BMC
Systems Biology

RESEARCH ARTICLE

Crosstalk between transcription factors and microRNAs in human protein interaction network

Chen-Ching Lin[1†], Ya-Jen Chen[1†], Cho-Yi Chen[2,3], Yen-Jen Oyang[1], Hsueh-Fen Juan[1,4*] and Hsuan-Cheng Huang[5*]

Abstract

Background: Gene regulatory networks control the global gene expression and the dynamics of protein output in living cells. In multicellular organisms, transcription factors and microRNAs are the major families of gene regulators. Recent studies have suggested that these two kinds of regulators share similar regulatory logics and participate in cooperative activities in the gene regulatory network; however, their combinational regulatory effects and preferences on the protein interaction network remain unclear.

Methods: In this study, we constructed a global human gene regulatory network comprising both transcriptional and post-transcriptional regulatory relationships, and integrated the protein interactome into this network. We then screened the integrated network for four types of regulatory motifs: single-regulation, co-regulation, crosstalk, and independent, and investigated their topological properties in the protein interaction network.

Results: Among the four types of network motifs, the crosstalk was found to have the most enriched protein-protein interactions in their downstream regulatory targets. The topological properties of these motifs also revealed that they target crucial proteins in the protein interaction network and may serve important roles of biological functions.

Conclusions: Altogether, these results reveal the combinatorial regulatory patterns of transcription factors and microRNAs on the protein interactome, and provide further evidence to suggest the connection between gene regulatory network and protein interaction network.

Background

A gene regulatory network (GRN) is a comprehensive collection of regulatory relationships that controls the global gene expression and the dynamics of protein output in a living cell [1-6]. These regulatory relationships may be derived from different layers in the gene regulatory system. Hence, a GRN can be roughly separated into two major levels: the transcriptional and the post-transcriptional levels.

At the transcriptional level, a class of DNA-binding proteins, known as transcription factors (TFs), plays a major role in regulating gene expression. By binding to specific regions of DNA sequences, TFs can control the transcription activities of target genes, thus regulating

the production of mRNA transcripts [7-9]. Since it has been widely believed that TFs are the primary regulators of gene expression, previous research on GRNs has mainly focused on the regulatory relationships at the transcriptional level [5,10,11]. However, there is increasing evidence suggesting that, at the post-transcriptional level, microRNAs (miRNAs) may also contribute to modulation of gene expression on a large scale [1-3]. miRNAs are small non-coding, single stranded RNAs of ~22 nucleotides in length that are abundantly found in eukaryotic cells [1-3]. By binding to complementary sequences (a.k.a. miRNA binding-sites) on target messenger RNA transcripts (mRNAs), miRNAs can trigger translational repression or gene silencing, thus regulating the expression of their target genes at the post-transcriptional level [12,13]. In recent years, miRNAs have been reported to control many biological processes, such as development, differentiation, growth, and even cancer development and progression [1-3]. Therefore, it has become critical to construct an integrated GRN that

* Correspondence: yukijuan@ntu.edu.tw; hsuancheng@ym.edu.tw
† Contributed equally
[1]Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei 106, Taiwan
[5]Institute of Biomedical Informatics, Center for Systems and Synthetic Biology, National Yang-Ming University, Taipei 112, Taiwan
Full list of author information is available at the end of the article

comprises both transcriptional and post-transcriptional regulatory interactions.

Similar to other biological networks, a GRN usually consists of several types of sub-network patterns known as network motifs, such as feedback and feedforward loops. Previous studies [5,10,11] have shown that certain types of network motifs are more overrepresented in GRNs[14]. These network motifs, such as feedback loops and co-regulation, are found to play pivotal roles in gene regulation [15-17]. For example, in *E. coli*, ~35% TFs participate in negative autoregulation motifs which can significantly speed up the transcriptional response time [15] and smooth the fluctuations of protein expression [16]. In addition to TFs, miRNAs may also form specific network motifs in the GRN. Previous studies [17-20] investigating the co-regulation between miRNAs and TFs found a variety of significant network motifs overrepresented in the co-regulation network, suggesting that the gene regulatory system requires close cooperation between transcriptional and post-transcriptional layers. These studies each proposed that the network motifs might be used as building blocks in GRNs. In order to understand how these motifs in the GRN influence the downstream biological processes, further studies on the protein interactome are essential.

Proteins are the major functional units in living cells, and usually do not work alone. Protein-protein interactions (PPIs), formed by two physically interacting proteins, are fundamental to most biological processes. In addition, proteins are translated from mRNAs, and therefore their abundance may be affected by upstream miRNAs and TFs. Consequently, investigating the correlations between PPIs and their upstream regulators could facilitate the understanding of biological mechanisms within living cells. Recently, the correlations between miRNAs and PPIs have been investigated [21,22]. Liang and Li [21] revealed that proteins regulated by more miRNAs tend to possess higher degree, more interacting partners, in a protein interaction network (PIN). Furthermore, Hsu *et. al.* [22] provided a comprehensive analysis and suggested that miRNAs could influence specific biological processes through regulating a small number of selected proteins in a PIN, such as hub and bottleneck proteins. These studies have revealed some connection principles between upstream regulators and downstream PINs. However, the specifics of the cooperation between TFs and miRNAs and their combinational regulatory effects on human PINs remain unclear.

In this study, we firstly collected human TF and miRNA regulatory relationships and integrated them into a global GRN. Next, we imported the human protein interactome into the GRN and screened the integrated network for four pre-defined regulatory patterns
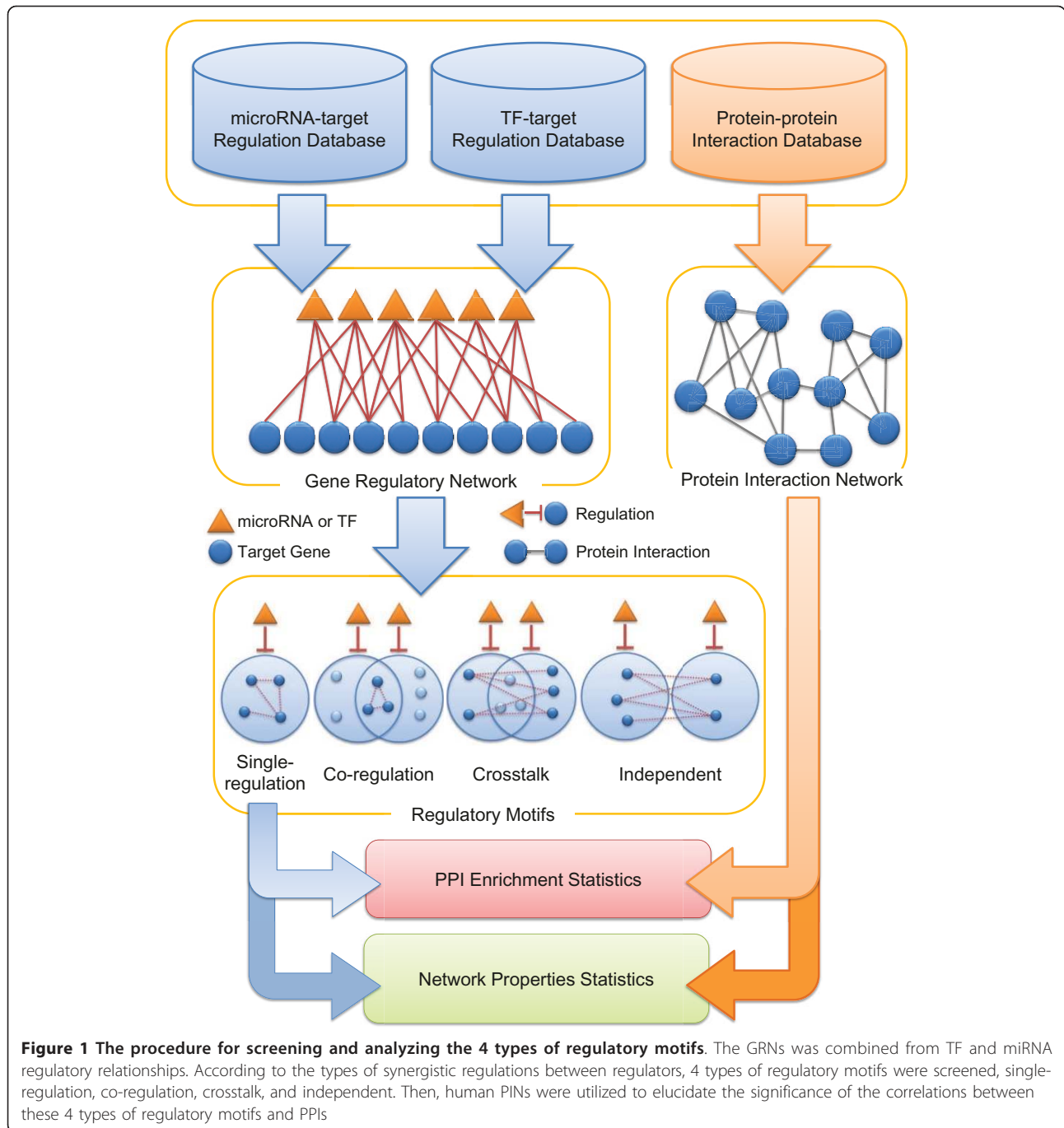
(Figure 1). Among the four patterns, the crosstalk was found to have the most enriched PPIs interconnected in their downstream regulatory target sets. Notably, the observed correlation between PPIs and the crosstalk motif has not been previously reported. Further investigation into the topological properties of the crosstalk motifs also revealed that they might serve important roles of biological functions. We thus propose that the crosstalk motifs may play significant roles in PINs, which may have important downstream effects on several biological processes in living cells via regulating corresponding PPIs.

Methods
Gene regulatory networks and protein-protein interactions
To construct the human GRN for our analysis, we collected TF and miRNA regulatory relationships from three online databases: TRED (Transcriptional Regulatory Element Database) [23], UCSC genome browser at http://genome.ucsc.edu/, and TargetScanHuman (release 6.0, November 2011) [24-27]. TRED contains transcriptional regulation information from experimental evidence and computational prediction. We collected 6,764 transcriptional regulation relationships between 133 TFs and 2,937 target genes from TRED. Additionally, we obtained the conserved binding sites of 125 TFs from UCSC genome browser. To identify the targets of these 125 TFs, the annotations of 21,368 human genes were downloaded from UCSC genome browser. We assigned a target gene to a TF if its promoter region (1000 bp upstream and 500 bp downstream of the transcription start site) covered at least one conserved binding site of the TF [28]. After this process, we identified 52,301 regulations between 125 TFs and 12,383 targets. Then, the union of these two transcriptional regulatory networks from TRED and UCSC was considered as the GRN of transcriptional level in this study, containing 58,711 regulations between 211 TFs and 13,402 targets. For miRNA target prediction programs, previous study had noticed that TargetScan possessed relatively higher precision and sensitivity than other programs [29]. We collected 144,490 post-transcriptional regulatory relationships between 153 miRNA families and 11,161 target genes with conserved binding sites of corresponding miRNAs. Next, these regulatory relationships collected from the three databases were merged together to construct our global GRN, in which nodes represent regulators (TFs/miRNAs) or target genes/proteins, and edges represent the regulatory relationships between regulators and targets.

Human PPI data were obtained from HPRD (Human Protein Reference Database) [30], which contained experimentally validated physical interactions among

Figure 1 The procedure for screening and analyzing the 4 types of regulatory motifs. The GRNs was combined from TF and miRNA regulatory relationships. According to the types of synergistic regulations between regulators, 4 types of regulatory motifs were screened, single-regulation, co-regulation, crosstalk, and independent. Then, human PINs were utilized to elucidate the significance of the correlations between these 4 types of regulatory motifs and PPIs

human proteins. In this study, we collected 37,080 interactions between 9,465 proteins.

Considering the incompleteness of current human PPI data, we performed an analogous analysis with an expanded PIN, a union of BioGRID [31] and HPRD PPI data, to verify our results. Additionally, since limited reproducibility of miRNA target prediction has also been reported [32-35], we further independently repeated our study with another miRNA target

prediction database, miRBase [36], to confirm the robustness of our conclusions.

Regulatory motif screening and analysis

We screened four types of regulatory motifs from GRN: single-regulation, co-regulation, crosstalk, and independent, considering possible synergistic regulation between regulators. These regulatory motifs are depicted in Figure 1. The synergistic regulation defined here is

determined by whether the regulators shared at least two common targets. A single-regulation motif consists of one regulator and its targets. The other three motifs consist of two regulators. The co-regulation motif is formed by two synergistic regulators and their shared targets. The crosstalk motif is formed by two synergistic regulators and their private (non-shared) target sets. The independent motif contains two non-synergistic regulators and their respective target sets.

Next, the PPI enrichment for each type of regulatory motif was analyzed. Specifically, for single-regulation, PPIs between every paired target genes were analyzed; for co-regulation motifs, only PPIs between common target genes were analyzed; and for crosstalk and independent motifs, only PPIs between two private target gene sets were analyzed. Additionally, the PPI enrichment analysis was performed from two directions: top-down and bottom-up. In the bottom-up model, genes were firstly classified into four categories analogous to four types of regulatory motifs, and each category was provided with one significance score. In the top-down model, significance scores were assigned to each regulator (for single-regulation motif) or to every pair of regulators (for the other three types of motifs). In this study, a significance score was defined as the z-score (standard score) derived from statistical analysis (Methods in Additional file 1). Furthermore, we also analyzed the significance of several selected network properties (Methods in Additional file 1) for each type of regulatory motif based on similar approaches adopted in the PPI enrichment analysis. The procedures of regulatory motif screening and analysis are depicted in Figure 1. In addition, the functional enrichment analysis of crosstalk motifs was performed to investigate the underlying biological roles for crosstalk motifs in human PINs (Methods in Additional file 1).

Results
Gene regulatory network properties
In order to provide a global view of human GRN, both transcriptional and post-transcriptional regulations were analyzed jointly (the global GRN) and respectively (local GRNs) in this study. Within the local GRNs, TF- and miRNA-regulation displayed similar patterns of distribution with respect to the number of target genes (Figure 2A and 2B). Most TFs and miRNAs possessed relatively fewer targets, and only a small fraction of TFs and miRNAs possessed a large number of targets.

To investigate the synergistic relationships between regulators, we further analyzed the distributions of the number of synergistic partners of miRNAs and/or TFs (Figure 2C-F). Herein, we defined two regulators as having a synergistic relationship if they shared at least two common targets. Most TFs and miRNAs have at least one synergistic TF and/or miRNA partner. In other words, they tended to form synergistic regulations with other regulators. Although we noticed that a small fraction of TFs did not form synergistic regulations with other TFs or miRNAs (Figure 2D and 2F), this could be due to the lack of sufficient TF-regulation information.

PPI enrichment of regulatory motifs
From the global GRN, we screened four types of regulatory motifs: single-regulation, co-regulation, crosstalk, and independent. Next, the PPI enrichment of regulatory motifs was investigated from two directions: top-down and bottom-up. Based on the combinations of regulators, the regulatory motifs of TF-TF, miRNA-miRNA, and TF-miRNA were analyzed separately.

In the top-down analysis, the significance scores of the motifs regulated by single TF or paired-TF are described in Figure 3A. The single-regulation and crosstalk motifs showed significantly enriched PPIs between regulated genes, while the co-regulation and independent motifs did not. Similarly, motifs regulated by miRNA or paired-miRNA also showed significantly enriched PPIs between regulated genes involved in the single-regulation and crosstalk motifs, but co-regulation and independent did not (Figure 3B). The single-regulation motif has been reported to be highly correlated with PPIs [21,22], which is consistent with our results. However, although correlations between PPIs and the co-regulation motif have also been reported and well-discussed [37,38], our analysis was inconsistent with these findings. For co-regulation motifs, the tested sample is the common targets between two regulators. Thus, we doubted that the insignificance of the co-regulation motifs was due to limited sample sizes of the common targets. To test this hypothesis, we gradually adjusted the threshold of synergistic regulation (i.e. the minimum number of shared targets) for co-regulation motifs. As the threshold increased, the z-scores of PPI enrichment of the co-regulation motifs also increased (Figure S1 in Additional file). This result suggested that the significance scores of the co-regulation motifs were truly affected by the sample size. In other words, the PPI enrichment would emerge if we adopted a stricter definition of co-regulation which means more common targets. For example, if the threshold of the synergistic regulation between TFs increases to 40 targets, the significance score would be elevated to 2. This result suggested that regulator pairs with common targets tend to regulate private targets with PPIs. The motifs regulated by TFs and miRNAs simultaneously were also investigated. Consistent with the results of the TF-TF and miRNA-miRNA analyses, the crosstalk motifs showed a significant correlation with PPI enrichment, while the co-regulation motifs did not (Figure 3C). Similarly, significance scores of the
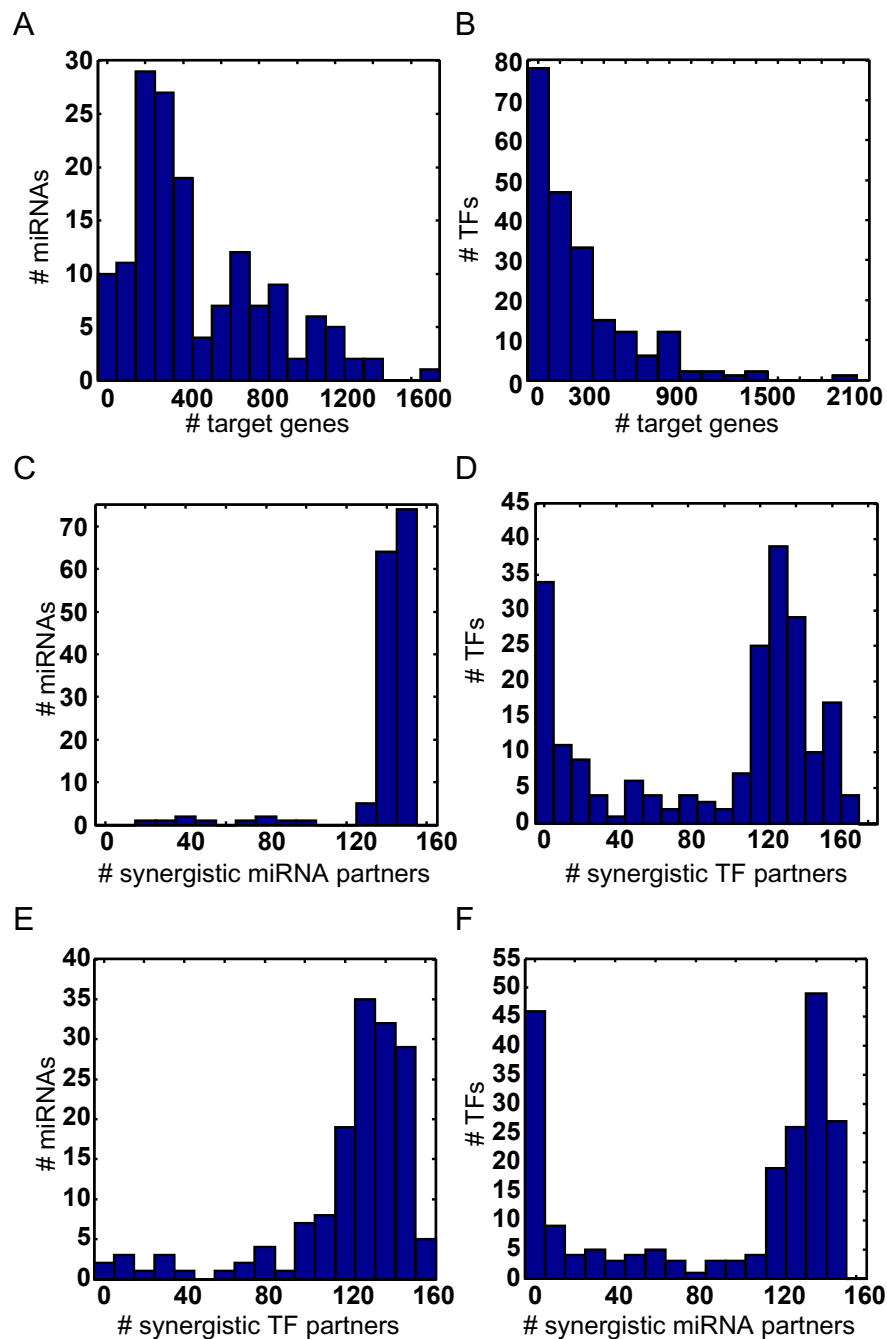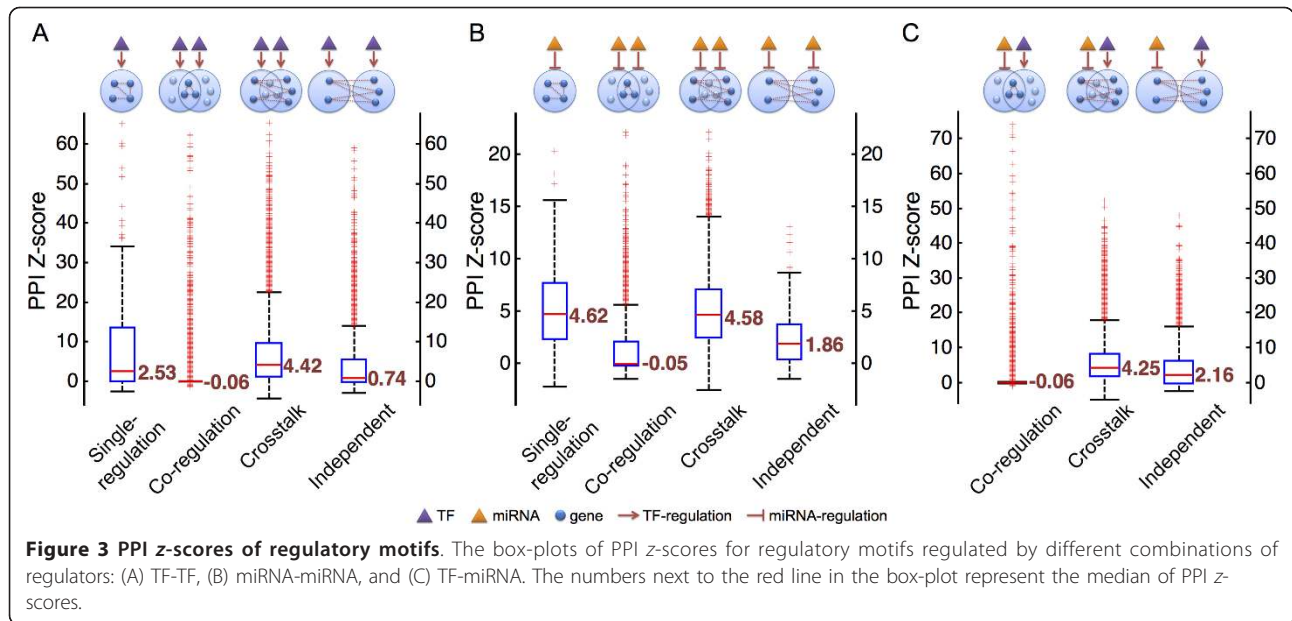
Figure 2 Properties of gene regulatory networks. The distributions of the number of target genes and synergistic partners for miRNAs and TFs.

co-regulation motifs regulated by TF-miRNA combinations were affected by the sample size as well (Figure S1 in Additional file).

Considering the reported TF-TF and TF-miRNA interactions [20], we divided each proposed motif into two subcategories, with or without known interactions of the regulator pairs. Motifs regulated by interacting regulator pairs displayed higher PPI z-scores than those

without known interactions (Figure 4). Notably, the crosstalk motifs showed the highest z-scores. This observation further confirmed our suggestions that regulatory motifs with synergistic relationships tend to regulate genes with PPIs, especially for crosstalk motifs.

For the bottom-up analysis, we investigated the PPI enrichment of genes sets in four types of regulatory motifs. The significance score and corresponding

Figure 3 PPI *z*-scores of regulatory motifs. The box-plots of PPI *z*-scores for regulatory motifs regulated by different combinations of regulators: (A) TF-TF, (B) miRNA-miRNA, and (C) TF-miRNA. The numbers next to the red line in the box-plot represent the median of PPI *z*-scores.

coverage of each type of regulatory motifs are presented in Table 1. In all combinations of regulator pairs, target genes of the single-regulation and co-regulation motifs showed significant PPI enrichment, but those of the independent motifs showed significantly insufficient PPI contents. This suggested that the targets of the single-regulation and co-regulation motifs preferred to form PPIs, but those of the independent motifs did not. With respect to the crosstalk motifs, targets regulated by TF-TF, miRNA-miRNA, and TF-miRNA pairs showed insignificant PPI enrichment. Here, we noticed that the coverage of the crosstalk motifs was much higher than
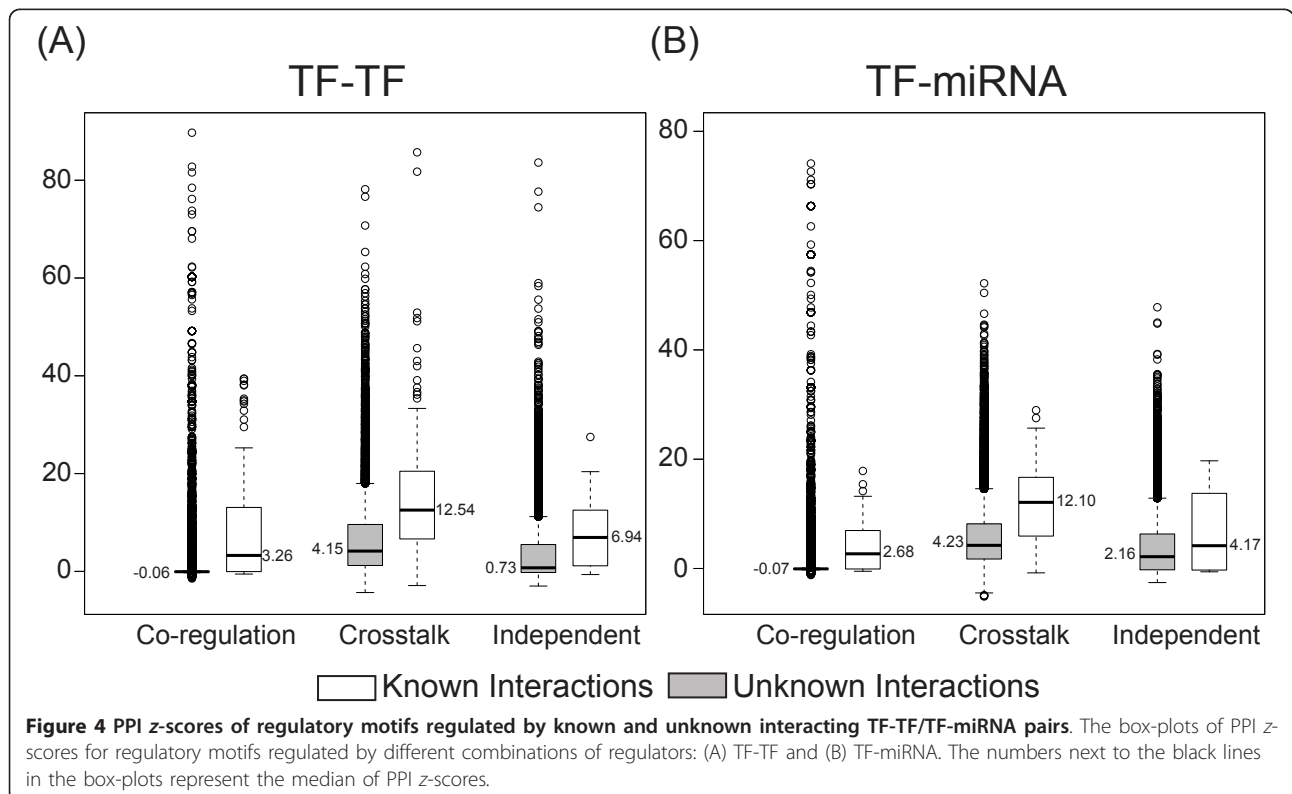

Figure 4 PPI *z*-scores of regulatory motifs regulated by known and unknown interacting TF-TF/TF-miRNA pairs. The box-plots of PPI *z*-scores for regulatory motifs regulated by different combinations of regulators: (A) TF-TF and (B) TF-miRNA. The numbers next to the black lines in the box-plots represent the median of PPI *z*-scores.

Table 1 PPI *z*-score and the coverage of gene sets involved in regulatory motifs

	Single-regulation	Co-regulation	Crosstalk	Independent
	TF-TF			
z-score	62.17	62.10	-0.80	-27.32
coverage	16.19%	4.31%	96.65%	83.81%
z-score*	56.43	51.37	11.37	-15.14
coverage*	6.72%	0.92%	70.60%	93.28%
	miRNA-miRNA			
z-score	35.37	38.15	1.52	-20.79
coverage	25.68%	9.14%	96.60%	74.32%
z-score*	34.74	32.16	14.37	-12.08
coverage*	10.79%	2.15%	69.08%	89.21%
	TF-miRNA			
z-score	-	62.44	-0.09	-55.91
coverage	-	6.65%	97.42%	73.52%
z-score*	-	47.94	15.71	-33.91
coverage*	-	1.18%	70.49%	88.31%

*The asterisk represents the filtered values. The upper limit of target gene size for filtration of TF-TF, miRNA-miRNA, and TF-miRNA motifs is 781, 789, and 789, respectively.
*The coverage represents the ratio of gene pairs in regulatory motifs to all possible gene pairs in HPRD PIN.

Table 2 *Z*-scores of each network properties for regulatory motifs

Motif	Single-regulation		Co-regulation			Crosstalk		
Regulator(s)	TF	miR	TF	miR	TF-miR	TF	miR	TF-miR
Degree	1.52	1.29	1.95	0.72	0.83	2.35	1.80	1.59
Closeness	1.55	1.86	2.02	1.22	1.26	2.45	2.73	2.12
Density	2.31	3.88	1.60	0.96	1.28	13.74	4.33	5.27
Clique level	1.51	1.46	1.53	1.19	1.00	2.14	2.02	1.63
Path length	-15.58	-24.10	-10.25	-6.66	-7.52	-21.27	-38.25	-50.47

*For single-regulation motifs, there are two types of regulators, TF and miRNA (miR). For co-regulation and crosstalk motifs, there are three types of regulators, TF-TF, miRNA-miRNA, and TF-miRNA. Here, we used TF and miR to represent TF-TF and miRNA-miRNA respectively. Path length represents the characteristic path length of the gene set.

the other three types of motifs, at almost 100%. This high coverage means that the tested sample is nearly identical to the whole population; therefore, the enrichment could be insignificant owing to the loose definition of crosstalk motifs. To test this hypothesis, we removed those regulators whose target sizes were excessively larger than other regulators; in other words, we removed the top outliers from the target-size distribution of regulators. Indeed, after this procedure, the PPI enrichment of crosstalk motifs emerged, and the results of other types of motifs remained the same (Table 1 Figure S2 in Additional file). In summary, genes under the control of regulatory motifs tend to form PPIs, except for those genes regulated by independent motifs.

According to the results of the top-down and the bottom-up analyses, we came to three conclusions: 1) the single-regulation motifs tend to regulate genes with PPIs. 2) Regulatory motifs with synergistic relationships (i.e. co-regulation and crosstalk) favor gene regulation with PPIs, especially for crosstalk motifs. 3) Gene pairs regulated by independent regulators (i.e. without synergistic relationships), in contrast, show no preference to form PPIs.

Regulatory motifs tend to regulate pivotal proteins in PIN

Genes encoding proteins with meaningful network properties in PINs have been proposed to play very important roles in living cells [39-44]. PPI enrichment analysis suggested that the single-regulation, co-regulation, and crosstalk motifs are highly correlated with PPIs. Herein, we further investigated the network properties of target genes (Methods in Additional file 1) involved in these three types of motifs. The *z*-scores of each of the network properties for regulatory motifs are summarized in Table 2 (more details in Figure S3 in Additional file). Network properties can be classified into two categories: 1) for individual genes–degree and closeness centrality; 2) for gene sets–density, clique level, and path length.

With respect to individual genes, most regulatory motifs tend to regulate those genes with higher degree and closeness centrality (*z*-score > 1). Degree represents the connectivity of proteins in a PIN, and closeness centrality represents how close proteins are to the center of a PIN. These results suggested that the regulatory motifs tend to regulate hub and central proteins. On the other hand, most regulatory motifs tend to regulate those gene sets with higher density (*z*-score > 1), larger clique levels (*z*-score > 1), and significantly shorter path lengths (*z*-score < -2). Density provides a quantitative measure of how tested gene sets group together to form a community in a PIN, clique level represents the level of maximal clique in which a tested gene can join, and path length describes how close tested proteins are to each other in a PIN. Briefly, these three network properties were usually used to evaluate the modularity of tested proteins. Hence, the results presented here imply that the regulatory motifs tend to control those proteins that form biological communities. Notably, the crosstalk motifs showed more significant *z*-scores than other types of regulatory motifs, suggesting they play more roles that are important in PINs.

Biological processes of the crosstalk motifs

After investigating the PPI enrichment and network properties of the screened regulatory motifs, we noticed

that the crosstalk motifs played a pivotal role in human PINs, and hence further studied their biological processes. First, we analyzed the functional similarity between two regulators. The results are shown in Figure 5A. Functional similarities between regulator pairs of crosstalk motifs, ranked in descending order, are as follows: miRNA-miRNA, TF-miRNA, and TF-TF (avg., 0.62, 0.42, and 0.37, respectively). Lower functional similarity in TF-TF and TF-miRNA pairs might reflect the dominant positions of TFs in global regulatory system (i.e. at the transcriptional level) [45]. Contrarily, higher functional similarity in miRNA-miRNA pairs might be due to the downstream positions of miRNAs in the global regulatory system (i.e. at the post-transcriptional level) [45]. In addition, we observed that a notable proportion of crosstalk motifs are with zero functional similarity (TF-TF: 29%; TF-miRNA: 22%;

miRNA-miRNA: 14%, Figure S4 in additional file), i.e. no common enriched functions between two regulators. To investigate this observation further, we compared the PPI z-scores and the averaging network property z-scores of crosstalk motifs with zero versus non-zero functional similarity. For all regulator combinations, the crosstalk motifs with non-zero functional similarity showed significantly higher PPI z-scores and network property z-scores than those with zero functional similarity (Figure 5B-G and Table 3). Therefore, the zero functional similarity might be due to the lack of PPIs between regulated private targets involved in the crosstalk motifs. This result suggests that the functional synergistic regulations of the crosstalk motifs could be based on the PPIs between regulated private target genes, highlighting the functional features of the crosstalk motifs.
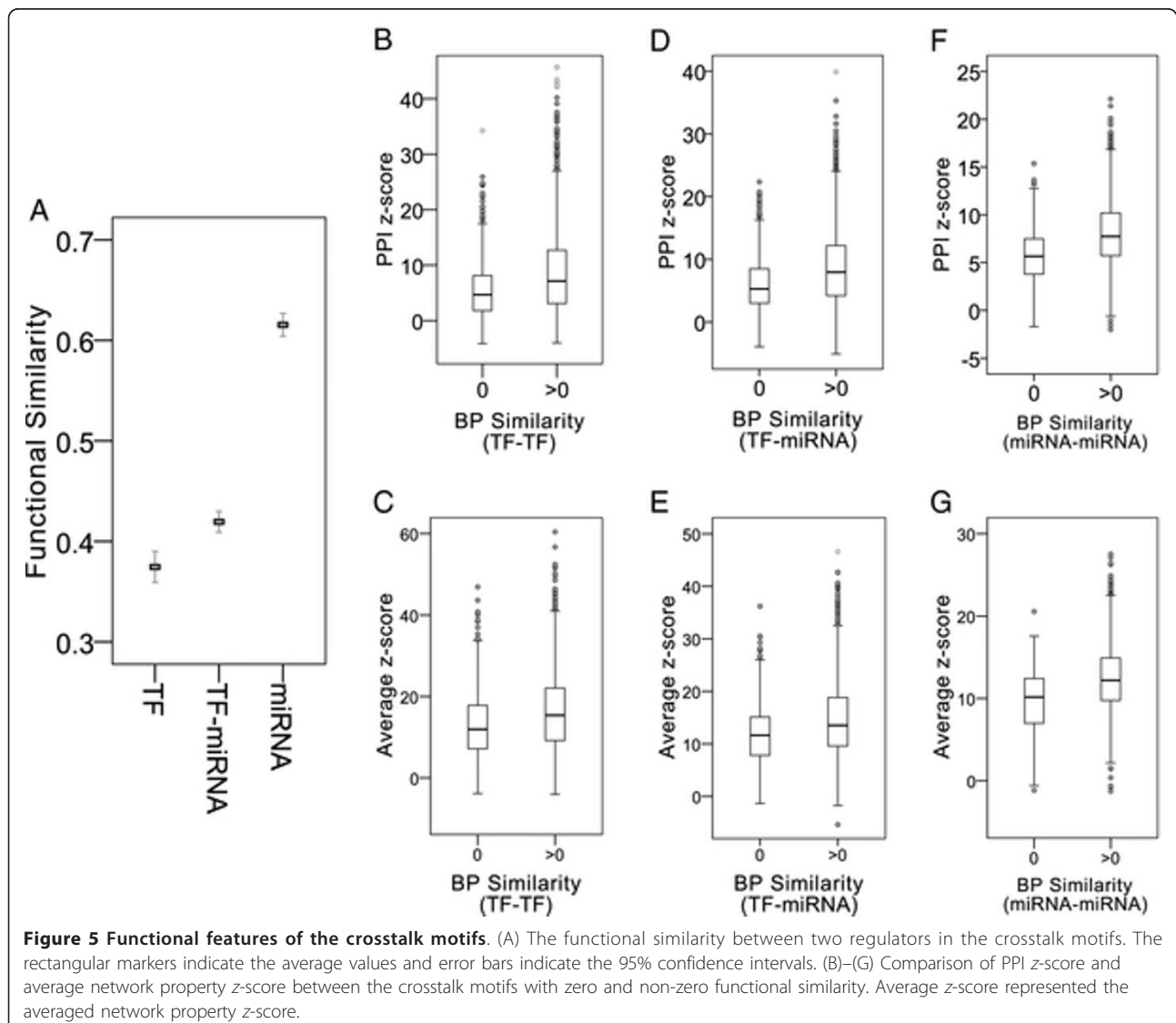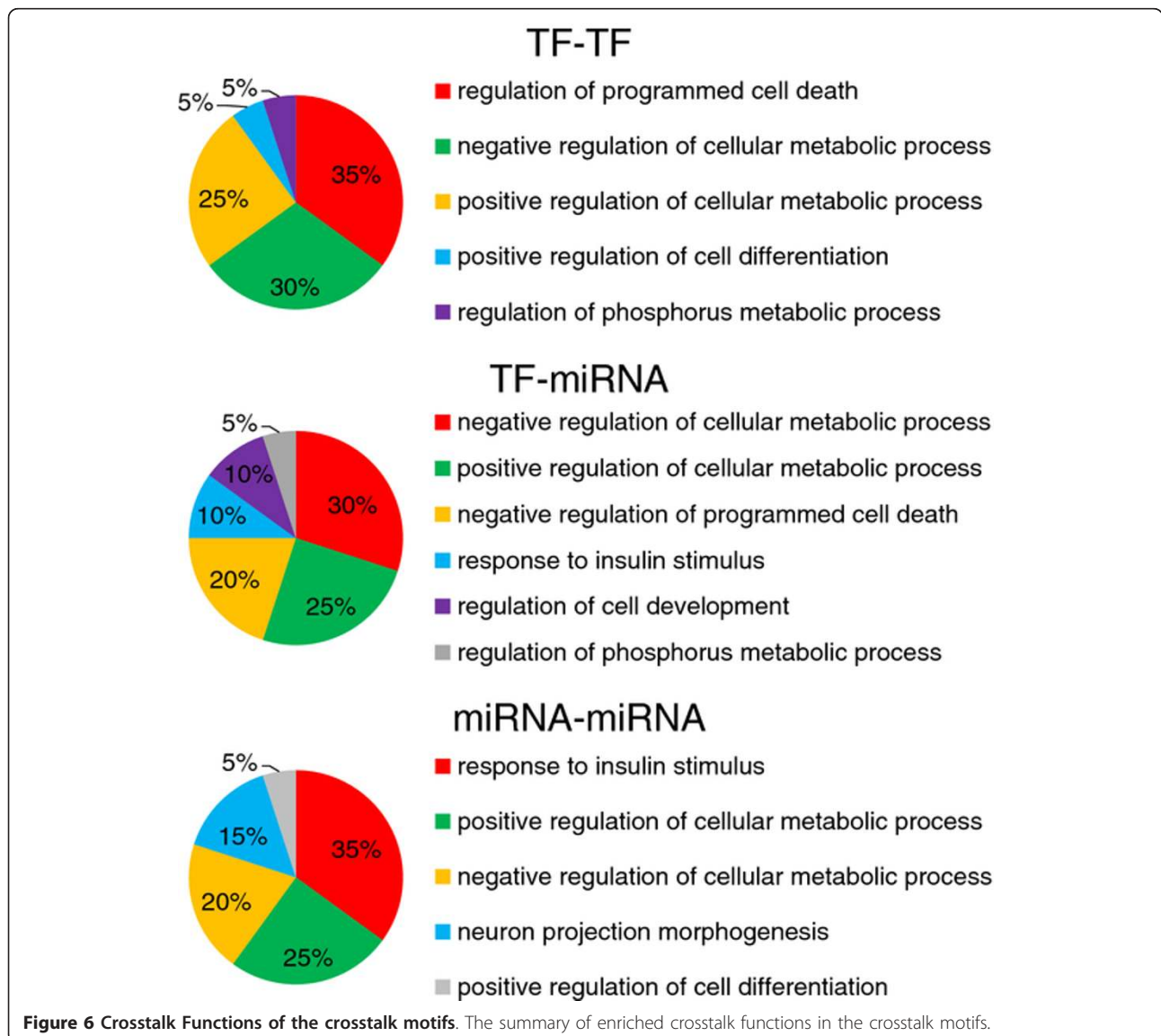

Figure 5 Functional features of the crosstalk motifs. (A) The functional similarity between two regulators in the crosstalk motifs. The rectangular markers indicate the average values and error bars indicate the 95% confidence intervals. (B)–(G) Comparison of PPI z-score and average network property z-score between the crosstalk motifs with zero and non-zero functional similarity. Average z-score represented the averaged network property z-score.

Table 3 PPI and network property *z*-score of crosstalk motifs with zero and non-zero functional similarity

	TF		TF-miRNA		miRNA	
z-score	PPI	Network	PPI	Network	PPI	Network
0	5.71	12.96	6.15	11.81	5.68	9.63
> 0	8.89	16.43	8.68	14.59	8.03	12.38
p-value	2.07E-15	1.89E-11	8.29E-29	2.31E-13	3.87E-42	2.02E-24

*P-values were derived from Wilcoxon rank-sum test. TF and miRNA represented TF-TF and miRNA-miRNA combinations respectively. Network represented the average *z*-score of the network properties. *Z*-scores in this table were averaged.

We then studied the underlying biological processes between private targets of crosstalk motifs (Methods in Additional file 1). For each combination of regulator pair, we selected top 20 biological processes ranked by proportions of involved motifs, respectively (Additional file 1 Figure S5 - S7). These biological processes covered nearly all the crosstalk motifs (TF-TF: 98.02%; TF-miRNA: 100%; miRNA-miRNA: 99.88%). Figure 6 shows a summary of these processes. The majority of selected processes for all three types of regulator pairs are associated with positive/negative regulation of cellular metabolic process. Notably, TF-TF crosstalk motifs also favor the processes associated with regulation of programmed cell death (apoptosis); miRNA-miRNA ones favor those with response to insulin stimulus; and TF-miRNA with both. These results not only displayed the functional homogeneity between regulators of crosstalk motifs, but also demonstrated the difference between TF and miRNA at regulatory level.

Figure 6 Crosstalk Functions of the crosstalk motifs. The summary of enriched crosstalk functions in the crosstalk motifs.

Figure 7 The TP53-miR-200bc/429/548a crosstalk motif. The visualization of the TP53-miR-200bc/548a crosstalk motif. There are 1,918 PPIs between target genes in this motif. 25.02% and 25.50% PPIs formed intra-connections within TP53 and mi-200bc/429/548a regulated-private target genes, respectively; 39.05% PPIs formed inter-connections between two private target sets.

We further inspected the potential biological processes of the TP53-miR-200bc/429/548a crosstalk motif (Figure 7) as it possessed the highest PPI *z*-score. In this crosstalk motif, there are 1,918 PPIs between target genes: around 25.02% and 25.50% PPIs formed intra-connections within TP53 and miR-200bc/429/548a regulated-private target genes, respectively, and around 39.05% PPIs formed inter-connections between two private target sets. The enriched interconnected PPIs within these motifs might imply massive crosstalk between regulators in their downstream regulatory pathways. PPIs between private targets of TP53-miR-200bc/429/548a were enriched in positive/negative regulation of cell death, response to insulin stimulus, epidermal growth factor receptor signaling pathway, toll-like receptor signaling pathway, positive regulation of cell differentiation/proliferation, regulation of protein kinase activity, protein phosphorylation, and regulation of cell

migration. TP53 is a well-studied cancer-related gene which encodes the tumor-suppressor protein p53 [46-48], and miR-200bc/429/548a has been reported to be significantly down-regulated in and related to several cancers [49-53]. For example, Shimon Y. *et al.* reported that miR-200bc/429/548a suppressed the ability of tumor formation driven by human breast cancer stem cell in vivo [50]; and Hu X. *et al.* reported that the over-expression of miR-200bc/429/548a could inhibit the cell migration of ovarian cancer cell and thus suggested that this miRNA should be strongly associated with cancer recurrence and overall survival [51]. Therefore, the TP53-miR-200bc/429/548a crosstalk motif might be a potential cancer-related regulatory motif.

Discussion and conclusions
In this study, we incorporated miRNAs into a traditional GRN to investigate the correlations between PPIs and

regulatory motifs formed by miRNAs, TFs, and target proteins/genes. The regulatory motifs were classified into four types: single-regulation, co-regulation, crosstalk, and independent. Traditionally, random sampling methods are usually applied to evaluate the significance of PPI numbers among a group of proteins, but this is very time-consuming. In addition, random sampling is not suitable for analyzing complicated regulatory networks, because the whole process should be redesigned for different motif members. In order to improve the efficiency of the evaluation process without loss of general applicability, we calculated the significance of PPI enrichment for different motifs based on the Bernoulli distribution; in other words, we regarded PPI gain and lost as a Bernoulli process. This allowed the whole evaluation process to be kept under constant time ($O(1)$).

Among the four types of motifs, the strong correlation between single-regulation and PINs has been well-discussed [21,22], and a correlation with the co-regulation type has also been reported [37,38]. Single-regulation motifs analyzed here showed consistent conclusions with previous studies. Our investigation into co-regulation motifs has further provided complementary analysis and given insights that have not been addressed in any previous studies. More importantly, we proposed that the third type of motif – the crosstalk motif – could be another prominent pattern in GRNs. Crosstalk motifs were defined as the private target gene sets of two corresponding regulators, TFs and/or miRNAs, which shared at least two targets. In human PINs, crosstalk motifs were significantly enriched in PPI contents and network properties. To summarize the analysis of network properties, crosstalk motifs displayed several features: 1) high degree, 2) high closeness, 3) high density, 4) high clique level, and 5) short characteristic path length. In PINs, proteins with a high degree are usually called "hub proteins", those with high closeness centrality are usually called "central proteins", and those with high density, short characteristic path length, and high clique level are usually called "modular proteins". Therefore, the regulators which participate in crosstalk motifs tend to regulate hub proteins, which are usually more essential than non-hub proteins [39-41], and modular proteins, which usually form important protein complexes or modules in human PINs [42-44]. Additionally, we investigated the enriched functions of the crosstalk motifs. For all three types of regulator pairs, the majority of enriched crosstalk functions are associated with positive/negative regulation of cellular metabolic processes. Notably, miRNA-miRNA crosstalk motifs are not only associated with regulation-related functions, but also response to insulin stimulus. This is consistent with previous findings that miRNAs preferentially regulate downstream components, such as TFs, in signaling

networks [19]. Moreover, we demonstrated the functional features within the crosstalk motifs with the highest PPI *z*-score and proposed a potential cancer-related motif, TP53-miR-200bc/429/548a. Consequently, this crosstalk motif might play an important role in living cells through regulating those essential or pivot proteins in human PINs.

Since our analysis relies on limited data sources from online databases to construct human PINs and GRNs, we carried out further examinations to test the robustness of our conclusions. With respect to miRNA regulation, all current online databases which provide predicted human miRNA targets still have room for improvement both in approach and performance [32-35]. Accordingly, we repeated our analysis with another database, miRBase [36], and were able to reach a consistent conclusion (Figure S8-S13, Table S1 and S2 in Additional file 1). Considering the incomplete and noisy human PPI data, we performed the same analysis with combined PPI data from HPRD and BioGRID [31] databases and also obtained consistent conclusions (Figure S14-S22, Table S3-S5 in Additional file 1). Therefore, these re-analyses provide further evidence to support the robustness of our conclusions. With ongoing efforts to improve the completeness of PPI data and GRNs, we will be able to further investigate and confirm the correlations between PPIs and regulatory motifs in the future.

In summary, we proposed a computational approach to investigate the significance of regulatory motifs formed by TFs/miRNAs and their corresponding targets in human PINs. With this approach, we screened four types of regulatory motifs, single-regulation, co-regulation, crosstalk, and independent, from human GRNs and investigated their correlations with PPIs. Among the four types of motifs, the crosstalk motif emerged as a potentially significant motif with important roles in PINs, which has not been previously reported. We suggested that this motif might play an important role in living cells because of its strong correlations with PPIs and significant network properties in human PINs.

Additional material

Additional file 1: Supplementary methods, figures, and tables. This file contains supplementary methods and results, and the repeat analysis for confirming the robustness of our results with different datasets, miRNA-target prediction and PPI data.

Author details
[1]Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei 106, Taiwan. [2]Genome and Systems Biology Degree Program, National Taiwan University, Taipei 106, Taiwan. [3]Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan. [4]Department of Life Science and Institute of Molecular and Cellular Biology, National Taiwan University, Taipei 106, Taiwan. [5]Institute of Biomedical Informatics, Center for Systems and Synthetic Biology, National Yang-Ming University, Taipei 112, Taiwan.

Authors' contributions
CCL, YJC, CYC, and HCH carried out the analysis, and drafted the manuscript. YJO, HFJ, and HCH conceived and directed the project, participated in the design and coordination of the study, and edited the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

References
1. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116**:281-297.
2. Flynt AS, Lai EC: **Biological principles of microRNA-mediated regulation: shared themes amid diversity.** *Nat Rev Genet* 2008, **9**:831-842.
3. Kim VN, Nam JW: **Genomics of microRNA.** *Trends Genet* 2006, **22**:165-173.
4. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I, *et al*: **Transcriptional regulatory networks in Saccharomyces cerevisiae.** *Science* 2002, **298**:799-804.
5. Shen-Orr SS, Milo R, Mangan S, Alon U: **Network motifs in the transcriptional regulation network of Escherichia coli.** *Nat Genet* 2002, **31**:64-68.
6. Yu H, Gerstein M: **Genomic analysis of the hierarchical structure of regulatory networks.** *Proc Natl Acad Sci USA* 2006, **103**:14724-14731.
7. Latchman DS: **Transcription factors: an overview.** *Int J Biochem Cell Biol* 1997, **29**:1305-1312.
8. Mitchell PJ, Tjian R: **Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins.** *Science* 1989, **245**:371-378.
9. Ptashne M, Gann A: **Transcriptional activation by recruitment.** *Nature* 1997, **386**:569-577.
10. Mangan S, Alon U: **Structure and function of the feed-forward loop network motif.** *Proc Natl Acad Sci USA* 2003, **100**:11980-11985.
11. Milo R, Itzkovitz S, Kashtan N, Levitt R, Shen-Orr S, Ayzenshtat I, Sheffer M, Alon U: **Superfamilies of evolved and designed networks.** *Science* 2004, **303**:1538-1542.
12. Doench JG, Sharp PA: **Specificity of microRNA target selection in translational repression.** *Genes Dev* 2004, **18**:504-511.
13. Guo H, Ingolia NT, Weissman JS, Bartel DP: **Mammalian microRNAs predominantly act to decrease target mRNA levels.** *Nature* 2010, **466**:835-840.
14. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U: **Network motifs: simple building blocks of complex networks.** *Science* 2002, **298**:824-827.
15. Rosenfeld N, Elowitz MB, Alon U: **Negative autoregulation speeds the response times of transcription networks.** *J Mol Biol* 2002, **323**:785-793.
16. Becskei A, Serrano L: **Engineering stability in gene networks by autoregulation.** *Nature* 2000, **405**:590-593.
17. Chen CY, Chen ST, Fuh CS, Juan HF, Huang HC: **Coregulation of transcription factors and microRNAs in human transcriptional regulatory network.** *BMC Bioinf* 2011, **12(1)**:S41.
18. Cui Q, Yu Z, Pan Y, Purisima EO, Wang E: **MicroRNAs preferentially target the genes with high transcriptional regulation complexity.** *Biochem Biophys Res Commun* 2007, **352**:733-738.
19. Cui Q, Yu Z, Purisima EO, Wang E: **Principles of microRNA regulation of a human cellular signaling network.** *Mol Syst Biol* 2006, **2**:46.
20. Qiu C, Wang J, Yao P, Wang E, Cui Q: **microRNA evolution in a human transcription factor and microRNA regulatory network.** *BMC Syst Biol* 2010, **4**:90.
21. Liang H, Li WH: **MicroRNA regulation of human protein protein interaction network.** *RNA* 2007, **13**:1402-1408.
22. Hsu CW, Juan HF, Huang HC: **Characterization of microRNA-regulated protein-protein interaction network.** *Proteomics* 2008, **8**:1975-1979.
23. Jiang C, Xuan Z, Zhao F, Zhang MQ: **TRED: a transcriptional regulatory element database, new entries and other development.** *Nucleic Acids Res* 2007, **35**:D137-D140.
24. Lewis BP, Burge CB, Bartel DP: **Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.** *Cell* 2005, **120**:15-20.
25. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP: **MicroRNA targeting specificity in mammals: determinants beyond seed pairing.** *Mol Cell* 2007, **27**:91-105.
26. Friedman RC, Farh KK, Burge CB, Bartel DP: **Most mammalian mRNAs are conserved targets of microRNAs.** *Genome Res* 2009, **19**:92-105.
27. Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP: **Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs.** *Nat Struct Mol Biol* 2011, **18**:1139-1146.
28. Cheng C, Yan KK, Hwang W, Qian J, Bhardwaj N, Rozowsky J, Lu ZJ, Niu W, Alves P, Kato M, *et al*: **Construction and analysis of an integrated regulatory network derived from high-throughput sequencing data.** *PLoS Comput Biol* 2011, **7**:e1002190.
29. Alexiou P, Maragkakis M, Papadopoulos GL, Reczko M, Hatzigeorgiou AG: **Lost in translation: an assessment and perspective for computational microRNA target identification.** *Bioinformatics* 2009, **25**:3049-3055.
30. Goel R, Muthusamy B, Pandey A, Prasad TS: **Human protein reference database and human proteinpedia as discovery resources for molecular biotechnology.** *Mol Biotechnol* 2011, **48**:87-95.
31. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M: **BioGRID: a general repository for interaction datasets.** *Nucleic Acids Res* 2006, **34**:D535-D539.
32. Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N: **Widespread changes in protein synthesis induced by microRNAs.** *Nature* 2008, **455**:58-63.
33. Min H, Yoon S: **Got target? Computational methods for microRNA target prediction and their extension.** *Exp Mol Med* 2010, **42**:233-244.
34. Bartel DP: **MicroRNAs: target recognition and regulatory functions.** *Cell* 2009, **136**:215-233.
35. Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP: **The impact of microRNAs on protein output.** *Nature* 2008, **455**:64-71.
36. Kozomara A, Griffiths-Jones S: **miRBase: integrating microRNA annotation and deep-sequencing data.** *Nucleic Acids Res* 2011, **39**:D152-D157.
37. Thuault S, Tan EJ, Peinado H, Cano A, Heldin CH, Moustakas A: **HMGA2 and Smads co-regulate SNAIL1 expression during induction of epithelial-to-mesenchymal transition.** *J Biol Chem* 2008, **283**:33437-33446.
38. Boross G, Orosz K, Farkas IJ: **Human microRNAs co-silence in well-separated groups and have different predicted essentialities.** *Bioinformatics* 2009, **25**:1063-1069.
39. Yook SH, Oltvai ZN, Barabasi AL: **Functional and topological characterization of protein interaction networks.** *Proteomics* 2004, **4**:928-942.
40. Reguly T, Breitkreutz A, Boucher L, Breitkreutz BJ, Hon GC, Myers CL, Parsons A, Friesen H, Oughtred R, Tong A, *et al*: **Comprehensive curation and analysis of global interaction networks in Saccharomyces cerevisiae.** *J Biol* 2006, **5**:11.
41. Jeong H, Mason SP, Barabasi AL, Oltvai ZN: **Lethality and centrality in protein networks.** *Nature* 2001, **411**:41-42.
42. Lin CC, Juan HF, Hsiang JT, Hwang YC, Mori H, Huang HC: **Essential core of protein-protein interaction network in Escherichia coli.** *J Proteome Res* 2009, **8**:1925-1931.
43. Rives AW, Galitski T: **Modular organization of cellular networks.** *Proc Natl Acad Sci USA* 2003, **100**:1128-1133.
44. Barabasi AL, Oltvai ZN: **Network biology: understanding the cell's functional organization.** *Nat Rev Genet* 2004, **5**:101-113.
45. Hobert O: **Gene regulation by transcription factors and microRNAs.** *Science* 2008, **319**:1785-1786.
46. Bourdon JC: **p53 and its isoforms in cancer.** *Br J Cancer* 2007, **97**:277-282.
47. Toledo F, Wahl GM: **Regulating the p53 pathway: in vitro hypotheses, in vivo veritas.** *Nat Rev Cancer* 2006, **6**:909-923.
48. Vogelstein B, Kinzler KW: **Tumour-suppressor genes. X-rays strike p53 again.** *Nature* 1994, **370**:174-175.

49. Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, Vadas MA, Khew-Goodall Y, Goodall GJ: **The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1.** *Nat Cell Biol* 2008, **10**:593-601.

50. Shimono Y, Zabala M, Cho RW, Lobo N, Dalerba P, Qian D, Diehn M, Liu H, Panula SP, Chiao E, *et al*: **Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells.** *Cell* 2009, **138**:592-603.

51. Hu X, Macdonald DM, Huettner PC, Feng Z, El Naqa IM, Schwarz JK, Mutch DG, Grigsby PW, Powell SN, Wang X: **A miR-200 microRNA cluster as prognostic marker in advanced ovarian cancer.** *Gynecol Oncol* 2009, **114**:457-464.

52. Korpal M, Lee ES, Hu G, Kang Y: **The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2.** *J Biol Chem* 2008, **283**:14910-14914.

53. Tryndyak VP, Ross SA, Beland FA, Pogribny IP: **Down-regulation of the microRNAs miR-34a, miR-127, and miR-200b in rat liver during hepatocarcinogenesis induced by a methyl-deficient diet.** *Mol Carcinog* 2009, **48**:479-487.