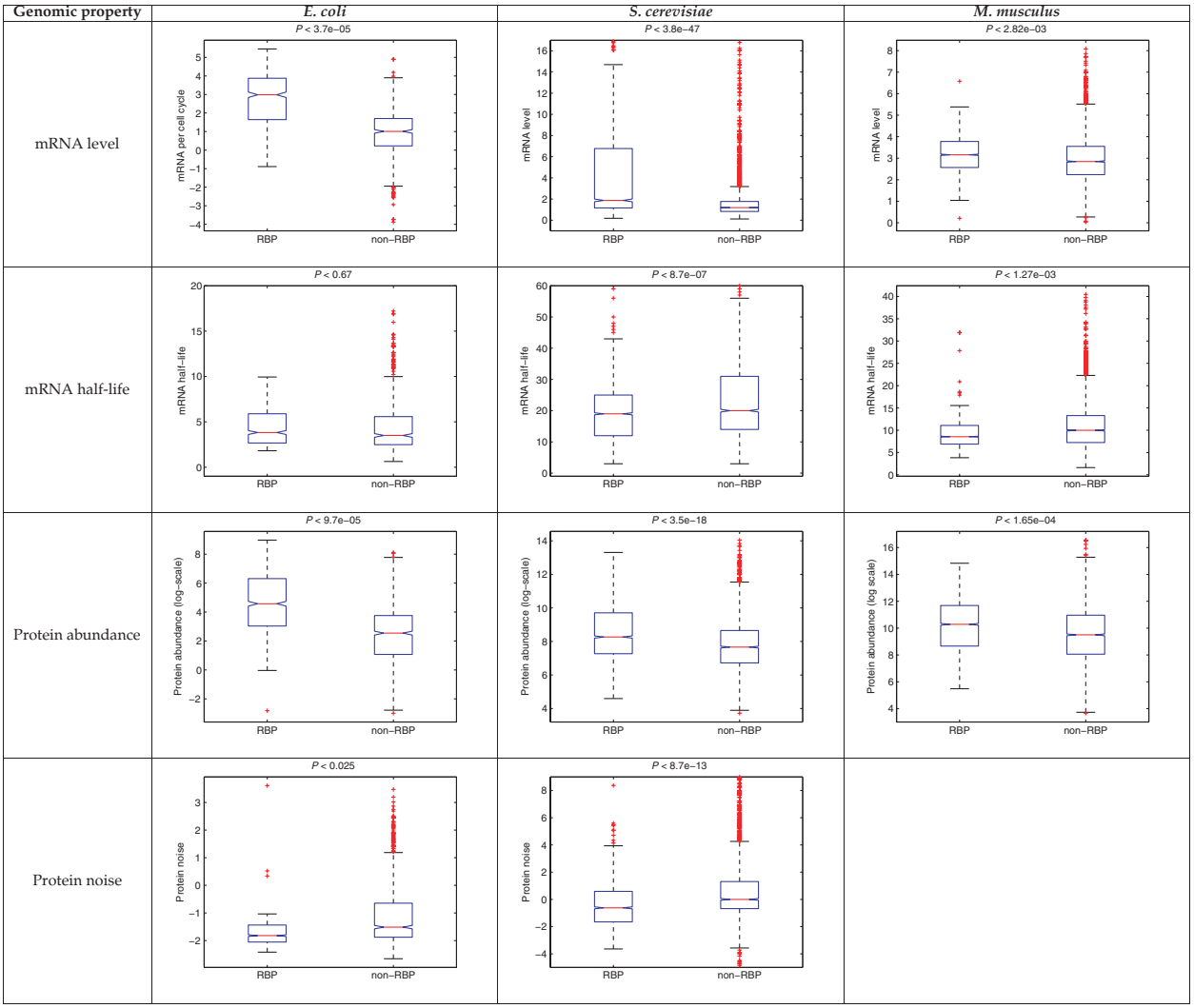


Fig. 1. Comparison of RBP expression dynamics versus non-RBPs across multiple species. The *P*-value in each figure corresponds to a Wilcoxon rank-sum test.

Role of RNA-binding proteins in post-transcriptional regulatory networks

If RBPs play a central role in the reduction of transcriptional noise, they do so through an extensive network of post-transcriptional regulatory interactions. In yeast, a conservative estimate is that post-transcriptional noise reduction affects a quarter of all proteins [9] and that 70% of the transcriptome has significant associations with at least one of 40 tested RBPs [12]. Extensive auto-regulation and feedback are present within the subnetwork of RBP–RBP post-transcriptional interactions [17] and specific post-transcriptional regulatory network motifs are associated with post-transcriptionally regulated co-expression clusters [7]. In the present study, we assembled networks of RBP–target interactions for yeast, mouse and human and integrated these with TF–target interactions (Table 1). We performed a comparative network motif analysis between yeast, mouse and human for composite motifs with a potential role in post-transcriptional noise reduction.

We first analyzed mixed bifan motifs, where a TF and RBP jointly regulate one or more target genes (Fig. 2, first row). Target genes in mixed bifans are co-regulated at transcriptional and post-transcriptional levels and, in yeast, such gene sets overlap significantly with post-transcriptionally regulated expression clusters [7]. In the context of RNA regulon theory [13], we expect that specific TFs and RBPs operate together

Table 1. Regulatory network statistics in four organisms showing, for each interaction type, the number of regulators, the number of targets and the number of edges.

Organism	Regulators	Targets	Edges
TF			
E. coli	195	1583	3909
Yeast	210	3416	10 403
Mouse	125	20 713	126 470
Human	131	20 513	114 818
RBP			
Yeast	44	3879	8881
Mouse	7	8867	15 477
Human	19	15 249	96 919
ncRNA			
E. coli	65	221	331
Yeast	24	86	130
Mouse	148	5720	11 600
Human	165	6561	12 977

and mixed bifans are not equally distributed over all possible (TF, RBP)-pairs. To test this hypothesis, we generated randomized networks preserving the incoming and outgoing degree distributions of each network (and consequently also preserving the total number of mixed bifans) and compared the number of (TF, RBP)-pairs that share common targets in the real versus the randomized networks. In yeast, the number of real coregulating (TF, RBP)-pairs is significantly lower than expected by chance (Fig. 2, first row), supporting the notion that RBPs often partner with specific TFs. In mouse and humans, the difference is not significant (Fig. 2, first row). However, although in yeast RBP-targets are known for 44 RBPs coming from different families [12], the number of RBPs with known targets is much smaller in mouse and humans (Table 1). Moreover, in humans, seven of the 19 RBPs originate from just two protein families, the eukaryotic translation initiation factor 2C family and the insulin-like

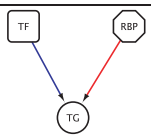
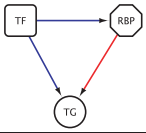
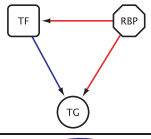
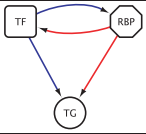
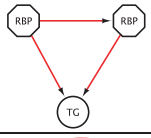
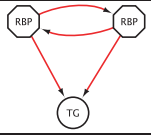
Motif	Yeast			Mouse			Human		
	N_m	Z_e	P	N_m	Z_e	P	N_m	Z_e	P
	3,368‡	–5.485	<0.001†	685‡	–0.294	0.366†	2,151‡	0.321	0.612†
	503	2.408	0.009	7,550	–0.074	0.527	39,377	1.903	0.032
	2,302	–1.356	0.901	9,145	0.632	0.288	88,352	–0.364	0.686
	97	1.111	0.129	2,326	1.597	0.068	9,928	1.140	0.130
	4,868	1.369	0.081	8,245	39.679	< 0.001	674,298	2.532	< 0.001
	992	3.382	0.002	—			299,103	2.118	0.004

Fig. 2. Comparative analysis of post-transcriptional regulatory network motifs in integrated networks of TF–target (blue) and RBP–target (red) interactions in yeast, mouse and humans. N_m, number of motif instances; Z_e, enrichment Z-score; P, enrichment P-value. #N_m is the number of (TF,RBP)-pairs sharing at least one target. †P is defined with respect to the left tail of the distribution (for details, see Materials and methods). '—', no such motifs were present in the real networks.

growth factor 2 mRNA binding protein (IGF2BP) family. We computed the (TF, RBP)-pairs that share significantly more common targets in real than in randomized networks, and found that it is still true that a given RBP cooperates with only few TFs but some TFs cooperate with all RBPs in these protein families, explaining the nonsignificance of the mixed bifan motif. We therefore hypothesize that the significance of the mixed bifan is not specific to yeast but, instead, represents a general mechanism where TFs cooperate with specific RBPs (or RBP families) to regulate gene expression in a co-ordinated way. More RBP binding data will be needed to confirm this hypothesis.

We tested whether (TF, RBP)-pairs in human with a significantly high number of common targets between them regulate functionally coherent gene sets by performing network motif clustering [18,19] aiming to identify densely overlapping bifan structures. Mixed bifan clusters are often enriched for basic metabolic processes, as well as for more specific processes, such as neurogenesis, the cell cycle, protein localization and transport, cell morphogenesis, cellular component organization, and chromatin assembly, amongst others (see the Supporting information, Table S16), corresponding to the functions of known RNA regulons described in the literature [13].

Special cases of mixed bifans are those where there is an additional regulatory interaction between the TF and RBP, either transcriptional, post-transcriptional or both (Fig. 2, rows 2–4). These represent feedforward loops whose enrichment (i.e. a higher than expected number of motifs in real compared to randomized networks) has been associated with functional importance in regulatory networks [20]. The composite feedforward loop with a master transcriptional regulator (Fig. 2, row 2) is consistently enriched in yeast and humans. A configuration where a post-transcriptional regulator is transcriptionally co-regulated with targets of a TF can act as a noise buffer. Any fluctuation in TF levels will affect both the target mRNA production and its subsequent post-transcriptional regulation by the RBP, such that an increased or decreased production rate of the mRNA could be countered by a corresponding decreased or increased translation rate [21]. Such noise-buffering could be particularly effective if the RBP also regulates the TF in a feedback loop. These feedforward–feedback motifs are indeed more abundant in real than in random networks, although the difference is not significant at the $P < 0.05$ level (Fig. 2, row 4).

Apart from the mixed bifan, a second network motif has been associated with post-transcriptionally regulated co-expression clusters in yeast, namely the post-transcriptional feedback loop, which consists of two mutually regulating RBPs sharing a set of target genes (Fig. 2, last row) [7]. This motif is also highly abundant and strongly enriched in humans. (In mouse, no feedback loops between the seven RBPs were present.) Indeed, the network of RBP–RBP post-transcriptional regulatory interactions in humans has a very high density with abundant auto-regulatory and two-component feedback loops (Fig. 3). This is consistent with findings in detailed, low-throughput experiments [22] and analyses of high-throughput data in yeast [7,12,17], as well as with the fact that RBPs are themselves post-transcriptionally regulated to suppress noise in their protein levels. We performed network motif clustering aiming to identify gene clusters regulated by post-transcriptional feedback loops in humans. Although many functions that are enriched for targets of mixed bifans are also enriched in feedback loop clusters, an additional set of strongly enriched categories are 'regulation of …' categories, such as regulation of metabolic processes, regulation of the cell cycle and regulation of gene expression (see Supporting information, Table S18). This is in line with a recently proposed concept where RBPs act as 'regulators of regulators' [23]. Feedback loops are prototypical noise suppressing motifs in regulatory networks [24] and, from a noise management perspective, it makes sense

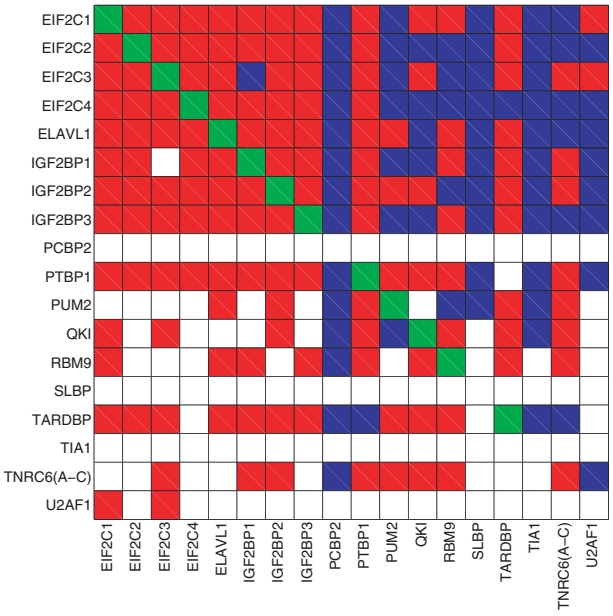


Fig. 3. RBP-RBP post-transcriptional regulatory interactions in human. A coloured cell in position (*i*, *j*) indicates an interaction RBP(*i*) → RBP(*j*). Green cells indicate autoregulatory loops; red cells interactions participating in two-component feedback loops; blue cells indicate all other interactions.

to use such motifs to minimize the noise of key regulators of downstream cellular processes.

Finally, we observed that the simple post-transcriptional feedforward loop (Fig. 2, row 5) is strongly enriched in mouse and humans, whereas the difference between real and random networks is not significant at the $P < 0.05$ level in yeast. However, because each feedback loop (Fig. 2, row 6) itself contains two feedforward loops, it follows that almost all feedforward loops in human originate from such feedback loops. Given the small number of RBPs in the mouse, it is therefore unclear whether the feedforward loop enrichment reflects a distinct mechanism from the feedback loop or not.

Role of noncoding small RNAs in post-transcriptional regulatory networks

In addition to RBPs, cells possess another mechanism for post-transcriptional regulation, namely base-pairing of small noncoding RNAs (ncRNAs) or micro-RNAs (miRNAs) to mRNAs, usually to inhibit translation [25]. Noncoding RNAs can filter transcriptional noise by reducing the effective lifetime of their target mRNAs [26] or by acting in concert with TFs via mixed feedforward and feedback loops [21,27–29]. Because they are not translated, they significantly shorten the response delay and more effectively buffer noise compared to regulation by TFs only [21,26].

We assembled networks of ncRNA–target interactions in *E. coli*, yeast, mouse and humans and miRNA–target interactions in mouse and humans and integrated these with TF–target interactions (Table 1). Similar to the RBP network, mixed bifans are not uniformly distributed over all (TF, ncRNA)-pairs, implying specificity in cooperativity between TFs and ncRNAs, which is found consistently in all of the organisms studied (Fig. 4, first row).

Although enrichment of composite feedforward loops between TFs and miRNAs in mammals has been reported in the literature [29,30], we only found significant enrichment of the corresponding ncRNA-regulated feedforward and feedback loops in *E. coli* (Fig. 4, rows 3 and 4). However, inconsistencies between the presence or absence of enrichment for these particular motifs have previously been observed between different miRNA target prediction methods [30] or between mouse and human datasets [29], indicating the need for more and better curated interaction datasets.

Co-operativity between transcription factors, RNA-binding proteins and small RNAs

RNA-binding proteins are involved in many human diseases [31] and several recent studies have shown that there is a complex interplay between regulation by RBPs and miRNAs [32–34]. We therefore analyzed the co-operativity between all three levels of regulation (by TFs, RBPs and miRNAs) in the human regulatory networks. We calculated all (TF, RBP), (TF, ncRNA) and (RBP, ncRNA) pairs that share significantly more common targets in the real compared to the randomized networks. Specificity of co-operativity at all three levels

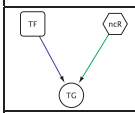
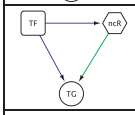
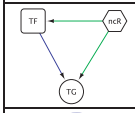
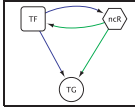
Motif	*E. coli*			Yeast			Mouse			Human		
	N_m	Z_e	P	N_m	Z_e	P	N_m	Z_e	P	N_m	Z_e	P
	455‡	–4.250	<0.001†	122‡	–6.770	<0.001†	13,398‡	–4.191	<0.001†	13,620‡	–5.905	<0.001†
	30	0.343	0.381	—			4663	1.334	0.085	3,043	0.469	0.316
	26	2.628	0.009	—			57	–1.929	0.982	132	–2.384	0.998
	14	8.367	0.002	—			7	0.513	0.259	15	0.599	0.244

Fig. 4. Comparative analysis of post-transcriptional regulatory network motifs in integrated networks of TF–target (blue) and ncRNA–target (green) interactions in *E. coli*, yeast, mouse and humans. N_m, number of motif instances; Z_e, enrichment *Z*-score; *P*, enrichment *P*-value. ‡N_m is the number of (TF,ncRNA)-pairs sharing at least one target. †*P* is defined with respect to the left tail of the distribution (for details, see Materials and Methods). '—', no such motifs were present in the real networks.

is reflected by the sparsity of the corresponding cooperativity network (Fig. 5A). This network has an approximate scale-free degree distribution (Fig. 5B), implying that the majority of regulators share significant targets only with a limited set of partners at the other regulatory levels (i.e. have low connectivity), whereas there are only a few hub regulators with many partners (i.e. have high connectivity). The cooperativity network contains several triples of a TF, RBP and miRNA, which all share significantly many targets with each other (Fig. 5C), some of which are indeed known to function together in key cellular disease processes.

The RNA-binding protein ELAVL1 (also called HuR) and the transcription factor HIF-1 are parts of the vascular endothelial growth factor (VEGF) pathway that plays a crucial role in tumour growth and angiogenesis [35]. VEGF is a direct target of the miRNA miR200b [36] and miR200b/c, which are induced during ischaemia (i.e. a restriction in blood supply) [37]. In our networks, ELAVL1, HIF-1 and miR200b/c all share more targets with each other than is expected by chance. VEGFA, one of the key genes in the VEGF pathway is a common target of ELAVL1 and HIF-1. In agreement with a role of the VEGF pathway in cancer [38], the genes regulated by at least two regulators of the (HIF-1, ELAVL1, miR200b/c) triple are over-represented for pancreatic cancer ($P < 0.03$) and chronic myeloid leukaemia ($P < 0.08$).

The transcription factor MYC, RNA-binding protein IGF2BP1 and miRNA miR26a all play a crucial role in apoptosis. On activation of β-catenin signalling, IGF2BP1 increases levels of MYC by targeting a recognition motif in the protein coding region [33]. MYC in turn acts as a repressor of miRNAs associated with tumour suppressor activity, such as miR26a [39]. The common targets of MYC and IGF2BP1 are enriched for cellular response to stress ($P < 10^{-3}$), response to DNA damage ($P < 10^{-3}$) and DNA repair ($P < 0.005$).

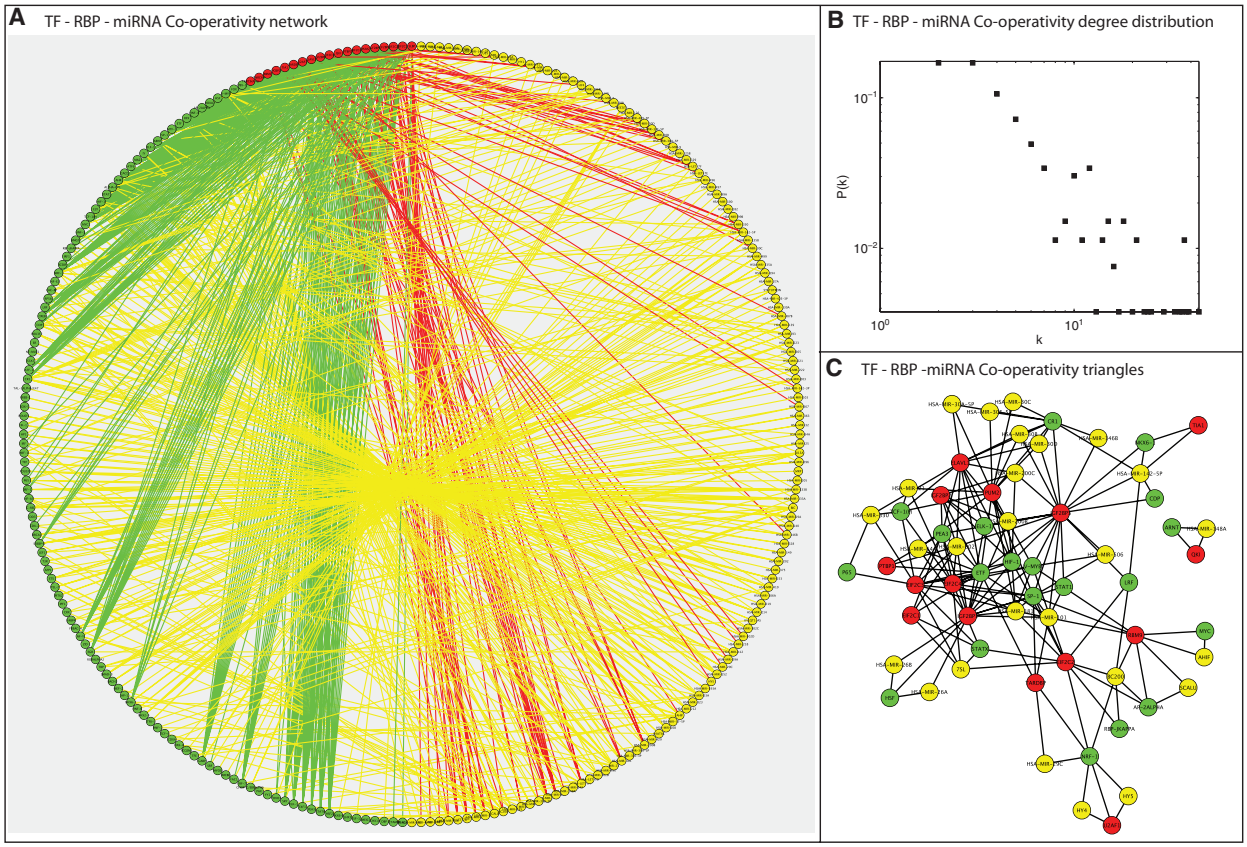


Fig. 5. Co-operativity specificity between (TF, RBP), (RBP, miRNA) and (TF, miRNA) pairs in humans. (A) Co-operativity network with edges connecting all pairs with a significantly higher number of common targets than expected by chance ($Z \geq 2.5$), with TFs in green, RBPs in red and miRNAs in yellow, and TF–RBP links in green, RBP–miRNA in red and TF–miRNA in yellow. (B) Degree distribution of the co-operativity network on a logarithmic scale. (C) Subgraph of the co-operativity network (with uncoloured edges) showing all fully connected (TF, RBP, miRNA) triangles.

Yet another example is given by the RNA-binding protein TNRC6 and the cell cycle transcription factor E2F1. Of their 127 common targets, 12 (10%) are periodically expressed ($P < 10^{-4}$) [40] and the whole set is enriched for cell cycle genes ($P < 10^{-5}$). Both miRNA 148b and 152 share targets with E2F1 and are enriched for periodic genes.

Discussion

Post-transcriptional regulation of mRNA transcript stability and translation by RNA-binding proteins is a key step in the regulation of gene expression, with important implications for development and disease. In the eukaryotic model organism *S. cerevisiae*, mounting evidence from genome-wide studies shows that a considerable fraction of the proteome is under the influence of post-transcriptional regulation, resulting in a reduced noise of protein expression levels compared to transcript levels. RNA-binding proteins, which are themselves tightly regulated at the post-transcriptional level, likely play an important role in this noise-buffering process. In the present study, we performed a comparative analysis (between four species) of the expression dynamics of RNA-binding proteins and their role in the integrated networks of transcriptional and post-transcriptional regulatory interactions.

We found that RNA-binding proteins in *E. coli*, yeast and mouse are expressed at significantly higher mRNA and protein levels, with a significantly lower mRNA half-life and protein noise, compared to non-RBP proteins. The fact that RNA-binding proteins are regulated to have minimal levels of protein noise supports the hypothesis that they are crucial regulators for the noise management of downstream cellular processes.

We further tested this hypothesis by combining DNA-binding data for multiple transcription factors with mRNA-binding data for multiple RNA-binding proteins, resulting in integrated regulatory networks for yeast, mouse and humans. Although the number of RNA-binding proteins with known binding targets concerns only a small fraction of the hundreds to thousands of known or predicted RNA-binding proteins, we found conserved organizational principles in these networks, which were shared between multiple species. In particular, composite feedforward loops, where an RNA-binding protein and its targets are co-regulated at transcriptional level, are significantly enriched in the real networks compared to suitably randomized networks, with $\sim$ 20–25% of these feedforward loops containing an additional feedback loop from the RNA-binding protein to the transcription factor. This type of network motif, where a post-transcriptional regulator can counter the effect of stochastic fluctuations in the activity level of a transcription factor by directly sensing the same fluctuations as the targets of the transcription factor, is indeed an effective noise-buffering network configuration. Co-operativity between RNA-binding proteins and transcription factors tends to be specific, as demonstrated by the fact that only a few (TF, RBP)-pairs share more targets than would be expected by chance. Targets co-regulated by such (TF, RBP)-pairs form so-called RNA regulons and are enriched for a variety of cellular functions, including neurogenesis, the cell cycle and chromatin assembly, amongst others.

Another network motif with a potential noise-buffering role consists of two RNA-binding proteins that mutually regulate each other in a feedback loop and share a significant number of common targets. Such feedback loops have previously been described in yeast, where they are associated with post-transcriptionally regulated co-expression clusters, and have been confirmed in humans as well in the present study. These two-component feedback loops are part of a dense network of RBP–RBP post-transcriptional regulatory interactions, with abundant auto-regulation, and their common targets are frequently involved in functions such as regulation of metabolism, regulation of the cell cycle or regulation of gene expression.

Taken together, our results point to a model where RNA-binding proteins are tightly regulated to minimize noise in their protein expression levels. The cell then exploits the extraordinary stability of RNA-binding proteins to buffer stochastic noise in downstream processes; on the one hand, by letting them partner with specific transcription factors in noise-reducing composite feedforward and feedback loops and, on the other hand, by letting them act as the regulators of regulators, including RNA-binding proteins themselves through extensive feedback mechanisms.

The tight regulation of RNA-binding proteins at the protein level puts them in contrast with small, ncRNAs, which exert their post-transcriptional regulatory function without being translated into protein. Similar to RNA-binding proteins, ncRNAs co-operate with specific transcription factors and, in several cases, these (TF, ncRNA)-pairs regulate each other in composite feedforward and feedback loops, although no network motifs were found to be consistently enriched in all of the species studied. It has been proposed that regulation by ncRNAs leads to faster response times compared to regulation by proteins as a result of the omission of an additional translation step. This makes it attractive to speculate on a model where ncRNAs act as the fast responders to fluctuating external signals and

RNA-binding proteins as a result of their own tight regulation at the protein level, comprising the slower but more effective maintainers of low overall noise levels for key cellular processes. This model, however, ignores the fact that RNA-binding proteins also play a crucial role in, for example, the biogenesis of microRNAs. Furthermore, we found several examples of (TF, RBP, ncRNA)-triples that share more targets than expected by chance and are jointly involved in the regulation of important disease pathways, suggesting that post-transcriptional regulation by RNA binding proteins and ncRNAs is neither redundant, nor disjoint of each other. The need to disentangle the different modes of post-transcriptional regulation of gene expression and also to understand the role of protein noise-buffering in disease and other processes thus remains an important challenge, both from an experimental and a computational point of view.

Materials and methods

Data sources

Data for *E. coli*

We obtained RNA-binding proteins from orthologous genes to yeast RBPs using OrthoMCL [41], genomic properties [42], TF–target interactions [43] (see Supporting information, Table S1) and ncRNA–target interactions [27,44,45] (see Supporting information, Table S2).

Data for *S. cerevisiae*

We used the same list of RBPs as that described previously [10], as well as previously described mRNA level and half-life data [46], protein abundance data [47], protein noise data [14], TF–target interactions [3] (see Supporting information, Table S3), RBP–target interactions [12] (see Supporting information, Table S4) and ncRNA–target interactions [44] (see Supporting information, Table S5).

Data for mouse and humans

Mouse RNA-binding proteins [48] and genomic properties [49] were as described previously. Mouse and human TF–target and miRNA–target networks [50,51] (see Supporting information, Tables S6, S8, S9 and S11) and RBP–target networks [52,53] (see Supporting information, Tables S7 and S10) were also as described previously.

Network motif calculations

We created a list of three-node post-transcriptional regulatory network motifs consisting of two regulators, (at least one post-transcriptional) and one target gene in the integrated networks of TF–target and RBP-target interactions (Fig. 2) and TF–target and ncRNA–target interactions (Fig. 4). Network motif enrichment was calculated by comparing the true number of motif instances with the number of instances in randomized networks that preserve the number of incoming and outgoing interactions for each type for each node. The enrichment Z-score is defined as:

$$Z_e = (N_{\mathrm{m}} - \mu)/\sigma$$

where N_{m} is the number of motif instances in the real network and μ and σ are the mean and standard deviation, respectively, of the number of motif instances over 1000 independently sampled random networks. The enrichment P-value is defined with respect to the right tail of the distribution (i.e. it is the fraction of random networks having at least N_{m} motif instances).

For the mixed bifan (Figs 2 and 4, first row), whose total number is preserved under degree-preserving randomizations, we defined N_{m} as the number of TF–RBP, and TF–ncRNA pairs, respectively, which share at least one target in the real networks and compared this with the number of such pairs in the randomized networks. The enrichment P-value is defined in this case with respect to the left tail of the distribution (i.e. it is the fraction of random networks having at most N_{m} co-operating pairs).

For a given input motif, network motif clustering identifies sets of nodes with a high density of motif instances between them, and was carried out as described previously [19]. Enrichment calculations and motif clustering were performed using the Network Motif Clustering Toolbox for MATLAB (http://omics.frias.uni-freiburg.de/software).

Co-operativity *Z*-scores

Co-operativity Z-scores for a given regulator pair were calculated by comparing their real number of common targets with the mean number of common targets of that pair in degree-preserving randomized networks.

Functional enrichment analysis

Functional enrichment analysis of network motif clusters with respect to GO functional classes was carried out using BINGO in batch mode [54]. Disease enrichment P-values were calculated by reference to the OMIM database (http://www.ncbi.nlm.nih.gov/omim) using a hypergeometric test.

References

1 Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D & Brown PO (2000) Genomic expression programs in the response of yeast

cells to environmental changes. *Mol Biol Cell* **11**, 4241–4257.
2 Wang Z, Gerstein M & Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**, 57–63.
3 Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, Hannett NM, Tagne JB, Reynolds DB, Yoo J *et al.* (2004) Transcriptional regulatory code of a eukaryotic genome. *Nature* **431**, 99–104.
4 Park PJ (2009) ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet* **10**, 669–680.
5 Lu P, Vogel C, Wang R, Yao X & Marcotte EM (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat Biotechnol* **25**, 117–124.
6 Halbeisen RE & Gerber AP (2009) Stress-dependent coordination of transcriptome and translatome in yeast. *PLoS Biol* **7**, e105.
7 Joshi A, Van de Peer Y & Michoel T (2011) Structural and functional organization of RNA regulons in the posttranscriptional regulatory network of yeast. *Nucleic Acids Res* **39**, 9108–9117.
8 Foss EJ, Radulovic D, Shaffer SA, Goodlett DR, Kruglyak L & Bedalov A (2011) Genetic variation shapes protein networks mainly through non-transcriptional mechanism. *PLoS Biol* **9**, e1001144.
9 Lee MV, Topper SE, Hubler SL, Hose J, Wenger CD, Coon JJ & Gasch AP (2011) A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. *Mol Syst Biol* **7**, 514.
10 Mittal N, Roy N, Madan Babu M & Janga SC (2009) Dissecting the expression dynamics of RNA-binding proteins in posttranscriptional regulatory networks. *PNAS* **106**, 20300–20305.
11 Glisovic T, Bachorik JL, Yong J & Dreyfuss G (2008) RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett* **582**, 1977–1986.
12 Hogan DJ, Riordan DP, Gerber AP, Herschlag D & Brown PO (2008) Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. *PLoS Biol* **6**, e255.
13 Keene JD (2007) RNA regulons: coordination of post-transcriptional events. *Nat Rev Genet* **8**, 533–543.
14 Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL & Weissman JS (2006) Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441**, 840–846.
15 Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, Pilpel Y & Barkai N (2006) Noise in protein expression scales with natural protein abundance. *Nat Genet* **38**, 636–643.
16 Paulsson J (2005) Models of stochastic gene expression. *Phys Life Rev* **2**, 157–175.
17 Mittal N, Scherrer T, Gerber AP & Chandra Janga S (2011) Interplay between posttranscriptional and posttranslational interactions of RNA-binding proteins. *J Mol Biol* **409**, 466–479.
18 Audenaert P, Van Parys T, Pickavet M, Demeester P, Van de Peer Y & Michoel T (2011) CyClus3D: a Cytoscape plugin for clustering network motifs in integrated networks. *Bioinformatics* **27**, 1587–1588.
19 Michoel T, Joshi A, Nachtergaele B & Van de Peer Y (2011) Enrichment and aggregation of topological network motifs are independent organizational principles of integrated interaction networks. *Mol BioSyst* **7**, 2769–2778.
20 Alon U (2007) Network motifs: theory and experimental approaches. *Nat Rev Genet* **8**, 450–461.
21 Tsang J, Zhu J & van Oudenaarden A (2007) MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Mol Cell* **26**, 753–767.
22 Pullmann R Jr, Kim HH, Abdelmohsen K, Lal A, Martindale JL, Yang X & Gorospe M (2007) Analysis of turnover and translation regulatory RNA-binding protein expression through binding to cognate mRNAs. *Mol Cell Biol* **27**, 6265.
23 Kanitz A & Gerber AP (2010) Circuitry of mRNA regulation. *Wiley Interdiscip Rev Syst Biol Med* **2**, 245–251.
24 Bruggeman FJ, Blüthgen N & Westerhoff HV (2009) Noise management by molecular networks. *PLoS Comp Biol* **5**, e1000506.
25 Estelle M (2011) Non-coding RNAs in human disease. *Nat Rev Genet* **12**, 861–874.
26 Mehta P, Goyal S & Wingreen NS (2008) A quantitative comparison of sRNA-based and protein-based gene regulation. *Mol Syst Biol* **4**, 221.
27 Shimoni Y, Friedlander G, Hetzroni G, Niv G, Altuvia S, Biham O & Margalit H (2007) Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol Syst Biol* **3**, 138.
28 Osella M, Bosia C, Cor'a D & Caselle M (2011) The role of incoherent microRNA-mediated feedforward loops in noise buffering. *PLoS Comp Biol* **7**, e1001101.
29 Cheng C, Yan KK, Hwang W, Qian J, Bhardwaj N, Rozowsky J, Lu ZJ, Niu W, Alves P, Kato M *et al.* (2011) Construction and analysis of an integrated regulatory network derived from high-throughput sequencing data. *PLoS Comp Biol* **7**, e1002190.
30 Shalgi R, Lieber D, Oren M & Pilpel Y (2007) Global and local architecture of the mammalian microRNA-transcription factor regulatory network. *PLoS Comp Biol* **3**, e131.
31 Lukong KE, Chang K, Khandjian EW & Richard S (2008) RNA-binding proteins in human genetic disease. *Trends Genet* **24**, 416–425.
32 Jacobsen A, Wen J, Marks DS & Krogh A (2010) Signatures of rna binding proteins globally coupled to

effective microrna target sites. *Genome Res*, **20** 1010–1019.
33 van Kouwenhove M, Kedde M & Agami R (2011) Microrna regulation by rna-binding proteins and its implications for cancer. *Nat Rev Cancer* **11**, 644–656.
34 Chang SH & Hla T (2011) Gene regulation by rna binding proteins and micrornas in angiogenesis. *Trends Mol Med* **17**, 650–658.
35 Hicklin DJ & Ellis LM (2005) Role of the vascular endothelial growth factor pathway in tumor growth and angiogenesis. *J Clin Oncol* **23**, 1011.
36 McArthur K, Feng B, Wu Y, Chen S & Chakrabarti S (2011) MicroRNA-200b regulates vascular endothelial growth factor-mediated alterations in diabetic retinopathy. *Diabetes* **60**, 1314.
37 Lee ST, Chu K, Jung KH, Yoon HJ, Jeon D, Kang KM, Park KH, Bae EK, Kim M, Lee SK *et al.* (2010) MicroRNAs induced during ischemic preconditioning. *Stroke* **41**, 1646–1651.
38 Ellis LM & Hicklin DJ (2008) Vegf-targeted therapy: mechanisms of anti-tumour activity. *Nat Rev Cancer* **8**, 579–591.
39 Bui TV & Mendell JT (2010) Myc, maestro of microRNAs. *Genes Cancer* **1**, 568.
40 Bar-Joseph Z, Siegfried Z, Brandeis M, Brors B, Lu Y, Eils R, Dynlacht BD & Simon I (2008) Genome-wide transcriptional analysis of the human cell cycle identifies genes differentially regulated in normal and cancer cells. *Proc Nat Acad Sci* **105**, 955.
41 Li L, Stoeckert CJ & Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* **13**, 2178–2189.
42 Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A & Xie XS (2010) Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533–538.
43 Salgado H, Gama-Castro S, Peralta-Gil M, Diaz-Peredo E, Sanchez-Solano F, Santos-Zavaleta A, Martinez-Flores I, Jimenez-Jacinto V, Bonavides-Martinez C, Segura-Salazar J *et al.* (2006) RegulonDB (version 5.0): *Escherichia coli* K-12 transcriptional regulatory network, operon organization, and growth conditions. *Nucleic Acids Res* **34**, D394–D397.
44 Wu T, Wang J, Liu C, Zhang Y, Shi B, Zhu X, Zhang Z, Skogerb G, Chen L, Lu H *et al.* (2006) NPInter: the noncoding RNAs and protein related biomacromolecules interaction database. *Nucleic Acids Res* **34**, D150–D152.
45 Cao Y, Wu J, Liu Q, Zhao Y, Ying X, Cha L, Wang L & Li W (2010) sRNATarBase: a comprehensive database of bacterial sRNA targets verified by experiments. *RNA* **16**, 2051–2057.
46 Wang Y, Liu CL, Storey JD, Tibshirani RJ, Herschlag D & Brown PO (2002) Precision and functional specificity in mRNA decay. *PNAS* **99**, 5860–5865.
47 Ghaemmaghami S, Huh W-K, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK & Weissman JS (2003) Global analysis of protein expression in yeast. *Nature* **425**, 737–741.
48 McKee AE, Minet E, Stern C, Riahi S, Stiles CD & Silver PA (2005) A genome-wide in situ hybridization map of RNA-binding proteins reveals anatomically restricted expression in the developing mouse brain. *BMC Dev Biol* **5**, 14.
49 Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W & Selbach M (2011) Global quantification of mammalian gene expression control. *Nature* **473**, 337–342.
50 Friard O, Re A, Taverna D, De Bortoli M & Cora D (2010) CircuitsDB: a database of mixed microRNA/transcription factor feed-forward regulatory circuits in human and mouse. *BMC Bioinformatics* **11**, 435.
51 Dweep H, Sticht C, Pandey P & Gretz N (2011) miRWalk–database: prediction of possible miRNA binding sites by 'walking' the genes of three genomes. *J Biomed Inform* **44**, 839–847.
52 Yang JH, Li JH, Shao P, Zhou H, Chen YQ & Qu LH (2011) starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. *Nucleic Acids Res* **39**, D202–D209.
53 Halbeisen RE, Galgano A, Scherrer T & Gerber AP (2008) Post-transcriptional gene regulation: from genome-wide studies to principles. *Cell Mol Life Sci* **65**, 798–813.
54 Maere S, Heymans K & Kuiper M (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**, 3448–3449.

Supporting information

The following supplementary material is available:
Table S1. *E. coli* TF network.
Table S2. *E. coli* ncRNA network.
Table S3. Yeast TF network.
Table S4. Yeast RBP network.
Table S5. Yeast ncRNA network.
Table S6. Mouse TF network.
Table S7. Mouse RBP network.
Table S8. Mouse ncRNA network.
Table S9. Human TF network.
Table S10. Human RBP network.
Table S11. Human ncRNA network.
Table S12. Human TF-RBP co-operativity *Z*-scores.
Table S13. Human TF-ncRNA co-operativity *Z*-scores.
Table S14. Human RBP-ncRNA co-operativity *Z*-scores.

Table S15. Human TF-RBP bifan clusters.
Table S16. Human TF-RBP bifan clusters, functional enrichment of the target sets.
Table S17. Human RBP feedback loop clusters.
Table S18. Human RBP feedback loop clusters, functional enrichment of the target sets.

This supplementary material can be found in the online version of this article.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be reorganized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.