

SUPPLEMENTARY DATA

The Application of miRTarBase in Others Database and Web tools

Table S2 lists the databases that incorporate miRTarBase which are categorized into six different groups 1) well-known databases; 2) information databases for MTIs; 3) MiRNA target prediction databases; 4) SNP or mutation related databases, 5) pathway related resources, and 6) others database or web tools. In category 1, the well-known databases, including miRBase (1), GeneCards (2), Ensembl (3) and Mouse Genome Database MGI (4) are kept which collect MTIs from miRTarBase as empirical evidence. TarBase (5) is included in category 2, and is considered to be another provider of complete MTIs. Furthermore, the empirical data in miRTarBase are also taken from validated criteria in relevant miRNA databases such as starBase (6), miRGator (7) and TSmiR (8), in terms of CLIP-seq data, gene expression profiles, and tissue specificity, respectively. In category 3, miRWalk2.0 documents the miRNA binding regions from existing target prediction programs to build comparative platforms of binding sites for the promoter, CDS and UTR regions. It also documents experimentally verified MTI information collected from miRTarBase (9). miRGate is a database of human, virus, mouse and rat MTI which is designed to analyze miRNA and gene isoform lists under a common and consistent set of annotations. Our database contains validated MTIs that were included in miRGate as experimentally validated data (10). mirMark is a tool based on machine learning method for miRNA target prediction at the site of UTR regions. The positive dataset used in mirMark is obtained from miRTarBase (11). With regard to the SNP or mutation related databases in category 4, PolymiRTS is a database of DNA variations in the miRNA seed regions of miRNA, as well as experimentally supported MTIs collected from miRecords and miRTarBase (12). SomamiR is a comprehensive database of somatic mutations in cancer from miRNA and target sites, which collects the MTIs from TarBase, miRecords and miRTarBase (13). miRTarBase has been incorporated into an increasing number of biological resources for pathway- and gene regulatory network related resources, such as miRPathDB (14), DIANA-miRPath (15), miRNet (16), and associations of miRNA related

disease, regulatory circuits, and cancer, such as HMDD (17), MAGIA2 (18), NCG (19), which are categorized in group 5 and 6. In brief, miRTarBase performs a significant role in providing validated experimental data. We believe the release of new version will provide great support in updating the databases mentioned above.

SUPPLEMENTARY FIGURES

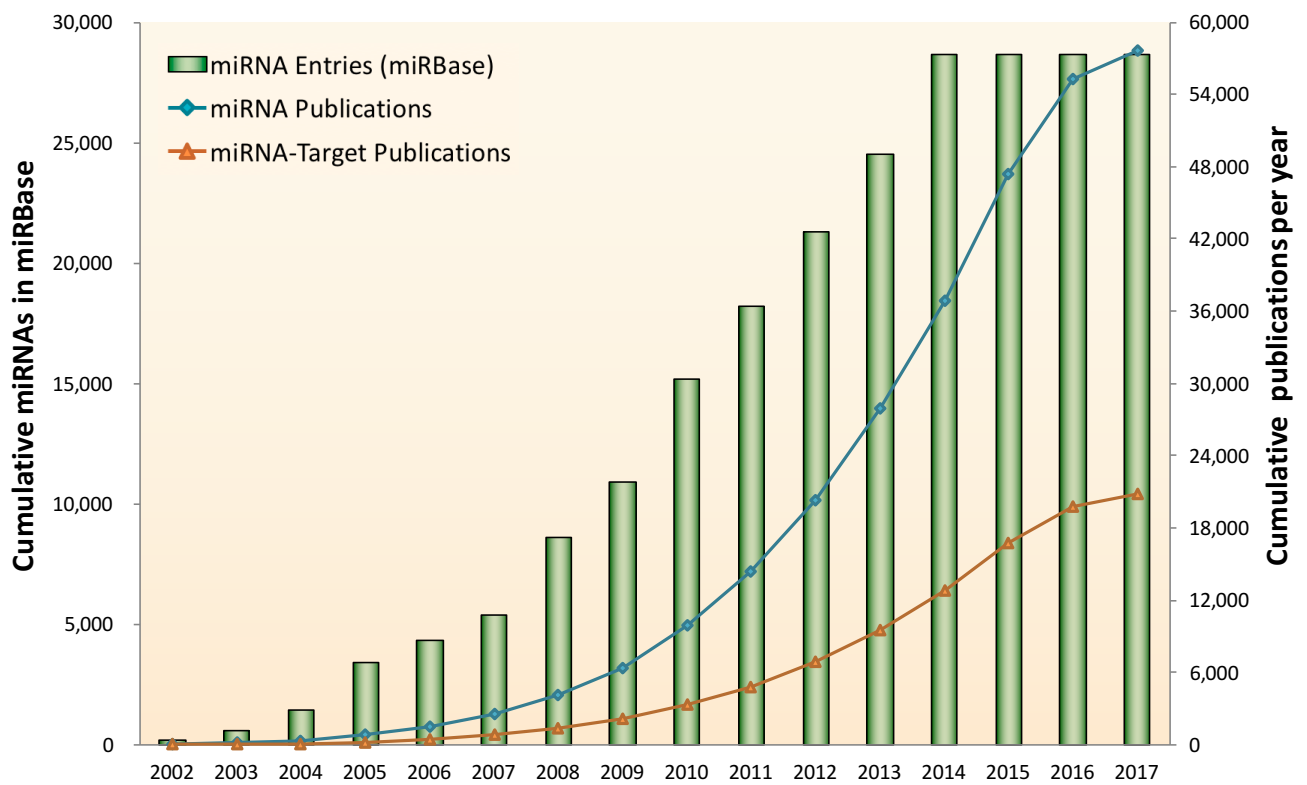


Figure S1. The growth of miRNA and it's publication.

The amount of miRNAs collected in miRBase, miRNA publications, and miRNA-target interaction publication collected in PubMed since 2002.

SUPPLEMENTARY TABLES

Table S1. Update MTI datasets validated by high throughput HITS-CLIP or PAR-CLIP sequencing

Source	NGS method	Species	RBP	Accession	Samples	Tissue/cell line	Treatment/Name
Luna, J. M., et al. Mol Cell. 2017 (20)	HITS-CLIP	Mouse	AGO2	GSE97058	GSM2550642	Liver	WT Female 1 CLIP
					GSM2550643		WT Female 2 CLIP
					GSM2550644		WT Female 3 CLIP
					GSM2550645		WT Male 4 CLIP
					GSM2550646		WT Male 5 CLIP
					GSM2550647		KO Female 1 CLIP
					GSM2550648		KO Male 2 CLIP
					GSM2550649		KO Male 3 CLIP
					GSM2550650		KO Male 4 CLIP
					Human		AGO2
		GSM2550605	Patient 1 Normal 2				
		GSM2550606	Patient 2 Normal 1				
		GSM2550607	Patient 2 Normal 2				
		GSM2550608	Patient 3 Normal 1				
		GSM2550609	Patient 3 Normal 2				
		GSM2550610	Patient 4 Normal 1				
		GSM2550611	Patient 4 Normal 2				
		GSM2550612	Patient 5 Normal 1				
		GSM2550613	Patient 5 Normal 2				
		GSM2550614	Patient 6 Normal 1				
GSM2550615	Patient 7 Normal 1						

					GSM2550616		Patient 8 Normal 1
					GSM2550617		Patient 9 Normal 1
					GSM2550618		Patient 1 Tumor 1
					GSM2550619		Patient 1 Tumor 2
					GSM2550620		Patient 2 Tumor 1
					GSM2550621		Patient 2 Tumor 2
					GSM2550622		Patient 3 Tumor 1
					GSM2550623		Patient 3 Tumor 2
					GSM2550624		Patient 4 Tumor 1
					GSM2550625		Patient 4 Tumor 2
					GSM2550626		Patient 5 Tumor 1
					GSM2550627		Patient 5 Tumor 2
					GSM2550628		Patient 6 Tumor 1
					GSM2550629		Patient 7 Tumor 1
					GSM2550630		Patient 8 Tumor 1
					GSM2550631		Patient 9 Tumor 1
Hamilton, M. P., et al. Neoplasia. 2016 (21)	PAR-CLIP	Human	AGO2	SRP075075	SRX1760641	LNCaP	AGO-CLIP-LNCaP-MDV_B
					SRX1760639	LNCaP	AGO-CLIP-LNCaP-MDV_A
					SRX1760638	PC3	AGO-CLIP-PC3-miR148
					SRX1760637	DU145	AGO-CLIP-DU145_A
					SRX1760632	22RV1	AGO-CLIP-22RV1_C
					SRX1760631	22RV1	AGO-CLIP-22RV1_B
					SRX1760630	22RV1	AGO-CLIP-22RV1_A
					SRX1760628	LAPC4	AGO-CLIP-LAPC4_B
					SRX1760620	LAPC4	AGO-CLIP-LAPC4_A
					SRX1760618	PC3	AGO-CLIP-PC3_B

					SRX1760616	PC3	AGO-CLIP-PC3_A
					SRX1760597	LNCaP	AGO-CLIP-LNCaP_C
					SRX1760591	LNCaP	AGO-CLIP-LNCaP_B
					SRX1760583	LNCaP	AGO-CLIP-LNCaP_A
Krell, J., et al. Genome Res. 2016 (22)	HITS-CLIP	Human	AGO2	PRJEB3396	ERX177599	HCT116	p53_D_AGO_CLIP_2_1
					ERX177600		p53_V_Ago_CLIP_2_2
					ERX177601		KO_D_Ago_CLIP_2_3
					ERX177602		KO_V_AGO_CLIP_2_4
					ERX177603		p53_D_AGO_CLIP_2_5
					ERX177604		p53_V_AGO_CLIP_2_6
					ERX177605		KO_D_AGO_CLIP_2_7
					ERX177606		KO_V_AGO_CLIP_2_8
					ERX177607		p53_D_AGO_CLIP_2_9
					ERX177608		p53_V_AGO_CLIP_2_10
					ERX177609		KO_D_AGO_CLIP_2_11
					ERX177610		KO_V_AGO_CLIP_2_12
					ERX177611		p53_D_AGO_CLIP_3_1
					ERX177612		p53_V_AGO_CLIP_3_2
					ERX177613		KO_D_AGO_CLIP_3_3
					ERX177614		KO_V_AGO_CLIP_3_4
					ERX177615		p53_D_AGO_CLIP_3_5
					ERX177616		p53_V_AGO_CLIP_3_6
					ERX177617		KO_D_AGO_CLIP_3_7
					ERX177618		KO_V_AGO_CLIP_3_8
ERX177619	p53_D_AGO_CLIP_3_9						
ERX177620	p53_V_AGO_CLIP_3_10						

					ERX177621		KO_D_AGO_CLIP_3_11
					ERX177622		KO_V_AGO_CLIP_3_12
					ERX177623		p53_D_AGO_CLIP_4_1
					ERX177624		p53_V_AGO_CLIP_4_2
					ERX177625		KO_D_AGO_CLIP_4_3
					ERX177626		KO_V_AGO_CLIP_4_4
					ERX177627		p53_D_AGO_CLIP_4_5
					ERX177628		p53_V_AGO_CLIP_4_6
					ERX177629		KO_D_AGO_CLIP_4_7
					ERX177630		KO_V_AGO_CLIP_4_8
					ERX177631		p53_D_AGO_CLIP_4_9
					ERX177632		p53_V_AGO_CLIP_4_10
					ERX177633		KO_D_AGO_CLIP_4_11
					ERX177634		KO_V_AGO_CLIP_4_12
Spengler, R. M., et al. Nucleic Acids Res 2016 (23)	HITS-CLIP	Human	AGO2	GSE83410	GSM2202476	Cardiac tissue	S1_LV_54yo_Male_AGO2_bound_RNA
					GSM2202477		S2_LV_25yo_Male_AGO2_bound_RNA
					GSM2202478		S3_LV_36yo_Male_AGO2_bound_RNA
					GSM2202479		S4_LV_29yo_Male_AGO2_bound_RNA
					GSM2202480		S5_LV_36yo_Male_AGO2_bound_RNA
					GSM2202481		S6_LV_61yo_Male_AGO2_bound_RNA

Table S2. A list of resources that incorporate miRTarBase.

Type of resources	Database names
Well-known databases	miRBase (1), GeneCards (2), Ensembl (3), MGI (4)
Information databases for MTIs	Tarbase (5), starBase (6), miRGator (7), TSmiR (8)
MiRNA target prediction databases	miRWalk (9), miRGate (10), mirMark (11)
SNP or mutation related databases	PolymiRTS (12), SomamiR (13)
Pathway related resources	MiRPathDB (14), DIANA-miRPath (15), miRNet (16)
Others Database or Web tools	HMDD (17), MAGIA2 (18), NCG (19)

REFERENCES

1. Kozomara, A. and Griffiths-Jones, S. (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*, **42**, D68-73.
2. Safran, M., Dalah, I., Alexander, J., Rosen, N., Iny Stein, T., Shmoish, M., Nativ, N., Bahir, I., Doniger, T., Krug, H. *et al.* (2010) GeneCards Version 3: the human gene integrator. *Database (Oxford)*, **2010**, baq020.
3. Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S., Gil, L. *et al.* (2016) Ensembl 2016. *Nucleic Acids Res*, **44**, D710-716.
4. Eppig, J.T., Blake, J.A., Bult, C.J., Kadin, J.A., Richardson, J.E. and Mouse Genome Database, G. (2015) The Mouse Genome Database (MGD): facilitating mouse as a model for human biology and disease. *Nucleic Acids Res*, **43**, D726-736.
5. Vlachos, I.S., Paraskevopoulou, M.D., Karagkouni, D., Georgakilas, G., Vergoulis, T., Kanellos, I., Anastasopoulos, I.L., Maniou, S., Karathanou, K., Kalfakakou, D. *et al.* (2015) DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. *Nucleic Acids Res*, **43**, D153-159.
6. Li, J.H., Liu, S., Zhou, H., Qu, L.H. and Yang, J.H. (2014) starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*, **42**, D92-97.
7. Cho, S., Jang, I., Jun, Y., Yoon, S., Ko, M., Kwon, Y., Choi, I., Chang, H., Ryu, D., Lee, B. *et al.* (2013) MiRGator v3.0: a microRNA portal for deep sequencing, expression profiling and mRNA targeting. *Nucleic Acids Res*, **41**, D252-257.
8. Guo, Z., Maki, M., Ding, R., Yang, Y., Zhang, B. and Xiong, L. (2014) Genome-wide survey of tissue-specific microRNA and transcription factor regulatory networks in 12 tissues. *Sci Rep*, **4**, 5150.
9. Dweep, H., Sticht, C., Pandey, P. and Gretz, N. (2011) miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. *Journal of biomedical informatics*, **44**, 839-847.
10. Andres-Leon, E., Gonzalez Pena, D., Gomez-Lopez, G. and Pisano, D.G. (2015) miRGate: a curated database of human, mouse and rat miRNA-mRNA targets. *Database (Oxford)*, **2015**, bav035.
11. Menor, M., Ching, T., Zhu, X., Garmire, D. and Garmire, L.X. (2014) mirMark: a site-level and UTR-level classifier for miRNA target prediction. *Genome Biol*, **15**, 500.
12. Bhattacharya, A., Ziebarth, J.D. and Cui, Y. (2014) PolymiRTS Database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways. *Nucleic Acids Res*, **42**, D86-91.
13. Bhattacharya, A. and Cui, Y. (2016) SomamiR 2.0: a database of cancer somatic mutations altering microRNA-ceRNA interactions. *Nucleic Acids Res*, **44**, D1005-1010.
14. Backes, C., Kehl, T., Stockel, D., Fehlmann, T., Schneider, L., Meese, E., Lenhof, H.P. and Keller, A. (2017) miRPathDB: a new dictionary on microRNAs and target pathways. *Nucleic Acids Res*, **45**, D90-D96.
15. Vlachos, I.S., Zagkanas, K., Paraskevopoulou, M.D., Georgakilas, G., Karagkouni, D., Vergoulis, T., Dalamagas, T. and Hatzigeorgiou, A.G. (2015) DIANA-miRPath v3.0: deciphering microRNA function with experimental support. *Nucleic Acids Res*, **43**, W460-466.
16. Fan, Y., Siklenka, K., Arora, S.K., Ribeiro, P., Kimmins, S. and Xia, J. (2016) miRNet - dissecting miRNA-target interactions and functional associations through network-based visual analysis. *Nucleic Acids Res*, **44**, W135-141.
17. Li, Y., Qiu, C., Tu, J., Geng, B., Yang, J., Jiang, T. and Cui, Q. (2014) HMDD v2.0: a database for experimentally supported human microRNA and disease associations. *Nucleic Acids Res*, **42**, D1070-1074.
18. Bisognin, A., Sales, G., Coppe, A., Bortoluzzi, S. and Romualdi, C. (2012) MAGIA(2): from miRNA and genes expression data integrative analysis to microRNA-transcription factor mixed regulatory circuits (2012 update). *Nucleic Acids Res*, **40**, W13-21.
19. An, O., Dall'Olio, G.M., Mourikis, T.P. and Ciccarelli, F.D. (2016) NCG 5.0: updates of a manually curated repository of cancer genes and associated properties from cancer mutational screenings. *Nucleic Acids Res*, **44**, D992-999.
20. Luna, J.M., Barajas, J.M., Teng, K.Y., Sun, H.L., Moore, M.J., Rice, C.M., Darnell, R.B. and Ghoshal, K.

- (2017) Argonaute CLIP Defines a Deregulated miR-122-Bound Transcriptome that Correlates with Patient Survival in Human Liver Cancer. *Molecular cell*, **67**, 400-410.e407.
21. Hamilton, M.P., Rajapakshe, K.I., Bader, D.A., Cerne, J.Z., Smith, E.A., Coarfa, C., Hartig, S.M. and McGuire, S.E. (2016) The Landscape of microRNA Targeting in Prostate Cancer Defined by AGO-PAR-CLIP. *Neoplasia (New York, N.Y.)*, **18**, 356-370.
 22. Krell, J., Stebbing, J., Carissimi, C., Dabrowska, A.F., de Giorgio, A., Frampton, A.E., Harding, V., Fulci, V., Macino, G., Colombo, T. *et al.* (2016) TP53 regulates miRNA association with AGO2 to remodel the miRNA-mRNA interaction network. *Genome research*, **26**, 331-341.
 23. Spengler, R.M., Zhang, X., Cheng, C., McLendon, J.M., Skeie, J.M., Johnson, F.L., Davidson, B.L. and Boudreau, R.L. (2016) Elucidation of transcriptome-wide microRNA binding sites in human cardiac tissues by Ago2 HITS-CLIP. *Nucleic acids research*, **44**, 7120-7131.