Supplementary files of 'Functional dissection of virus-human crosstalk mediated by miRNAs based on the VmiReg database'

Supplementary Table S1 - The statistics of the number of experimentally validated and predicted miRNA-target interactions.

Method name	#viral miRNA:mRNA interactions	#human miRNA:mRNA interactions
Experimentally validated	265	3120
PITA	495,150(495,134)	-
Miranda	368,948(368,819)	-
RNAhybrid	207,026(207,012)	-
TagetScan	343,181(343,069)	-

Note: The number in bracket is presented the number of only computationally predicted miRNA:mRNA interactions.

Supplementary Table S2 - Sensitivity for each program.

Method Name	#predicted interactions	#viral miRNA	#interactions per viral miRNA	#validated interaction	Sensitivity
PITA	588	34	17	16	7.92%
Miranda	2072	34	61	129	63.86%
RNAhybrid	276	29	10	14	6.93%
TagetScan	1039	34	31	112	55.45%

Note: Sensitivity=TP/(TP+FN) 'TP' is defined as the number of experimentally supported miRNA-target interactions that are predicted by a prediction method, and 'FN' is defined as the number of experimentally supported miRNA-target interactions that are not predicted by the prediction method.

Supplementary Figure S1. The screenshot showing the tools of the VmiReg database.
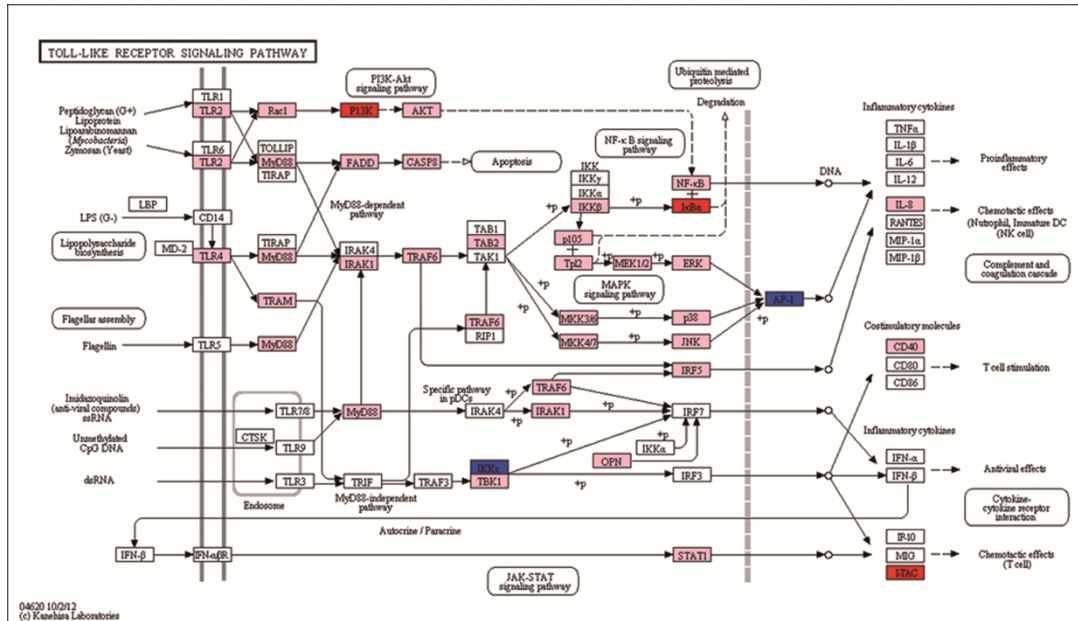
Supplementary Figure S2. Toll-like receptor signaling pathway enriched by targets of viral miRNA and human miRNA. The Toll-like receptor signaling pathway comes from KEGG. Genes targeted by human miRNAs are marked in pink, genes targeted by viral miRNAs are marked in red, and genes co-regulated by viral and human miRNAs are marked in blue.
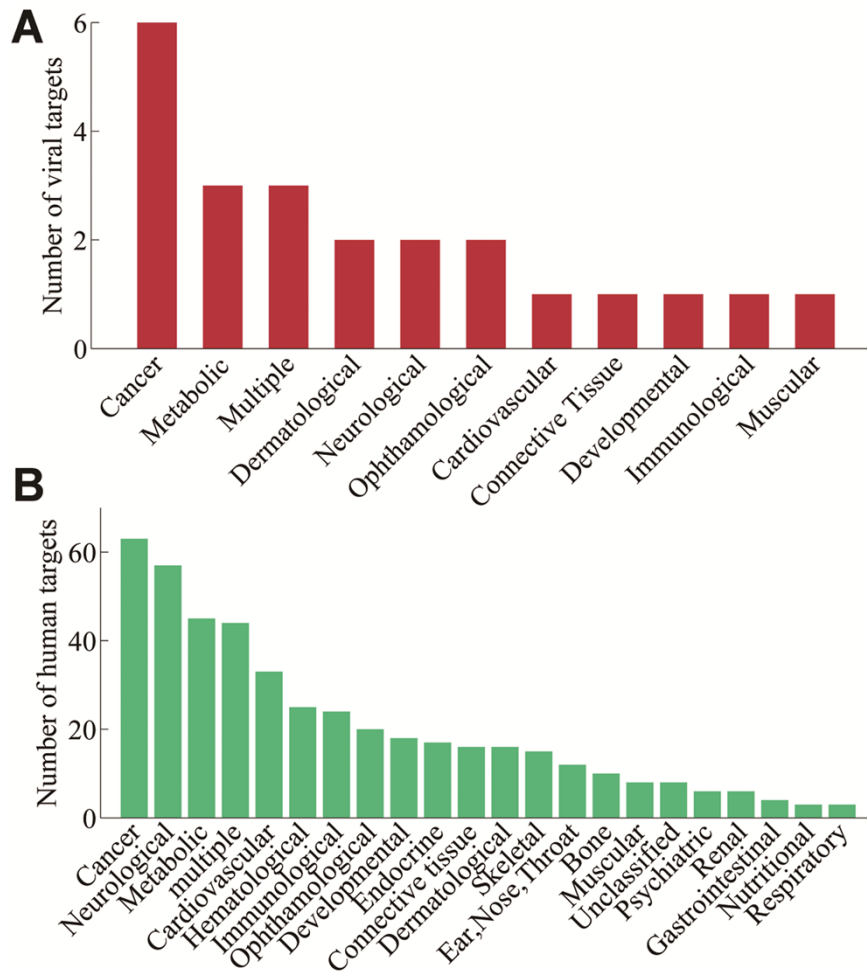
Supplementary Figure S3. Function of viral miRNAs and human miRNAs association with diseases. (A) The number of viral miRNA targets in 11 disease categories from OMIM. (B) The number of human miRNA targets in 22 disease categories from OMIM.
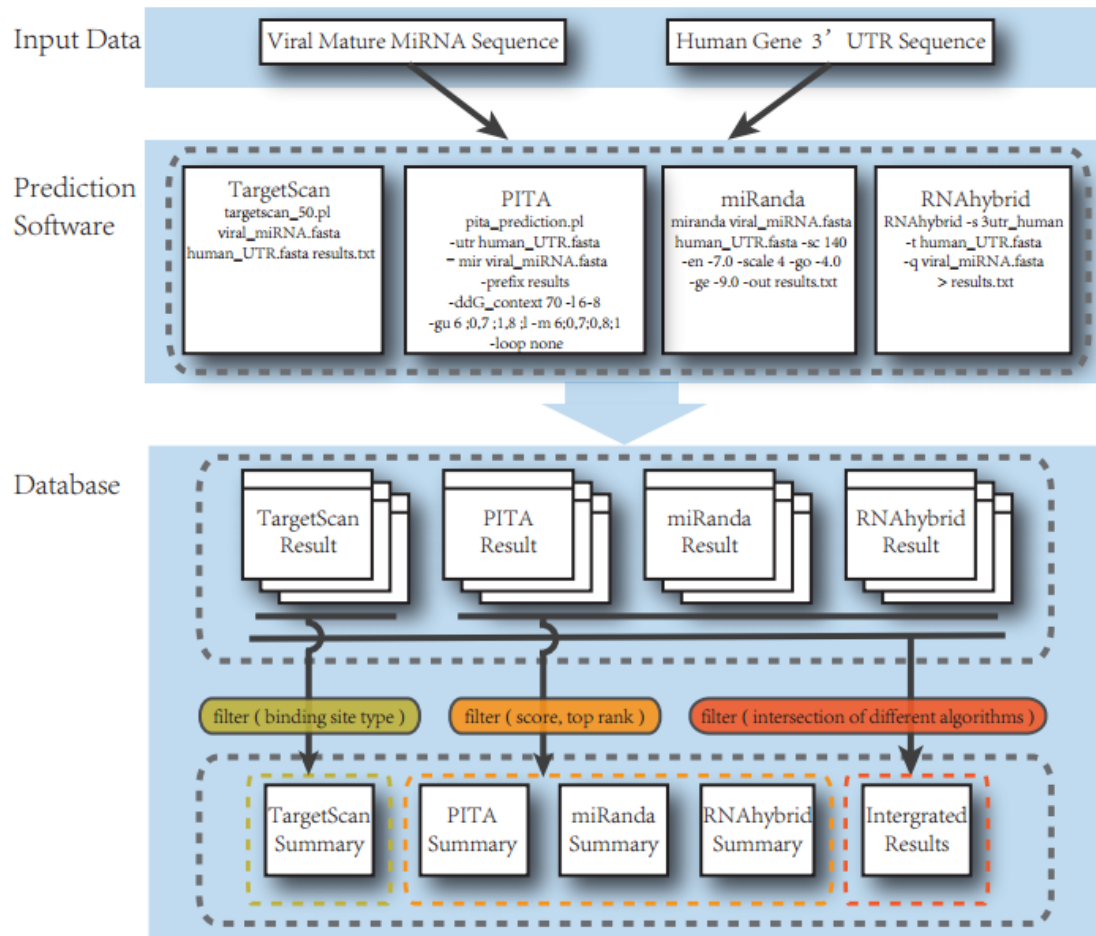
Supplementary Figure S4. Overall system architecture of the prediction part in VmiReg database. In the prediction part, four prediction programs, miRanda, TargetScan, RNAhybrid and PITA are used to predict the human targets of viral miRNAs. The predicted results are stored in VmiReg database. When the database is up and running the different filtering options can be chosen to get predicted results according to different needs. All the commands used to run the four prediction programs are shown in figure. "viral_miRNA.fasta" stands for the viral miRNA fasta file; "human_UTR.fasta" stands for the human 3'UTR sequences file; "results.txt" is the predicted result file.

Supplementary Text S1: Robustness analysis of the virus-human crosstalk principles mediated by miRNAs using updated data.

Data

In order to test the validation of the main results, here we updated the experimentally supported interactions by manually reviewing ~100 literatures published between April 2012 and October 2014. Totally, there were 236 regulations between 29 viral miRNAs and 181 human genes, and miRNAs of Adenovirus, KSHV, EBV, HCMV, Human immunodeficiency virus 1(HIV) and Herpes simplex virus 1(HSV1) were found to regulate human genes. These updated interactions were used to validate the main virus-human crosstalk patterns we found.

The crosstalk between viral and human miRNAs via their targets with similar functions

We first explored the crosstalk between viral and cellular miRNAs from the view of regulatory network. Viral and human miRNA-target regulatory network (abbreviated VHMiTN) is an objective representation of all validated regulations (Figure S7A). Two types of degree distributions were used to evaluate the global topological structure of the VHMiTN: out-degree distribution and in-degree distribution. As a result, out-degree distributions of viral and human miRNAs are similar, both revealing power-law trends, that is, most of miRNAs are lowly connected and only a few are relatively highly connected (Figure S7B). On the other hand, both the 'in-degree' distributions signify exponential distributions (Figure S7C). Therefore, like many large-scale networks, the VHMiTN displays scale-free characteristics, indicating that the VHMiTN is not random but is characterized by a core set of organizing principles in its structure that distinguish it from randomly linked networks. Based on the above observation, miRNAs encoded by viruses or human have the same regulatory patterns and their regulations are connected together to form the big network, the VHMiTN, which provide us a clue that co-regulated target genes might be one of the crosstalk ways between virus and human mediated by miRNAs. We further computed the extent of overlap between viral and human miRNA targets in the VHMiTN. 30% of transcripts targeted by viral miRNAs are also regulated by human miRNAs, which is six times greater than that of human miRNA targets in the whole genome (Figure S7D P<1.0e-32, computed by hypergeometric test). Therefore, target genes of viral miRNAs are also prone to be silenced by human miRNAs, implying that the means of the crosstalk between viral and human miRNAs might be via co-regulating targets.
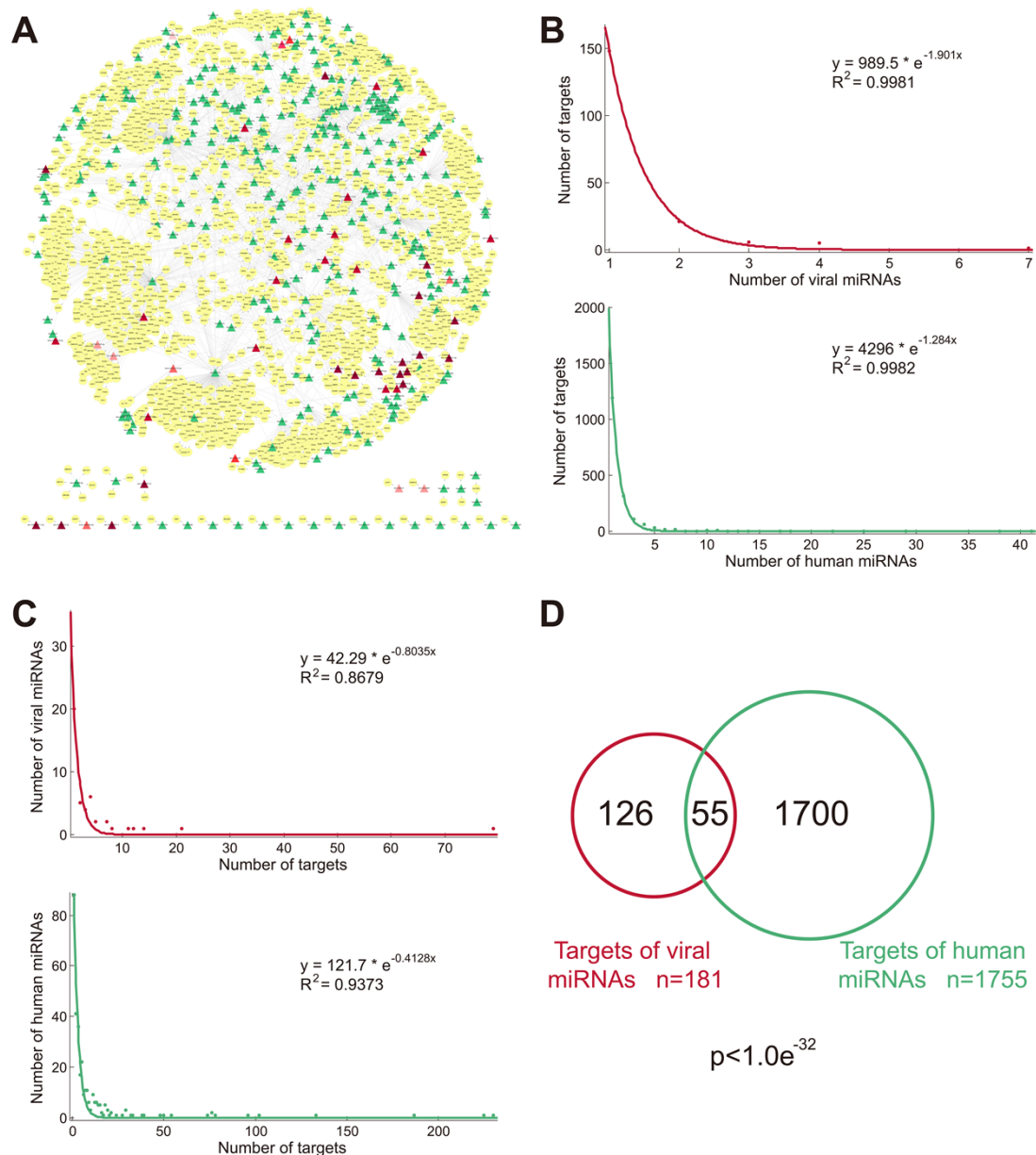
Supplementary Figure S7. The layout and topological features of the VHMiTN. (A) The VHMiTN generated by validated viral and human miRNA regulations. (B) Out-degree distribution of the VHMiTN. (C) In-degree distribution of the VHMiTN. (D) Overlap of targets between viral miRNAs and human miRNAs. Red represents viruses and green represents human.

Moreover, we performed enrichment analysis to discover biological themes significantly co-controlled by viral and human miRNAs and excavated functional crosstalk between them. Two kinds of common functional annotation data were analyzed, which were from GO and KEGG databases. Here, we only considered the ontology of biological process in the GO database. Figure S8B summarizes the top (12) GO themes significantly enriched by at least five targets of viral miRNAs. Most of these GO themes are associated with oncogenesis, including apoptosis, transcriptional regulation, signalling transduction and so on. On the other hand, there are also several biological pathways significantly regulated by virus-encoded miRNAs. We found that most of these pathways are related to signal transduction and oncogenesis (Figure S8B).

A

GO term	Description
GO:0007265	Ras protein signal transduction
GO:0045944	positive regulation of transcription from RNA polymerase II promoter
GO:0007165	signal transduction
GO:0000278	mitotic cell cycle
GO:0045892	negative regulation of transcription, DNA-dependent
GO:0019048	virus-host interaction
GO:0000122	negative regulation of transcription from RNA polymerase II promoter
GO:0048011	nerve growth factor receptor signaling pathway
GO:0045893	positive regulation of transcription, DNA-dependent
GO:0043066	negative regulation of apoptotic process
GO:0006915	apoptotic process
GO:0006917	induction of apoptosis

-Log(FDR)

B

map	Description
map:5162	Measles
map:4350	TGF-beta signaling pathway
map:5222	Small cell lung cancer
map:5160	Hepatitis C
map:5220	Chronic myeloid leukemia
map:5215	Prostate cancer
map:5200	Pathways in cancer
map:4722	Neurotrophin signaling pathway
map:4115	p53 signaling pathway
map:4110	Cell cycle

-Log(FDR)

Supplementary Figure S8. Viral and Human miRNAs have close crosstalk via genes with important functions. (A) Selected GO terms with significant enrichments of viral miRNA targets (red) and human miRNA targets (green). (B) Selected pathways with significant enrichments of viral miRNA targets (red) and human miRNA targets (green). A threshold of FDR ≤ 0.05 was used to judge the significantly enriched GO terms or pathways enriched by at least five targets of viral miRNA.

Next, we deciphered the functional crosstalk of viral miRNAs and human miRNAs via their cellular targets from two levels: association with diseases, and lethality. As a result, we found that 64% (117/181) of viral miRNA human targets are directly associated with diseases by analyzing disease data from the GAD database, significant higher than random conditions (P<1.0e-32, computed by hypergeometric test, Figure S9).As shown in Figure S9, cellular target genes of both viral and human miRNAs are mainly associated with cancers, metabolic diseases, immune diseases and cardiovascular diseases. Likewise, we got the same conclusion by using another set of disease data from the OMIM database (see Figure S9 and Figure S10). Lastly, we continued to focus on another kind of important genes, essential genes. Dysfunctions of essential genes would lead to death or severe disease phenotype. As a result, a preference for targeting essential genes was observed for both viral and human miRNAs (Figure S9 P<1.0e-32, computed by hypergeometric test).
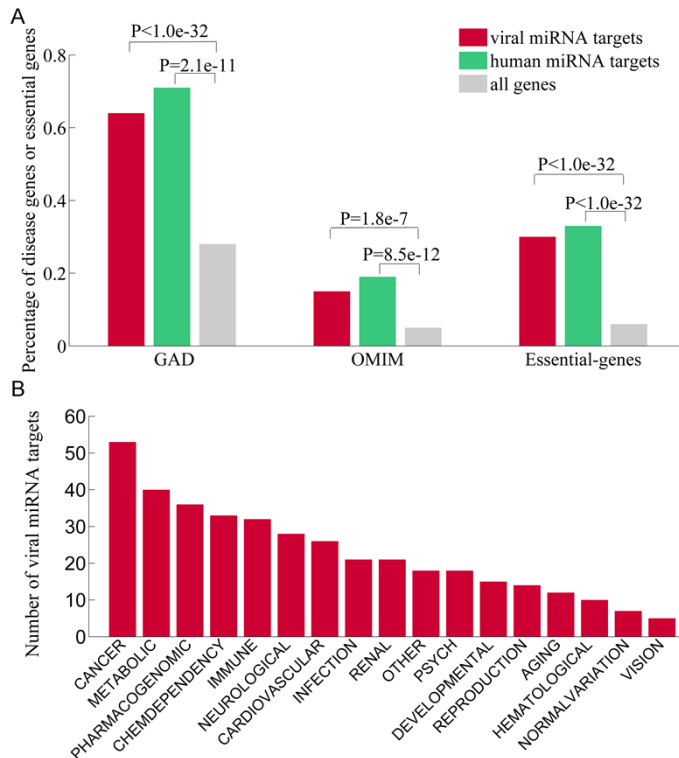
Supplementary Figure S9. Function of viral miRNAs and human miRNAs association with diseases and lethality. (A) Comparison of the percentage of disease genes or essential genes. Disease genes from GAD and OMIM are shown respectively. (B) The number of viral miRNA targets in 17 disease categories from GAD.
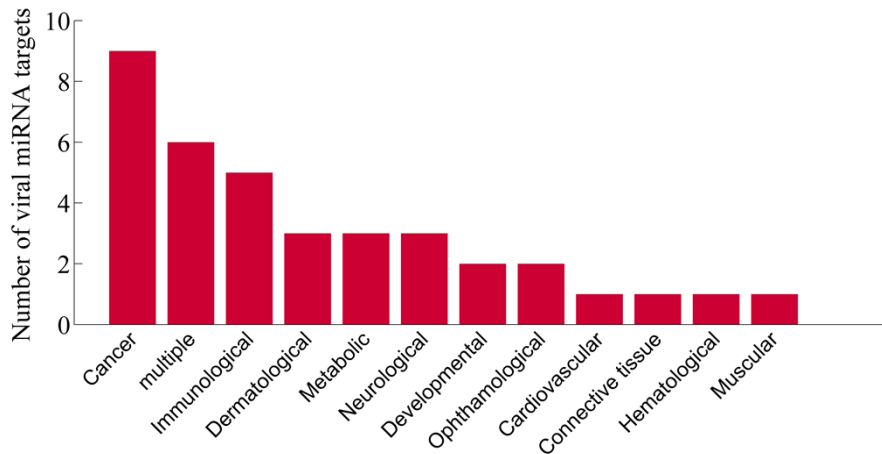
Supplementary Figure S10. Function of viral miRNAs association with diseases. The number of viral miRNA targets in 12 disease categories from OMIM.

The functional crosstalk between viral and human miRNAs in the context of the PPIN

First, we respectively calculated three topological centrality features of targets regulated by viral and human miRNAs in the PPIN, including degree, betweenness, and closeness. As a result, targets of both viral and human miRNAs have significantly higher degree, betweenness, and closeness than random conditions, suggesting that target genes tend to be located at the center of the PPIN (Table S7). Degree and betweenness of viral miRNA targets are 20 and 119,909, respectively, both of which are much larger than those of human miRNA targets. We furthermore observed an

enrichment of targets regulated by viral miRNAs in the top genes with high number of interactions (hubs), even when considering different gradients of degree (Supplementary Figure S11).

Supplementary Table S5 - Comparisons of topological features of viral miRNA targets and human miRNA targets.

	HPRD	Virus_miRNA_targets			Human_miRNA_targets		
	Mean	Mean	expectation	p-values	Mean	expectation	p-values
Degree	7.9453	20.2639	7.9544	0.0007	10.4715	7.9455	<10e-4
Betweenness $(*10^4)$	2.9137	11.9909	2.9206	<10e-4	4.4122	2.9147	<10e-4
Closeness	0.2310	0.2576	0.1062	<10e-4	0.2478	0.1063	<10e-4

p-value is calculated by the randomization test.
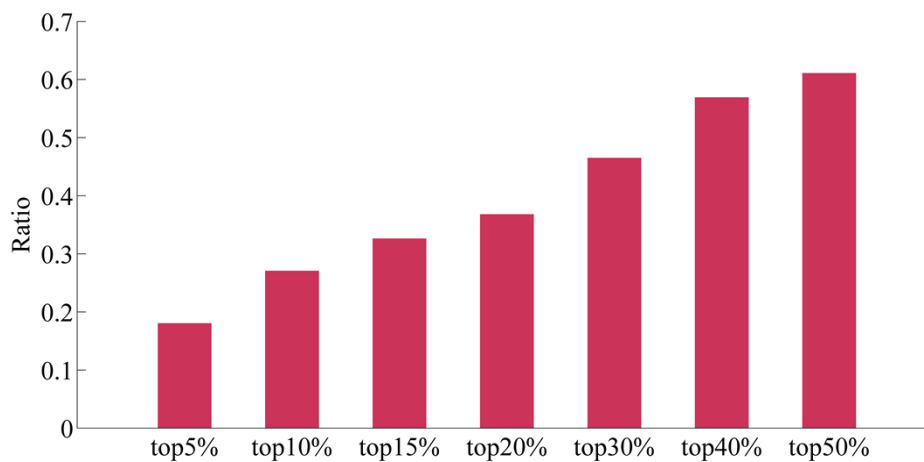
Supplementary Figure S11. The percentage of viral miRNA targets in the hubs. Genes are ranked by the degree in the PPIN and hubs are defined as the top ranked genes.

Next, we attempted to explore the efficiency of communication between viral miRNAs and human miRNAs based on the measurement of characteristic path length, which represents close proximity and consequently how quickly information can spread in a network. The characteristic path length between two kinds of target genes is 3.709 significantly lower than random conditions, indicating that target genes of viral miRNAs and those of human miRNAs are close proximity in the PPIN (Figure S12A, $P<10-4$). Indeed, by further calculating the number of direct interactions between these two types of targets, we found that there are up to 913 interactions between them, also remarkably higher than random gene sets (Figure S12B, $P<10-4$). In addition, although the genes co-regulated by viral and human miRNAs are limited, the one-step neighbours of genes targeted by viral miRNAs tend to be targets of human miRNAs, and the number reaches 467 (Figure S12C, $P<1.0e-4$).
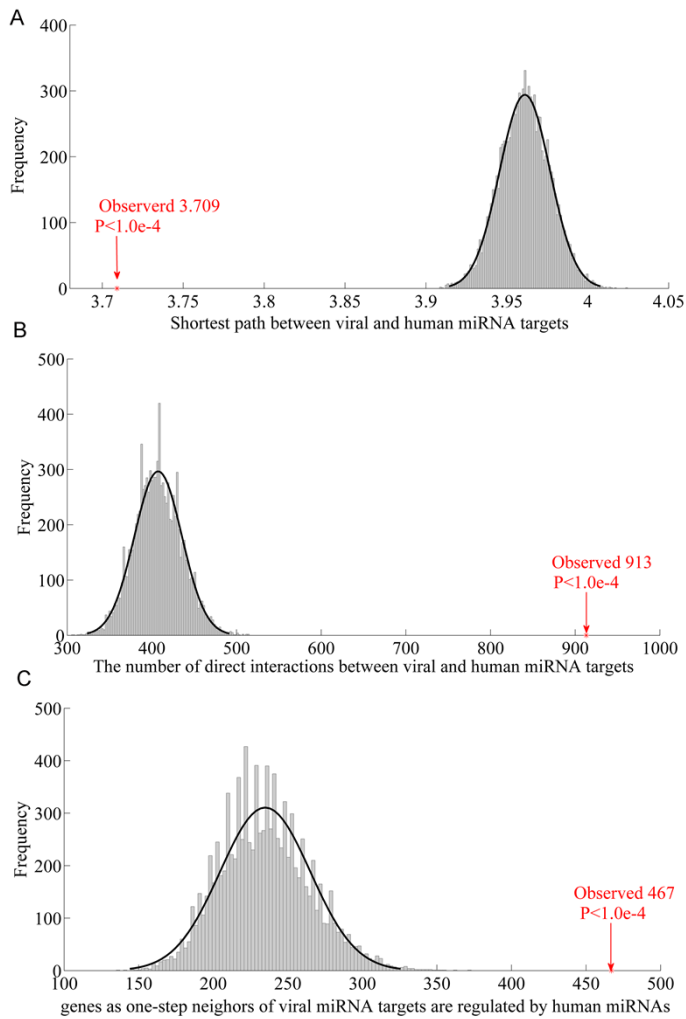
Supplementary Figure S12.Target genes of human and viral miRNAs are close proximity in PPIN. (A) The characteristic path length between viral miRNA targets and human miRNA targets is lower than random conditions. (B) The number of direct interactions between viral miRNA targets and human miRNA targets is higher than random conditions. (C) The number of genes as one-step neighbours of viral miRNA targets that are regulated by human miRNAs is higher than random conditions. The arrow represents the real value; the line is fitted using random distributions.

Next, we discovered regulatory modules to visually demonstrate the local crosstalk between viral and cellular miRNAs. Firstly, a subnetwork is constructed and consists of two kinds of genes, targets of viral miRNAs and their direct neighbours in the PPIN which were required to be regulated by at least one human miRNA. Interactions between these nodes were considered. Next, modules were detected by Markov Cluster algorithm (MCL algorithm) realized by its corresponding plugins in the Cytoscape software by using the default parameters[1]. After mapping miRNAs to these modules, we found that genes co-regulated by both viral and human miRNAs tend to be in the important locations of modules, whose deletion would disrupt the integrity of the corresponding modules (Figure 13). We proposed that after virus infection, some human miRNAs and viral miRNAs might have cumulative (synergistic) effects in disease states, and jointly promote occurrence of diseases.
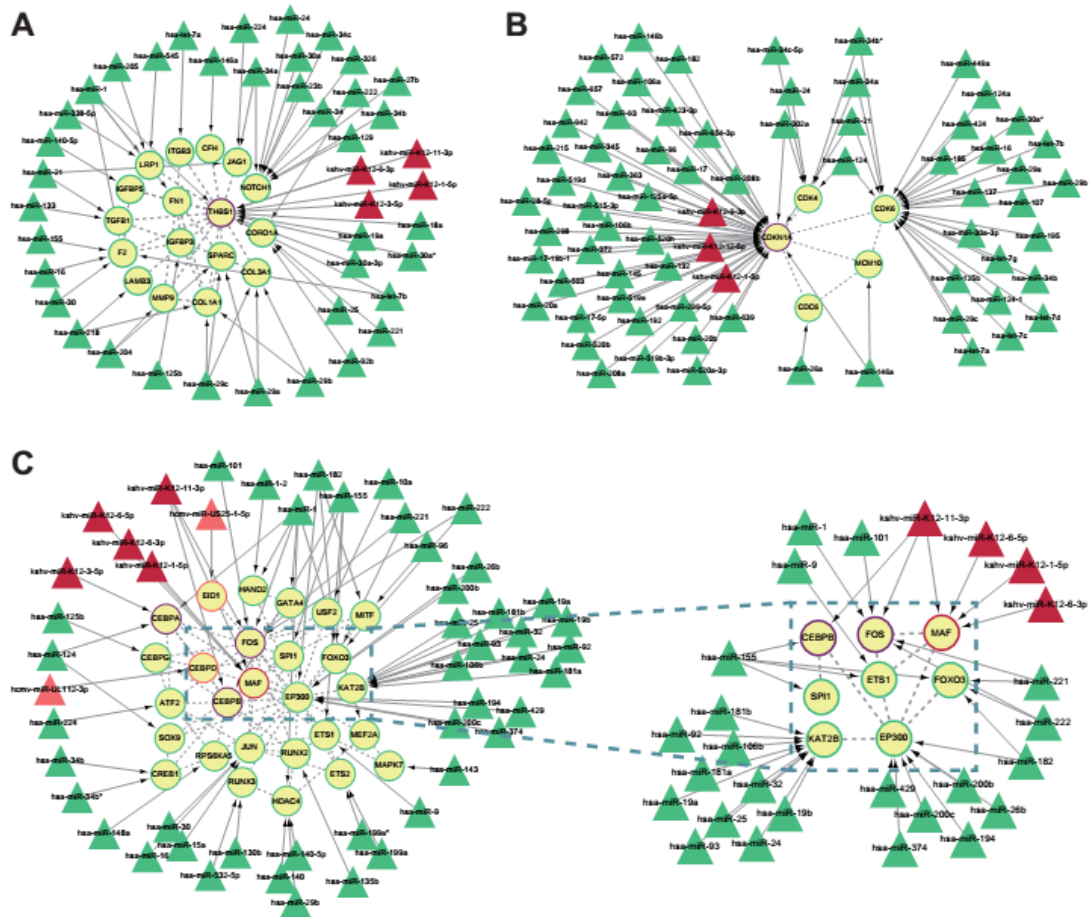
Supplementary Figure S13. Three composite modules composed of viral miRNAs, human miRNAs, targets and interaction between targets. A triangle node marks miRNA and a circle node marks gene. Red represents viruses and green represents human. Protein-protein interactions are shown as dashed lines and regulations are shown as solid lines. Circle nodes with green border, red border or purple border are human miRNA targets, viral miRNA targets or targets co-regulated by viral and human miRNAs respectively.

Supplementary Text S2: Robustness analysis of the virus-human crosstalk principles mediated by miRNAs using targetome sequenced by PAR-CLIP.

Data

All EBV miRNA-target pairs and human miRNA-target interactions were obtained from a published EBV and cellular miRNA targetome in latently infected EBV-B95-8 lymphoblastoid cell lines by using photoactivatable-ribonucleoside enhance crosslinking and immunoprecipitation (PAR-CLIP)[2]. High-confidence miRNA-interaction sites in that study which were present in at least two libraries were analyzed in this work.

The EBV miRNA regulatory network

The EBV miRNA-target regulatory network was assembled between 13 EBV microRNAs and 630 human target genes. As shown in (Supplementary Figure S14A), 99.69% of all targets are connected to a large component. Interestingly, in the EBV miRNA-target regulatory network we found most EBV miRNAs can regulate more than one human transcripts (Supplementary Figure S14B) and there are also several cellular targets regulated by different EBV miRNAs. For example, these seven genes *G3BP1*, *PANK3*, *PTAR1*, *IRF1*, *HUWE1*, *NSD1* and *PSMD11* are regulated by three EBV miRNAs respectively. These findings indicated that miRNA-mediated interactions between EBV and human are complex. These results are accordant with those generated by validated viral miRNA targets data.
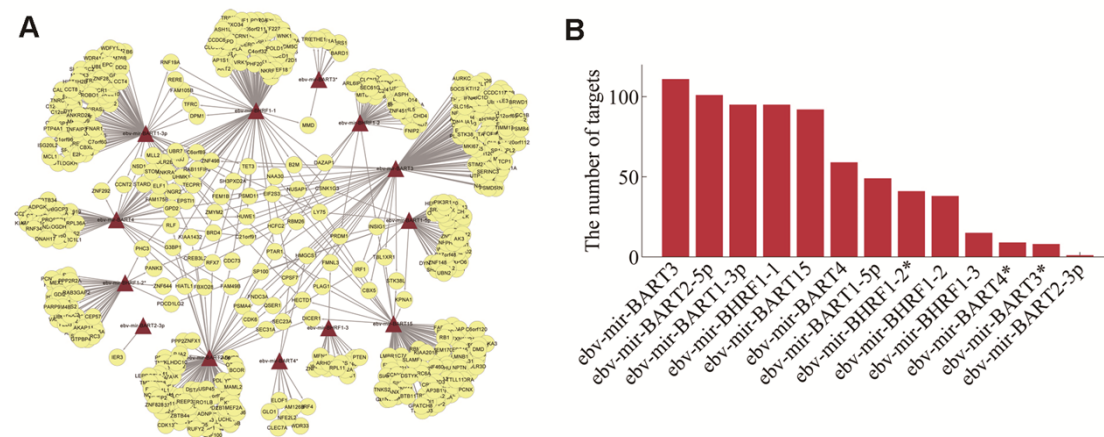
Supplementary Figure S14. The EBV miRNA-target regulatory network. (A) The EBV miRNA-target regulatory network consists of 13 EBV microRNAs and 630 human target genes. A triangle node marks miRNA and a circle node marks gene. (B) The number of cellular targets to individual EBV miRNAs.

The crosstalk between EBV and human miRNAs via their targets with similar functions

We explored the crosstalk between EBV and cellular miRNAs from the view of EBV and human miRNA-target regulatory network (abbreviated EHMiTN) which contains 28600 regulatory relationships between 3734 targets and 490 miRNAs including 477 human miRNAs and 13 EBV miRNAs as shown in Supplementary Figure S15A. Then, we evaluated the 'out-degree' of miRNAs, and the out-degree distribution of the EHMiTN reveals a power-law with a slope of -0.4733, but R^2=0.5459 (Supplementary Figure S15B). We also calculated the 'in-degree' of each target gene,

and the 'in-degree' distribution signifies an exponential distribution with an exponent of -1.762 and R^2=0.9795 (Supplementary Figure S15C). The phenomenon was found using validated viral miRNA targets data. On the other hand, 91% of transcripts targeted by EBV miRNAs are also regulated by human miRNAs, which is above ten times more than that of human miRNA targets in the whole genome (Supplementary Figure S15D P<1.0e-32, computed by hypergeometric test). Therefore, target genes of EBV miRNAs are also prone to be silenced by human miRNAs.
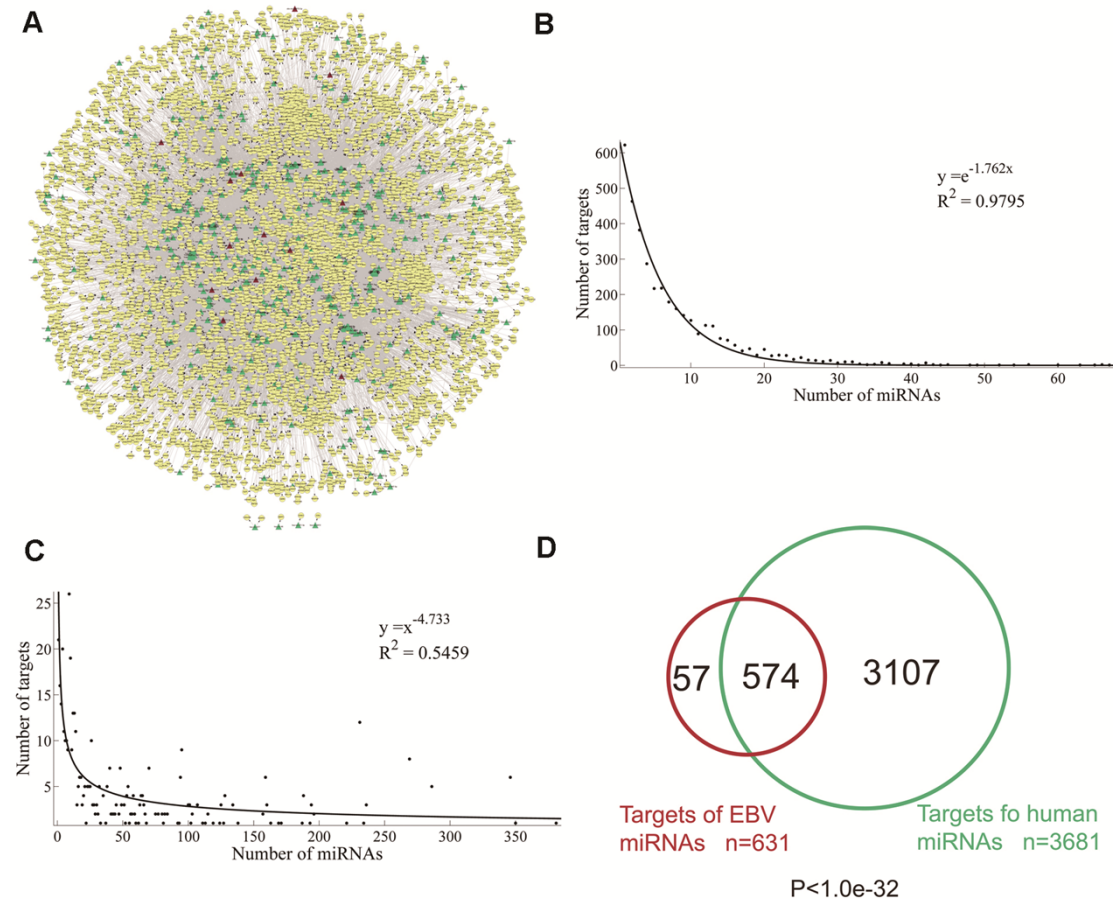
Supplementary Figure S15. The layout and topological features of the EHMiTN. (A) The EHMiTN consists of 28600 regulatory relationships between 490 miRNAs (477 human miRNAs and 13 EBV miRNAs) and 3734 targets. A triangle node marks miRNA and a circle node marks gene. (B) Out-degree distribution of the EHMiTN. (C) In-degree distribution of the EHMiTN. (D) Overlap of targets between EBV miRNAs and human miRNAs. Red represents EBV and green represents human.

Next, for targets regulated by EBV or human miRNAs, we computed their functional similarity and the extent of overlap between them respectively. Resultly, we found that the functional similarity score of these two kinds miRNAs reaches 0.64, significantly higher than random conditions (Supplementary Figure S16A, P<1.0e-3), indicating that cellular and EBV miRNAs tend to regulate genes with similar functions.
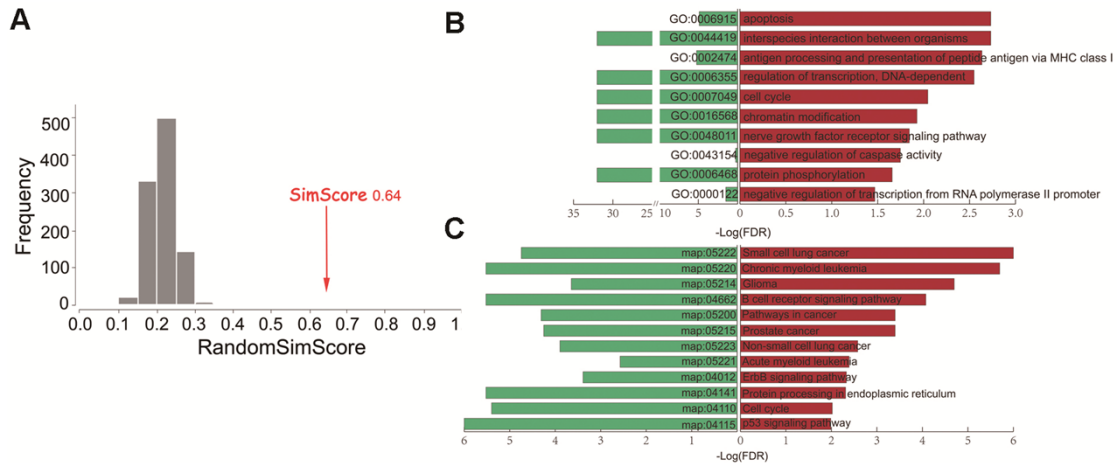
Supplementary Figure S16. EBV and Human miRNAs have close crosstalk via genes with important functions. (A) The functional similarity score between EBV and human miRNAs is larger than random conditions. The arrow represents the real functional similarity score between EBV and human miRNAs. (B) Selected GO terms with significant enrichments of EBV miRNA targets (red) and human miRNA targets (green). (C) Selected pathways with significant enrichments of EBV miRNA targets (red) and human miRNA targets (green). A threshold of FDR ≤ 0.05 was used to judge the significantly enriched GO terms or pathways enriched by at least five targets of EBV miRNAs.

Moreover, we performed enrichment analysis to discover specific biological themes significantly co-controlled by EBV- and human-encoded miRNAs and excavated functional crosstalk between them. Two kinds of common functional annotation data are analyzed, which are from GO and KEGG databases. Here, we only considered the ontology of biological process in the GO database. Supplementary Figure S16B summarizes the top (10) GO themes significantly enriched by at least five targets of EBV miRNAs. Most of these GO themes are associated with oncogenesis, including apoptosis, transcriptional regulation, cell cycle and so on. Unexpectedly, we found that almost all theses GO themes are also significantly enriched by target genes of human miRNAs. For example, 32 target genes are enriched in the term of cell cycle with FDR=0.009, while up to 151 human miRNA targets are also significantly enriched in this term with FDR<1.0e-32. Particularly, 26 genes from the 32 EBV miRNA targets are also regulated by human miRNAs. It is well known that a disregulation of the cell cycle components may lead to tumor formation. These miRNAs may co-control the cell cycle and cause the cell to multiply uncontrollably, forming tumor.

On the other hand, there are also several biological pathways significantly regulated by EBV-encoded miRNAs, all of which are also significantly regulated by human miRNAs. We found that most of these pathways are also related to signal transduction and cancers (Supplementary Figure S16C). These results are accordant with those generated by validated viral miRNA targets data. Fo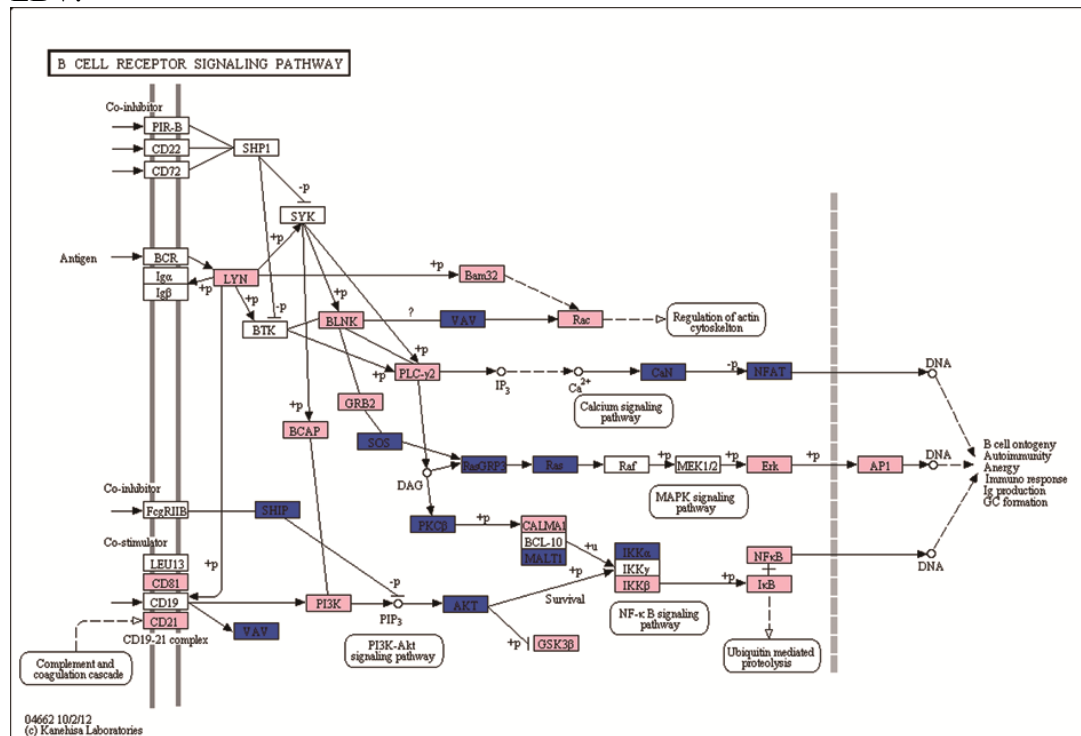r example, B cell receptor signaling pathway totally involving 13 targets of EBV miRNAs, this signaling ultimately results in the expression of immediate early genes that further activate the expression of other genes involved in B cell proliferation, differentiation and Ig production as well as other processes[3]. B cell receptor mediated signaling is important in the pathogenesis of a subset of diffuse large B cell lymphomas (DLBCL)[4]. Human miRNAs regulate most of the components of the pathway and 12 nodes of the pathway are co-regulated by EBV and human miRNAs(Supplementary Figure S17).

These results suggested that cooperating with cellular miRNAs to interfere in B Cell Receptor Signaling pathway might be one of major tumorigenesis mechanisms for EBV.

Supplementary Figure S17. B cell receptor signaling pathway enriched by targets of viral miRNA and human miRNA. The B cell receptor signaling pathway comes from KEGG. Genes targeted by human miRNAs are marked in pink, and genes co-regulated by EBV and human miRNAs are marked in blue.

Next, we deciphered the functional crosstalk of EBV miRNAs and human miRNAs via their cellular targets from two levels: association with diseases, and lethality. As a result, we found that, 57.37% (362/631) of cellular targets of EBV miRNAs are directly associated with diseases by analyzing disease data from the GAD database, significantly higher than random conditions (P=3.89e-11, computed by hypergeometric test, Supplementary Figure S18A). Human miRNA targets also significantly tend to be disease genes as shown in Supplementary Figure S18A. To further probe if EBV and human miNRAs are commonly involved in some specific disease categories, the information of disease classes in the database was considered and excess retention (ER) was calculated. As shown in Supplementary Figure S18B and S18C, cellular target genes of both EBV and human miRNAs are mainly associated with cancers, metabolic diseases, and immune diseases. Likewise, we got the same conclusion by using another set of disease data from the OMIM database (ER=2.29; P=7.53e-11, computed by hypergeometric test, Supplementary Figure S18A and S19). Cancer is the main disease class affected by EBV miRNAs in both disease datasets, and the corresponding scores of ER are 2.51 and 6.14 respectively. It has been known that EBV contributes to Burkitt's lymphoma, nasopharyngeal carcinoma and Hodgkin's disease[5]. So these EBV miRNAs may play some roles in cancer potential pathogenesis in the EBV-infected cell by targeting cancer-related genes.
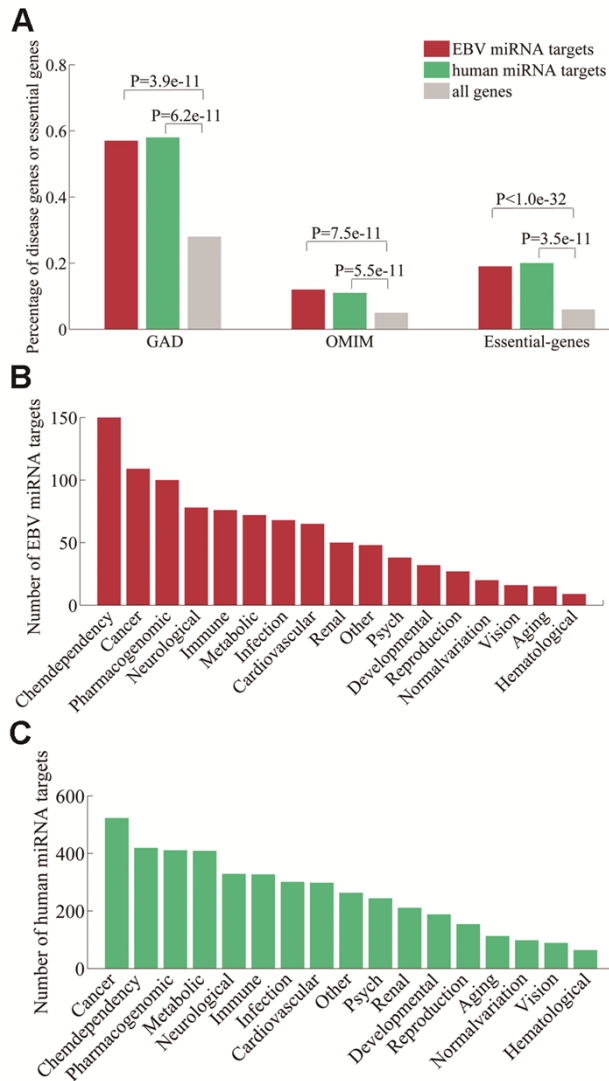
Supplementary Figure S18. Function of EBV miRNAs and human miRNAs association with diseases and lethality. (A) Comparison of the percentage of disease genes or essential genes. Disease genes from GAD and OMIM are shown respectively. (B) The number of EBV miRNA targets in 17 disease categories from GAD. (C) The number of human miRNA targets in 17 disease categories from GAD.
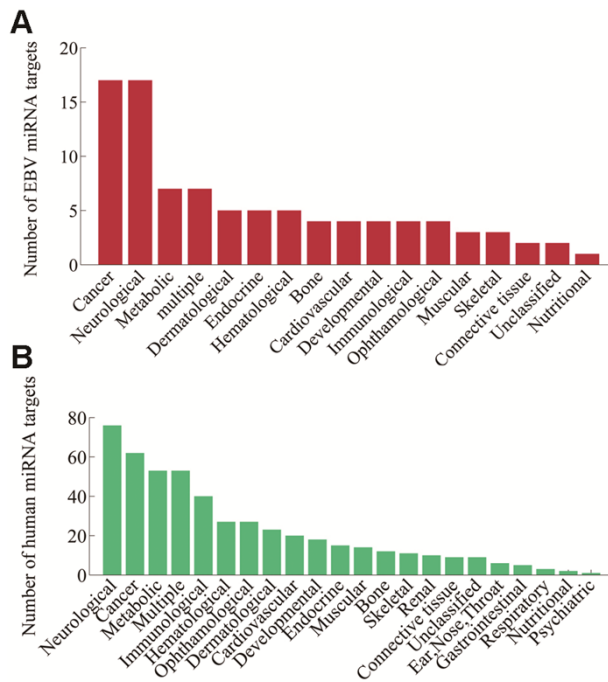
Supplementary Figure S19. Function of EBV miRNAs and human miRNAs association with diseases. (A) The number of EBV miRNA targets in 17 disease categories from OMIM. (B) The number of human miRNA targets in 22 disease categories from OMIM.

Lastly, we continued focusing on another kind of important genes, essential genes, to address whether both EBV and human miRNAs tend to regulate genes with important roles in host. As a result, a preference for targeting essential genes was observed for both EBV and human miRNAs(Supplementary Figure S18A P<1.0e-32, P=3.5e-11 for EBV and human miRNAs respectively, computed by hypergeometric test). For example, *G3BP1*, as an essential gene, has been found to be involved in a variety of growth-related signalling pathways and is targeted by three EBV miRNAs ebv-mir-BHRF1-1, ebv-mir-BART4 and ebv-mir-BART2-5p, meanwhile it is also regulated by 12 human miRNAs[6]. These EBV and human miRNAs co-repress the expression of *G3BP1* to regulate growth-related signalling pathways and are likely to contribute to the oncogenic potential of EBV.

The functional crosstalk between EBV and human miRNAs at the context of the PPIN

As discussed above, the regulations of EBV and human miRNAs are connected together to form a big post-transcriptional regulatory network and targets regulated by EBV and cellular miRNAs have high functional similarities. Thus, we next further explored the interaction patterns between these two types of target genes in the context of the PPIN for disclosing the crosstalk between EBV and human miRNAs.

First, we respectively calculated three topological centrality features of targets regulated by EBV and human miRNAs in the PPIN. As a result, targets of both EBV and human miRNAs have significantly higher degree, betweenness and closeness than random conditions, suggesting that target genes tend to be located at the center of the PPIN. Degree and closeness of EBV miRNA target genes respectively is 10 and 0.2503, both significantly higher than random (Supplementary Table S8). Betweenness of EBV miRNA targets is 37565, much lager than background. The results reveal a preference that EBV and host miRNAs tend to target hubs and bottlenecks. The tendency is found using validated viral miRNA targets data. We

furthermore observed an enrichment of targets regulated by EBV miRNAs in the top genes with high number of interactions (hubs), even when considering different gradients of degree (Supplementary Figure S20).

Supplementary Table S6 - Comparisons of topological features of EBV miRNA targets and human miRNA targets.

	HPRD	EBV_miRNA_targets			Human_miRNA_targets		
	Mean	Mean	expectation	p-values	Mean	expectation	p-values
Degree	7.9453	10.4100	7.9544	0.0007	10.4715	7.9455	<10e-4
Betweenness (*10^4)	2.9137	3.7565	2.9206	0.0886	4.4122	2.9147	<10e-4
Closeness	0.2310	0.2503	0.1062	<10e-4	0.2478	0.1063	<10e-4

p-value is calculated by the randomization test.
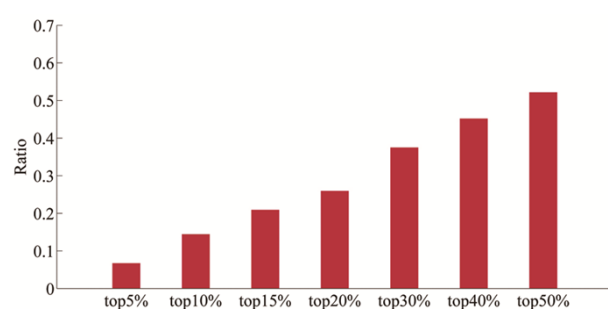
Supplementary Figure S20. The percentage of EBV miRNA targets in the hubs. Genes are ranked by the degree in the PPIN and hubs are defined as the top ranked genes.

As observed above, both EBV and host miRNAs tend to target genes located in the center of the PPIN. Next, we attempted to explore the efficience of communication between EBV miRNAs and human miRNAs based on the measurement of characteristic path length. Resultly, the characteristic path length between two kinds of target genes is 3.9188, significantly lower than random conditions, indicating that target genes of EBV miRNAs and those of human miRNAs are close proximity in the PPIN (Supplementary Figure S21A, P<10-4). Indeed, by calculating the number of direct interactions between these two types of targets, we found that there are up to 1577 interactions between them, also remarkable higher than random gene sets (Supplementary Figure S21B, P<10-4). In addition, the one-step neighbors of genes targeted by EBV miRNAs tend to be targets of human miRNAs, and the number reaches 1013 (Supplementary Figure S21C, P<10-4). These results all point that interactions between target genes provide necessary conditions for the crosstalk of EBV and human at miRNA level that is EBV miRNAs and human miRNAs are closely linked via physical interactions of their target genes.
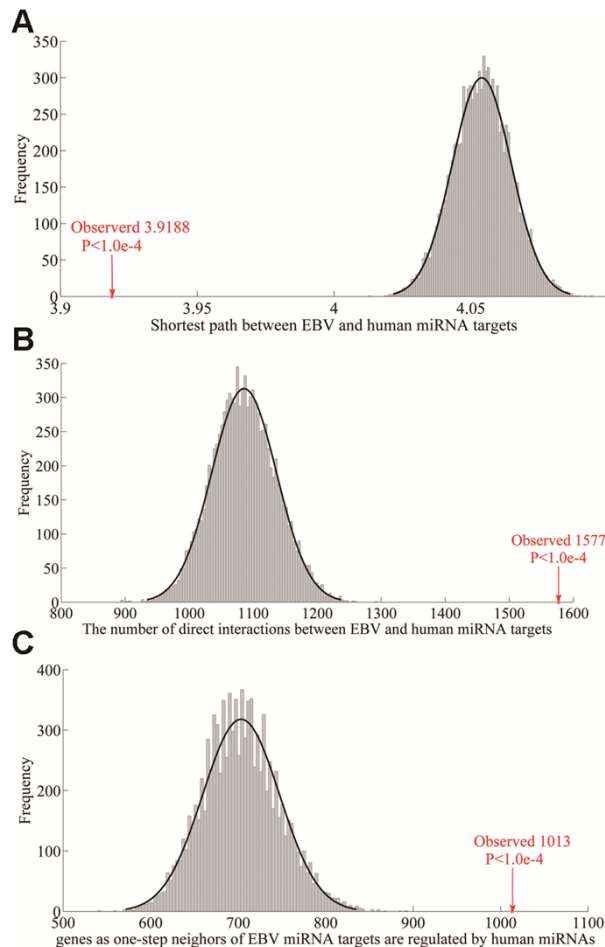
Supplementary Figure S21. Target genes of human and EBV miRNAs are close proximity in PPIN. (A) The characteristic path length between EBV miRNA targets and human miRNA targets is lower than random conditions. (B) The number of direct interactions between EBV miRNA targets and human miRNA targets is higher than random conditions. (C) The number of genes as one-step neighbours of EBV miRNA targets that are regulated by human miRNAs is higher than random conditions. The arrow represents the real value; the line is fitted using random distributions.

In the general context of the whole PPIN, several features of the crosstalk between EBV and human miRNAs are discovered. Next, we explored the local crosstalk between EBV and cellular miRNAs from the view of modular structure. Firstly, a subnetwork is constructed and consisted of two kinds of genes, targets of EBV miRNAs and their direct neighbors in the PPIN which are required to be regulated by at least one human miRNA. Interactions between these nodes are considered. Next, Markov Cluster algorithm is used to detect modules, and performed by its corresponding plugins in the cytoscape software by using the default parameters[1]. We found that these identified modules are involved in various important functions, including regulation of cell proliferation, regulation of apoptosis and transcription. After mapping miRNAs to these modules, we found that genes co-regulated by both EBV and human miRNAs tend to be in the important locations of modules, whose deletion would disrupt the integrity of the corresponding modules (Supplementary Figure S22A and B). These results are accordant with those generated by validated viral miRNA targets data. For example, a module including seventeen target genes is mainly involved in regulation of transcription. Especially, *RB1* is located at the center of this module and regulated by ebv-mir-BART15 and 26 human miRNAs at the same

time, which is a negative regulator of the cell cycle and a tumor suppressor gene (Supplementary Figure S22A). It was found that the module included 20 lymphoma-related disease miRNAs, up to 48% of the known, ten of which are Hodgkin Disease miRNAs. This could be explained by that EBV and human contest for controlling a number of important functions through competing for inhibition of center of functional modules. The above results indicate that regulation and control of EBV miRNAs and human miRNAs to this module play a catalytic role in lymphoma, so we speculated that after virus infection, some human miRNAs and EBV miRNAs have accumulate (synergistic) effects in disease states, and jointly promote occurrence of diseases. By calculating the proportion of disease miRNAs in every module, we found that most of human miRNAs are included in the mir2disease [7]and HMDD databases[88](data not shown). Therefore, these results suggest that EBV miRNAs tend to have crosstalk with cellular miRNAs, especially disease miRNAs.
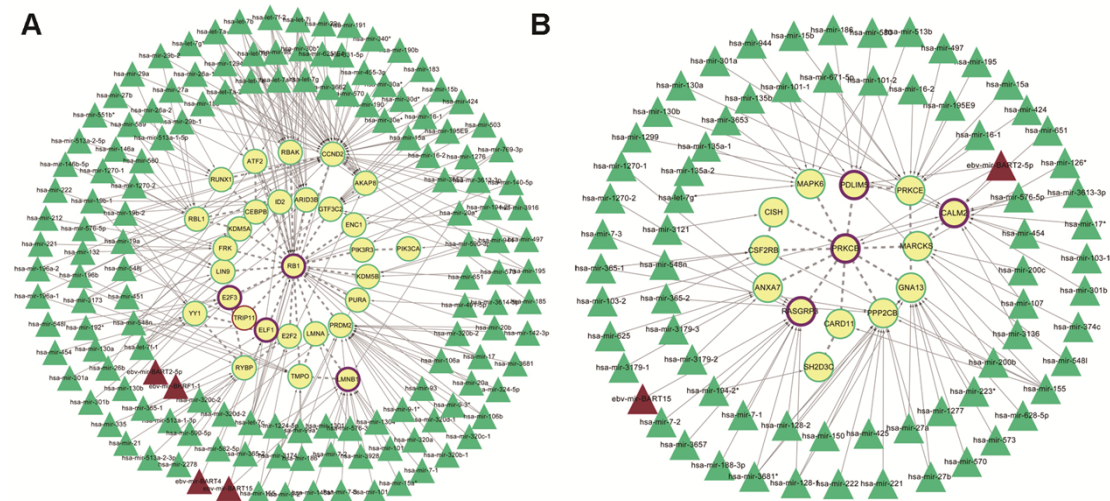
Supplementary Figure S22. Two composite modules composed of EBV miRNAs, human miRNAs, targets and interaction between targets. A triangle node marks miRNA and a circle node marks gene. Red represents EBV and green represents human. Protein-protein interactions are shown as dashed lines and regulations are shown as solid lines. Circle nodes with green border, red border or purple border are human miRNA targets, EBV miRNA targets or targets co-regulated by EBV and human miRNAs respectively.

Supplementary Text S3: Robustness analysis of the virus-human crosstalk principles mediated by miRNAs using predicted data.

Data

In order to test the validation of the main results, here we analyzed the predicted viral miRNA targets data. The predicted vial miRNA target data came from the VmiReg database. In order to reduce false positives, we used intersections of predicted results from the four algorithms miRanda, TargetScan, RNAhybrid and PITA. Totally, there were 8283 regulations between 95 viral miRNAs and 4381 human genes. The predicted human miRNA target data were downloaded from miRecords database. We also used intersections of predicted results from the same four algorithms with viral miRNA. Totally, there were 204768 regulations between 416 human miRNAs and 14595 human genes.

The crosstalk between viral and human miRNAs via their targets with similar functions

We first explored the crosstalk between viral and cellular miRNAs from the view of co-regulation. 83% of transcripts targeted by viral miRNAs are also regulated by human miRNAs. Therefore, target genes of viral miRNAs are also prone to be silenced by human miRNAs, implying that the means of the crosstalk between viral and human miRNAs might be via co-regulating targets.

   Moreover, we performed enrichment analysis to discover biological themes significantly co-controlled by viral and human miRNAs and excavated functional crosstalk between them. Two kinds of common functional annotation data were analyzed, which were from GO and KEGG databases. Here, we only considered the ontology of biological process in the GO database. Figure S23A summarizes the GO themes significantly enriched by at least three targets of viral miRNAs. Most of these GO themes are associated with signalling transduction, and transcriptional regulation. Unexpectedly, we found that almost all these GO themes are also significantly enriched by targets of human miRNAs. On the other hand, there are also several biological pathways significantly regulated by virus-encoded miRNAs. We found that most of these pathways are related to signal transduction and oncogenesis (Figure S23B). Most of these pathways are significantly regulated by human miRNAs. These results indicate that some specific important biological processes controlled by human miRNAs might be simultaneously interfered by virus-encoded miRNAs for their successful infection.
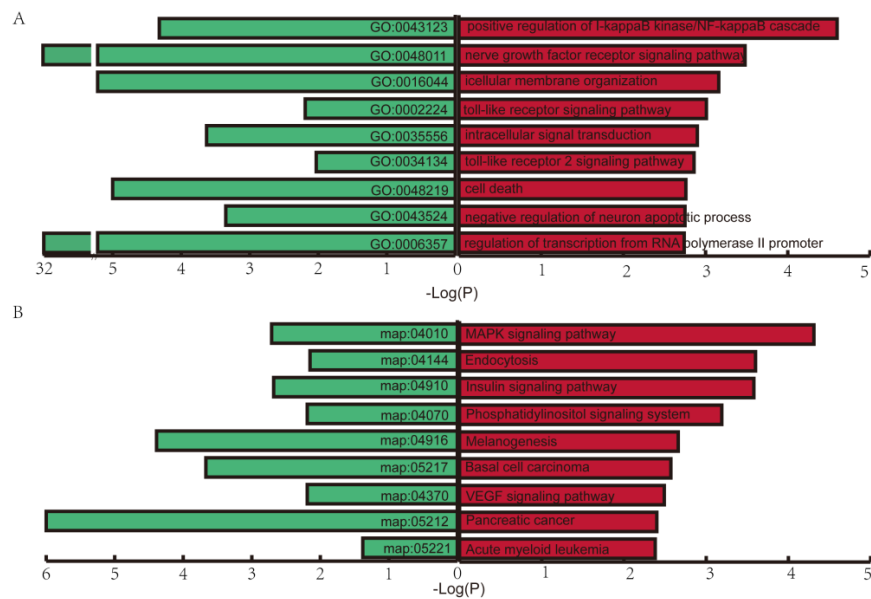
Supplementary Figure S23. Viral and Human miRNAs have close crosstalk via genes with important functions. (A) Selected GO terms with significant enrichments of viral miRNA targets (red) and human miRNA targets (green). (B) Selected pathways with significant enrichments of viral miRNA targets (red) and human miRNA targets (green). A threshold of FDR ≤ 0.05 was used to judge the significantly enriched GO terms or pathways enriched by at least three targets of viral miRNA.

Next, we systematically explored the crosstalk between viral and human miRNAs in the context of diseases and found that viral miRNAs significantly regulate disease genes by analyzing disease data from the GAD or OMIM database (P<1.0e-32 and <1.0e-32 respectively, computed by hypergeometric test, Supplementary Figure S24). As a result, we found that 54% of viral miRNA human targets are directly associated with diseases by analyzing disease data from the GAD database, significant higher than random conditions (P<1.0e-32, computed by hypergeometric test, Supplementary Figure S24). As shown in Figure S24, cellular target genes of both viral and human miRNAs are mainly associated with cancers, metabolic diseases, immune diseases and cardiovascular diseases. Likewise, we got the same conclusion by using another set of disease data from the OMIM database (see Supplementary Figure S24). Lastly, we continued to focus on another kind of important genes, essential genes. Dysfunctions of essential genes would lead to death or severe disease phenotype. As a result, a preference for targeting essential genes was observed for both viral and human miRNAs (Figure S24 P<1.0e-32, computed by hypergeometric test).
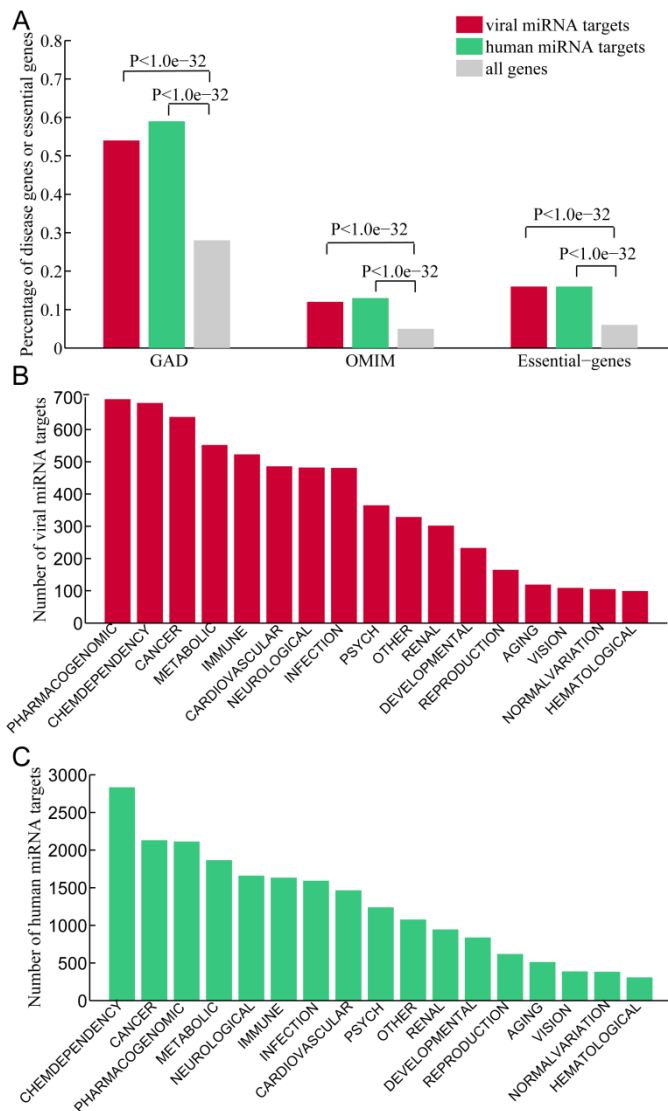
Supplementary Figure S24. Function of viral miRNAs and human miRNAs association with diseases and lethality. (A) Comparison of the percentage of disease genes or essential genes. Disease genes from GAD and OMIM are shown respectively. (B) The number of viral miRNA targets in 17 disease categories from GAD.

The functional crosstalk between viral and human miRNAs in the context of the PPIN

First, we respectively calculated three topological centrality features of targets regulated by viral and human miRNAs in the PPIN, including degree, betweenness, and closeness. As a result, targets of both viral and human miRNAs have higher degree, betweenness, and closeness than random conditions, suggesting that target genes tend to be located at the center of the PPIN (Table S9). Degree and betweenness of viral miRNA targets are respectively larger than those of human miRNA targets.

Supplementary Table S7. Comparisons of topological features of viral miRNA targets and human miRNA targets

	HPRD	Viral_miRNA_targets			Human_miRNA_targets		
	Mean	Mean	expectation	p-value	Mean	expectation	p-value
Degree	7.9453	8.5875	7.9442	0.0089	8.4174	7.9454	<10e-4

Betweenness (*10^4)	2.9137	3.1853	2.9123	0.1088	3.1361	2.9140	<10e-4
Closeness	0.2310	0.2426	0.2407	0.0003	0.2423	0.2407	<10e-4

Note: p-value is calculated by the randomization test. Expectation is mean in random cases.

Next, we attempted to explore the efficiency of communication between viral miRNAs and human miRNAs based on the measurement of characteristic path length, which represents close proximity and consequently how quickly information can spread in a network. The characteristic path length between two kinds of target genes is 4.166 significantly lower than random conditions, indicating that target genes of viral miRNAs and those of human miRNAs are close proximity in the PPIN (Figure S25A, P<10-4). Indeed, by further calculating the number of direct interactions between these two types of targets, we found that there are up to 17089 interactions between them, also remarkably higher than random gene sets (Figure S25B, P<10-4). In addition, although the genes co-regulated by viral and human miRNAs are limited, the one-step neighbours of genes targeted by viral miRNAs tend to be targets of human miRNAs, and the proportion reaches 84% (Figure S25C, P<1.0e-4).
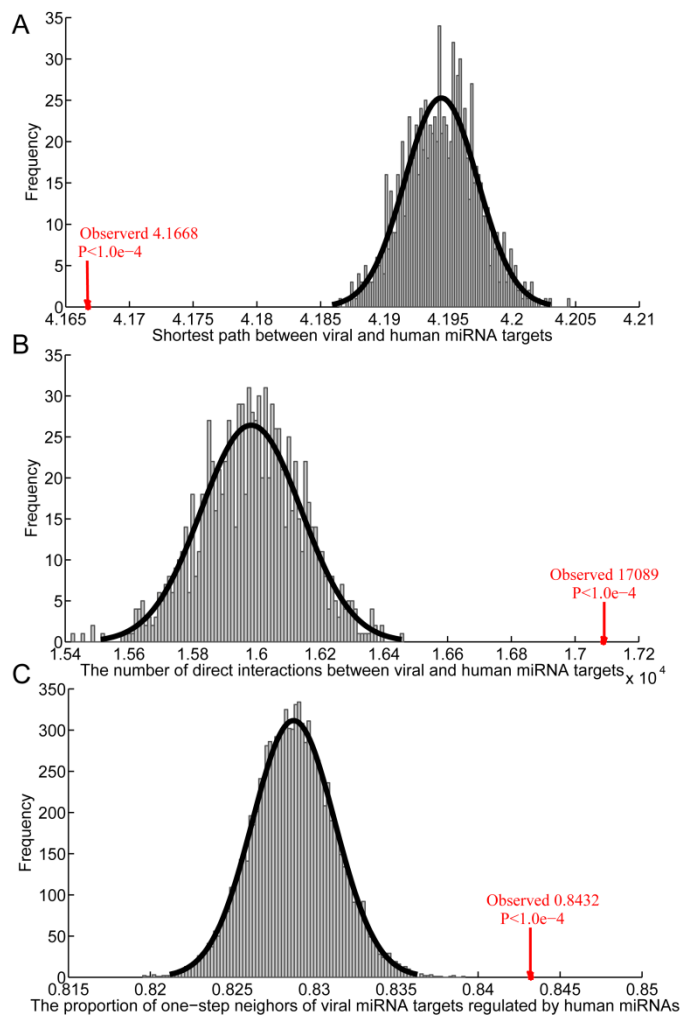
Supplementary Figure S21. Target genes of human and viral miRNAs are close proximity in PPIN. (A) The characteristic path length between viral miRNA targets and human miRNA targets is lower than random conditions. (B) The number of direct interactions between viral miRNA targets and human miRNA targets is higher than random conditions. (C) The proportion of genes as one-step neighbours of viral miRNA targets regulated by human miRNAs is higher than random conditions. The

arrow represents the real value; the line is fitted using random distributions.

REFERENCES

1.   S. Van Dongen, *Siam J Matrix Analysis Applications* 2008, 30, 121-141.
2.   R. L. Skalsky, D. L. Corcoran, E. Gottwein, C. L. Frank, D. Kang, M. Hafner, J. D. Nusbaum, R. Feederle, H. J. Delecluse, M. A. Luftig, T. Tuschl, U. Ohler and B. R. Cullen, *PLoS Pathog*, 2012, 8, e1002484.
3.   M. Kanehisa, S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki and M. Hirakawa, *Nucleic Acids Res*, 2006, 34, D354-357.
4.   A. M. Bogusz, R. H. Baxter, T. Currie, P. Sinha, A. R. Sohani, J. L. Kutok and S. J. Rodig, *Clin Cancer Res*, 2012, 18, 6122-6135.
5.   S. A. Rezk and L. M. Weiss, *Hum Pathol*, 2007, 38, 1293-1304.
6.   M. M. Kim, D. Wiederschain, D. Kennedy, E. Hansen and Z. M. Yuan, *Oncogene*, 2007, 26, 4209-4215.
7.   Q. Jiang, Y. Wang, Y. Hao, L. Juan, M. Teng, X. Zhang, M. Li, G. Wang and Y. Liu, *Nucleic acids research*, 2009, 37, D98-104.
8.   M. Lu, Q. Zhang, M. Deng, J. Miao, Y. Guo, W. Gao and Q. Cui, *PloS one*, 2008, 3, e3420.