

# SCIENTIFIC REPORTS



OPEN

## Psmir: a database of potential associations between small molecules and miRNAs

Fanlin Meng\*, Jing Wang\*, Enyu Dai, Feng Yang, Xiaowen Chen, Shuyuan Wang, Xuexin Yu, Dianming Liu & Wei Jiang

Received: 22 October 2015

Accepted: 10 December 2015

Published: 13 January 2016

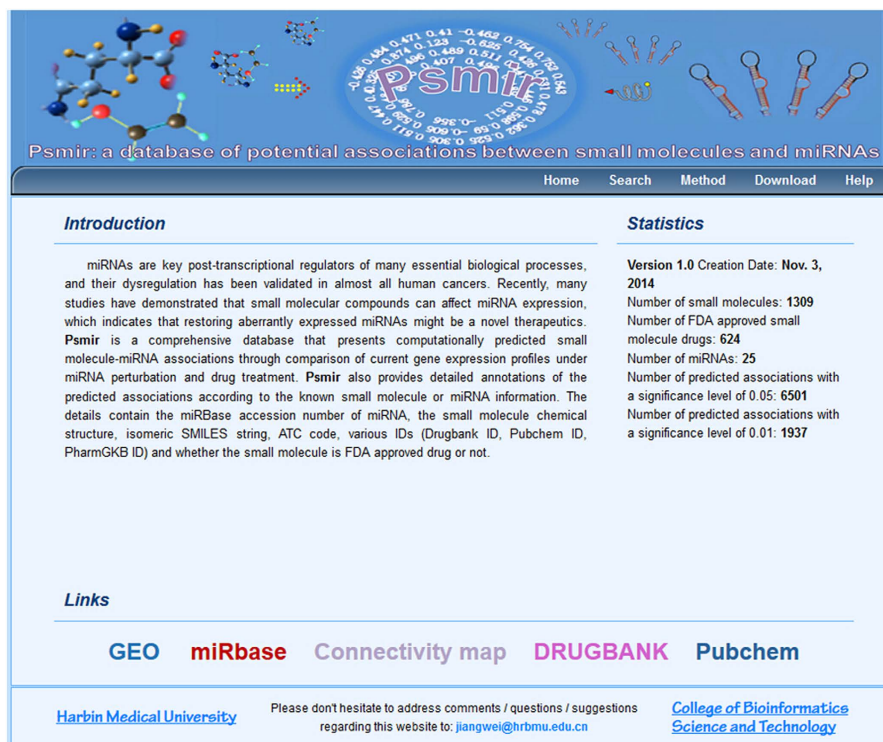
miRNAs are key post-transcriptional regulators of many essential biological processes, and their dysregulation has been validated in almost all human cancers. Restoring aberrantly expressed miRNAs might be a novel therapeutics. Recently, many studies have demonstrated that small molecular compounds can affect miRNA expression. Thus, prediction of associations between small molecules and miRNAs is important for investigation of miRNA-targeted drugs. Here, we analyzed 39 miRNA-perturbed gene expression profiles, and then calculated the similarity of transcription responses between miRNA perturbation and drug treatment to predict drug-miRNA associations. At the significance level of 0.05, we obtained 6501 candidate associations between 1295 small molecules and 25 miRNAs, which included 624 FDA approved drugs. Finally, we constructed the Psmir database to store all potential associations and the related materials. In a word, Psmir served as a valuable resource for dissecting the biological significance in small molecules' effects on miRNA expression, which will facilitate developing novel potential therapeutic targets or treatments for human cancers. Psmir is supported by all major browsers, and is freely available at <http://www.bio-bigdata.com/Psmir/>.

Dysregulation of miRNAs has been implicated in a plethora of diseases, giving miRNAs great potential in both cancer diagnostics and treatment. Modulation of miRNA expression levels has been demonstrated as a viable strategy for tumor therapeutics<sup>1</sup>. The well studied miRNA modulators include antisense oligonucleotides (anti-miRs), antagomirs, miRNA sponges, and so on<sup>2</sup>. However, the inefficient delivery into target tissues and suboptimal pharmacodynamics or pharmacokinetics properties are major hurdles in the oligonucleotide-based therapeutics, which highlighting the need and importance for small-molecule-based intervention strategies<sup>3</sup>. Small molecules have indeed immense therapeutic potential to modulate miRNA expression, because they are inclined to possess ideal drug properties, including good solubility, bioavailability and metabolism<sup>4</sup>. Gumireddy *et al.* applied cell-based assays to report small molecule diazobenzene as modifiers for miR-21, suggesting that miR-21 may become a druggable target<sup>5</sup>. Melo *et al.* provided evidence that the small molecule enoxacin modulated miRNA processing by enhancing TRBP (TAR RNA binding protein) and affected miRNA expression<sup>6</sup>.

Many types of public database are available for providing valuable information about miRNAs. HMDD and miR2Disease have shown their great help in functional analysis of miRNAs in various diseases<sup>7,8</sup>. miREnvironment mainly fills the gap among miRNA, environmental factors (such as exercise, alcohol, radiation and nutrition) and phenotypes (such as cytotoxicity, fat loss and diseases)<sup>9</sup>. SM2miR is the first manually curated database to comprehensively collect the verified miRNAs response to small molecules<sup>10</sup>. Here, we attempt to develop Psmir database to provide the potential small molecule-miRNA interactions which are inferred through computational algorithm in a large-scale based on similarity of genome-wide expression profiles. Among the predictions, the negative associations indicate that the small molecules may function by inhibiting corresponding miRNAs.

In our previous studies, we have proposed computational approaches to identify the potential connections between small molecules and miRNAs based on gene expression similarity in Alzheimer's diseases, as well as cancers<sup>11,12</sup>. Since gene expression profiles following miRNA perturbation were rare in the past, we simulated the presence and absence in miRNA perturbation using up-regulated and down-regulated miRNA in disease condition. Moreover, we built up-regulated/down-regulated miRNA-specific signature by intersecting differentially expressed genes and target genes. Along with abundance of miRNA transfection experiments, genome-wide data

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, P. R. China. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to W.J. (email: [jiangwei@hrbmu.edu.cn](mailto:jiangwei@hrbmu.edu.cn))



**Figure 1.** The snapshot of the interface of Psmir database.

regarding the effects of miRNA on gene expression are available, which facilitates obtaining miRNA-perturbed gene expression profiles. It's noted that the miRNA transfection experiments refer to promoting or inhibiting one miRNA expression using specific miRNA mimics or inhibitors. To build Psmir, we made efforts to comprehensively collect 39 miRNA affected gene expression profiles from Gene Expression Omnibus (GEO). We incorporated the intrinsic miRNA-specific signatures that were extracted from these miRNA-perturbed gene expression profiles into our computational pipeline. The gene expression profiles measured from the cell lines transfection of miRNA mimics or inhibitors will reflect the miRNA regulation effect more directly, which to some extent promotes the reliability of our prediction.

## Results

We searched the corresponding datasets with the keywords “miRNA transfection” or “microRNA transfection” in GEO database. As a result, 124 miRNA-perturbed gene expression profiles comprising both miRNA-transfected group and control group were manually collected as of Sep. 20th in 2013. At the same time, we obtained 6100 small molecule-transfected gene expression profiles from the Connectivity Map (cmap, build 02, <https://www.broadinstitute.org/cmap/>)<sup>13</sup>. Here, we only chose the datasets measured by the same platform (HG-U133A) with cmap for further analysis. Finally, we collected 39 miRNA-perturbed and 6100 small molecule-perturbed gene expression profiles involving 25 miRNAs and 1309 small molecules to perform expression profile similarity analysis (details in Methods). All candidate 51051 associations between 1309 small molecules and 39 miRNAs across 25 miRNAs transfecting in different experimental conditions were evaluated. At the significant level of 0.05 and 0.01, the number of significant associations was 6501 and 1937, respectively. The details of associations were provided in our online database known as Psmir.

Psmir is an accessible web interface for users to browse, search and download small molecule and miRNA associations via a user-friendly graphical user interface at the web address <http://www.bio-bigdata.com/Psmir/>. A snapshot of the database interface is provided in Fig. 1. In the Search section, by selecting miRNA or small molecule in corresponding dropdown menu, users will retrieve the entries that comprise of small molecule name, miRNA name, condition of the miRNA transfection experiment, AS (Association Score) and significance *p*-value. For small molecules, we provide an option that facilitates users focusing on FDA-approved drugs. We further annotated the predicted associations in details according to the known small molecule or miRNA information. The details contain the miRBase accession number of miRNA, the small molecule chemical structure, isomeric SMILES string, ATC code, various IDs (Drugbank ID, Pubchem ID, PharmGKB ID) and whether the small molecule is FDA approved drug or not. The significance of the small molecule and miRNA association was obtained from permutation test.

Psmir is developed in JSP. The Psmir website is deployed in tomcat 6.0.33 web server and runs under Cent OS 5.5 system, supported by a MySQL database to manage data.

**Psmir: 649 Records**

vigabatrin	FDA approved	<a href="#">miR-155</a>	<a href="#">GSE22002</a>	32h	0.576	0.049	<a href="#">more...</a>
vinburnine	FDA unapproved	<a href="#">miR-155</a>	<a href="#">GSE22002</a>	12h	0.7	0.006	<a href="#">more...</a>
vinpocetine	FDA unapproved	<a href="#">miR-155</a>	<a href="#">GSE13296</a>		0.553	0.04	<a href="#">more...</a>
vitexin	FDA unapproved	<a href="#">miR-155</a>	<a href="#">GSE22002</a>	32h	0.529	0.033	<a href="#">more...</a>
vitexin	FDA unapproved	<a href="#">miR-155</a>	<a href="#">GSE9264</a>		-0.569	0.038	<a href="#">more...</a>
vorinostat	FDA approved	<a href="#">miR-155</a>	<a href="#">GSE13296</a>		-0.613	0.001	<a href="#">more...</a>
withaferin A	FDA unapproved	<a href="#">miR-155</a>	<a href="#">GSE13296</a>		-0.657	0.048	<a href="#">more...</a>
xylometazoline	FDA approved	<a href="#">miR-155</a>	<a href="#">GSE22002</a>	12h	-0.691	0.004	<a href="#">more...</a>
xylometazoline	FDA approved	<a href="#">miR-155</a>	<a href="#">GSE22002</a>	32h	0.788	0.007	<a href="#">more...</a>
yohimbic acid	FDA unapproved	<a href="#">miR-155</a>	<a href="#">GSE9264</a>		-0.709	0.042	<a href="#">more...</a>
zalcitabine	FDA approved	<a href="#">miR-155</a>	<a href="#">GSE22002</a>	12h	-0.583	0.039	<a href="#">more...</a>
zidovudine	FDA approved	<a href="#">miR-155</a>	<a href="#">GSE22002</a>	32h	-0.633	0.007	<a href="#">more...</a>
zimeldine	FDA unapproved	<a href="#">miR-155</a>	<a href="#">GSE22002</a>	12h	0.658	0.005	<a href="#">more...</a>
zomepirac	FDA unapproved	<a href="#">miR-155</a>	<a href="#">GSE13296</a>		0.613	0.028	<a href="#">more...</a>
zoxazolamine	FDA unapproved	<a href="#">miR-155</a>	<a href="#">GSE9264</a>		-0.622	0.032	<a href="#">more...</a>

**Figure 2. The query results of miR-155.** At the significant level of 0.05, we retrieved the 649 records corresponding to miR-155. A record included small molecule name, whether the small molecule is FDA approved drugs or not, miRNA name, the GEOseries ID of the miRNA perturbed expression profiles, the condition of the miRNA transfection experiment, the scores of the associations, the significant *p*-values of the associations and the detail information for the associations.

### Case Study

Taking miR-155 as an example, we selected miR-155 in pull down menu, the result corresponding to miR-155 was illustrated in (Fig. 2). MiR-155 is an oncogenic miRNA which regulates several cancer-related pathways and is most significantly up-regulated miRNA in several cancers, such as breast cancer<sup>14</sup>, colon cancer<sup>15</sup> and lung cancer<sup>16</sup>. According to the results of Psmir database, we found that the association of miR-155 and vorinostat presented the negative correlation pattern, which meant that these two molecules regulated the opposite pattern on gene expression at the whole genome level. Vorinostat is a histone deacetylase inhibitor which is a drug under investigation for treatment of cutaneous T-cell lymphoma, and also many research reported its anticancer activity over variety of cancers<sup>17,18</sup>. In lung cancer, miR-155 is frequently up-regulated<sup>19</sup>, as well as able to predict recurrence and survival<sup>16</sup>. Lee *et al.* indicated that the expression level of miR-155 is obviously decreased following vorinostat treatment<sup>20</sup>. What's more, the target genes of miR-155 related to angiogenesis, such as AGTR1 and MYD88<sup>21,22</sup>, and cell proliferation and differentiation, such as BCL2, FGF7<sup>23,24</sup>. These biological processes are closely related to lung cancer. Vorinostat may regulate these target genes of the miRNA and reverse the expression of the miRNA from up-regulation to normal level in lung cancer. Furthermore, the association of miR-155 and vorinostat was also verified in the SM2miR database which is a manually curated database on small molecules' effects on miRNA expression<sup>10</sup>. Predicting associations between small molecules and miRNAs has potential to provide a valuable biological insight for cancer therapies.

### Discussion

The major goal of the present work was to provide potential drug-miRNA associations, which may be a valuable resource for drug-target studies. In this study, we utilized a computational method to infer small molecule-miRNA associations. Firstly, we collected the gene expression profiles perturbed by miRNAs by keywords "miRNA transfection" or "microRNA transfection" from the GEO database. Meanwhile, we obtained the small molecule-perturbed gene expression profiles from cmap. After screening the probes perturbed by one small molecule or one miRNA, we calculated the degree of the probes perturbed by one small molecule (or one miRNA) significantly distributing at the top/bottom of the entire miRNA-perturbed (or small molecule perturbed) gene expression profile. The size of the score determined the degree of similarity between the small molecule and the miRNA. Next, considering multiple gene expression profiles perturbed by the same small molecule, we combined the scores between the same small molecule and the miRNA. What's more, we developed a database known as Psmir, which presented all the results. Researchers who are interested in dissecting the potential relationship of pair-wise small molecules and miRNAs could obtain inference from Psmir which comprises statistically significant associations based on high throughput and public data. In order to expand the application, we annotated each entry according to several of the most popular databases such as Drugbank, miRBase and so on. The users can retrieve all records by selecting one miRNA or small molecule in pull down menu. Besides that, users also can obtain more detailed information about miRNA (such as miRBase accession number of miRNA) and small molecule (such as ATC code of drug) in 'Detail' item.

As far as we know, this is the first database to predict potential associations between small molecules and miRNAs. However, most of human miRNAs have not been designed for transfection experiments, our predicted small molecule-miRNA associations are limited by the number of the available miRNA-perturbed gene expression profiles. With the increasing of miRNA-perturbed or small molecule-perturbed gene expression profiles, we will constantly update Psmir through implementing prediction pipeline for more miRNAs or small molecules in the future.

We believe that Psmir is a valuable complement to the existing curated resources about small molecules and miRNAs, providing a reference to investigations on small molecule and miRNA associations. These associations will facilitate developing novel potential therapeutic targets for human diseases, and may be used to identify drug candidates and predict new indications for existing drugs.

## Methods

**MiRNA-perturbed gene expression profiles.** We searched the GEO database by keywords “miRNA transfection” or “microRNA transfection” and obtained the corresponding GEO series (GSEs) IDs and summary of the experiments. We only reserved the GSEs that comprised both miRNA-transfected group and corresponding control group. Finally, 124 miRNA-perturbed gene expression profiles were manually collected as of Sep. 20th in 2013. Individual gene expression profile is measured from specific miRNA mimics/inhibitors transfection into cultured cell lines. In line with the small molecule-perturbed gene expression profiles obtained from the cmap database, we applied 39 miRNA-perturbed gene expression profiles corresponding to Affymetrix HG-U133-generation microarrays, involving 25 different miRNAs and more than 10 cancer cell lines (Table S1). For simplicity, different cell lines, time points and doses for the same miRNA in different experiments are considered as independent gene expression profiles in the subsequent analysis.

For the miRNA-perturbed gene expression profiles which comprise only one sample in any one class (‘transfection’ or ‘non-transfection’, ‘non-transfection’ also known as ‘control’), we applied ‘fold change’ to measure the extent of differential expression of each probe (fold change  $\geq 2$ ). The expression profiles which consist of greater than or equal to two samples in each class are conducted differential expression analysis using significance analysis of microarray (SAM, FDR  $\leq 0.05$ ). We built a response signature that contains differentially expressed probes and a ranked response pattern which is a full list of probes ranked from the most up-regulated to the most down-regulated by fold change or SAM statistics.

**Small molecule-perturbed gene expression profiles.** cmap (build 02) contains 6100 genome-wide gene expression profiles that comprise the transcriptional responses to 1309 small molecules in cultured human cell lines. To compare the treatment group and control group, all the gene expression measurements were converted to amplitudes. Amplitude ( $a$ ) is used to measure the extent of the differential expression of a given probe.

$$a = \frac{(t - c)}{\frac{1}{2}(t + c)} \quad (1)$$

Where  $t$  is the expression value for the treatment and  $c$  is the expression value for the control. If  $a > 2/3$ , the expression is increased on treatment, i.e.  $t > 2c$ ; if  $a < -2/3$ , the expression is decreased on treatment, i.e.  $t < 0.5c$ .

Analogous to miRNA-perturbed gene expression profile, we also assigned a response signature and a ranked response pattern to each small molecule. It's worth noting that there are different gene expression profiles for the same small molecule due to various experiment conditions (such as cell lines, concentrations and batches). We considered several gene expression profiles for the same small molecule as independent datasets (also referred as “instances”), producing individual response signature and ranked response pattern.

In total, the small molecule-related response signature, as well as small molecule-related ranked response pattern is created for each small molecule-perturbed gene expression profile. The miRNA-related response signature and the miRNA-related ranked response pattern are constructed for each miRNA-perturbed gene expression profile.

**The Enrichment Score (ES) of small molecule/miRNA-related response signature relative to miRNA/small molecule-related ranked response pattern.** The Enrichment Score (ES) is the maximum deviation from zero of  $P_{hit} - P_{miss}$ .

$$P_{hit}(S, i) = \sum_{\substack{g_j \in S \\ j \leq i}} \frac{|r_j|^p}{N_R}, \quad N_R = \sum_{g_j \in S} |r_j|^p \quad (2)$$

$$P_{miss}(S, i) = \sum_{\substack{g_j \notin S \\ j \leq i}} \frac{1}{(N - N_H)} \quad (3)$$

Here,  $S$  is the aforementioned response signature, which is the probes related to the phenotypic distinction (small molecule perturbation or miRNA perturbation), and represents the transcriptional responses to small molecule (or miRNA).  $N_H$  is the number of genes in the gene set.  $L$  is the ranked response pattern, which signifies the genome-wide transcriptional profile.  $L$  comprises  $N$  probe members. For the  $i_{th}$  member in  $L$ , we evaluate the fraction of probes in  $S$  (“hit”) weighted by their correlation and the fraction of probes not in  $S$  (“miss”) which are both present up to the given position  $i$  in  $L$ . We set parameter  $p = 1$ , weighting the probes in  $S$  by their correlation

with perturbation (small molecule or miRNA) normalized by the sum of the correlations over all of the probes in  $S^{25}$ .

For both small molecule and miRNA, we divided the response signature into up-regulated probe sets and down-regulated probe sets to calculate  $ES$ , respectively. The up-regulated or down-regulated probes of the miRNA-related response signature were searched within the ranked response pattern of small molecule, resulting in two enrichment scores ( $ES_{mir2sm}^{up}$ ,  $ES_{mir2sm}^{down}$ ), respectively. Analogously, up-regulated and down-regulated probes of small molecule-related response signature were also searched within the ranked response pattern of miRNA, producing two enrichment scores ( $ES_{sm2mir}^{up}$ ,  $ES_{sm2mir}^{down}$ ), respectively. We implemented the computational procedure of enrichment score with R program.

**The similarity of small molecule-perturbed and miRNA-perturbed gene expression profiles (total enrichment score).** For a small molecule  $i$  and miRNA  $j$ , we firstly obtained  $ES_{sm2mir}^{up}$ ,  $ES_{sm2mir}^{down}$ ,  $ES_{mir2sm}^{up}$ ,  $ES_{mir2sm}^{down}$  as above stated. Then, we combined  $ES_{sm2mir}^{up}$  and  $ES_{sm2mir}^{down}$  to Total Enrichment Score ( $TES_{sm2mir}$ ), and combined  $ES_{mir2sm}^{up}$ ,  $ES_{mir2sm}^{down}$  to Total Enrichment Score ( $TES_{mir2sm}$ ). Lastly, we used the mean of  $TES_{sm2mir}$  and  $TES_{mir2sm}$  to measure the  $TES_{(i,j)}$  of the pair of small molecule  $i$  and miRNA  $j$ .

$$TES_{sm2mir} = \frac{ES_j^u - ES_j^d}{2} \quad (4)$$

$$TES_{mir2sm} = \frac{ES_i^u - ES_i^d}{2} \quad (5)$$

$$TES_{(i,j)} = \frac{TES_{sm2mir} + TES_{mir2sm}}{2} \quad (6)$$

Here,  $ES_x^y$  (with  $x \in \{i,j\}$ ,  $y \in \{u,d\}$  for small molecule and  $y \in \{u_m, d_m\}$  for miRNA) ranges in  $[-1, 1]$ .  $ES_x^y$  is the enrichment score of the  $y$  (up or down) response signature with respect to the ranked response pattern of  $x$ . It quantifies how much a set of genes (or probes) is at the top of a ranked full list. The closer this measure is to 1, the closer the genes are to the top of the ranked full list. The closer the value to  $-1$ , the closer the genes are to the bottom of the ranked full list.

$TES_{(i,j)}$  also ranges in  $[-1, 1]$ . It comprehensively quantifies how much the genes (or probes) in the up-regulated or down-regulated gene sets of ( $u$  or  $d$ ) relative to small molecule  $i$  are placed at the top or bottom of the miRNA  $j$  and how much the genes (or probes) in the up-regulated or down-regulated gene sets ( $u$  or  $d$ ) relative to miRNA  $j$  are placed at the top or bottom of the small molecule  $i$ . The closer these two statements are to the truth, the farther to 0 is the value of  $TES_{(i,j)}$ , that is to say the larger is the absolute value of  $TES_{(i,j)}$ .

**Pairwise association score (AS) of small molecules and miRNAs.** Since multiple gene expression profiles are relative to one small molecule (aforementioned ‘instances’), the final Association Score (AS) for a pair of small molecule and miRNA is a composite score by combining TES of all the instances corresponding to the same small molecule with one miRNA via Kolmogorov-Smirnov (KS) test:

$$a = \max_{1 \leq j \leq t} \left[ \frac{j}{t} - \frac{v(j)}{n} \right] \quad (7)$$

$$b = \max_{1 \leq j \leq t} \left[ \frac{v(j)}{n} - \frac{(j-1)}{t} \right] \quad (8)$$

$$\text{Association Score (AS)} = \begin{cases} a, & (a > b) \\ -b, & (b > a) \end{cases} \quad (9)$$

Here,  $t$  is the number of instances corresponding to one small molecule.  $n$  is the number of all the instances for all small molecules ( $n = 6100$ ).  $V(j)$  is the  $j$ th element of the vector  $V$ , and denotes the position of the  $j$ th instance which are in the set of  $t$  instances in the ordered list of all  $n$  instances ( $j = 1, 2, \dots, t$ ). All the  $n$  instances are ranked in decreasing order according to the magnitude of TES.

Additionally, we performed a permutation test for each pair of small molecule and miRNA. We calculated 1000 random ASs generated by randomly sampling probes in ranked response pattern to form response signature while keeping the number of the response signature per small molecule or miRNA unchanged.  $p$ -value is the fraction of the absolute value of the random AS which is larger than that of the observed AS.

## References

1. Medina, P. P., Nolde, M. & Slack, F. J. OncomiR addiction in an *in vivo* model of microRNA-21-induced pre-B-cell lymphoma. *Nature* **467**, 86–90 (2010).
2. van Rooij, E. & Kauppinen, S. Development of microRNA therapeutics is coming of age. *EMBO Mol Med* **6**, 851–864 (2014).
3. Castanotto, D. & Rossi, J. J. The promises and pitfalls of RNA-interference-based therapeutics. *Nature* **457**, 426–433 (2009).
4. Zhang, S., Chen, L., Jung, E. J. & Calin, G. A. Targeting microRNAs with small molecules: from dream to reality. *Clin Pharmacol Ther* **87**, 754–758 (2010).
5. Gumireddy, K. *et al.* Small-molecule inhibitors of microRNA miR-21 function. *Angew Chem Int Ed Engl* **47**, 7482–7484 (2008).

6. Melo, S. *et al.* Small molecule enoxacin is a cancer-specific growth inhibitor that acts by enhancing TAR RNA-binding protein 2-mediated microRNA processing. *Proc Natl Acad Sci USA* **108**, 4394–4399 (2011).
7. Jiang, Q. *et al.* miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res* **37**, D98–104 (2009).
8. Li, Y. *et al.* HMDD v2.0: a database for experimentally supported human microRNA and disease associations. *Nucleic Acids Res* **42**, D1070–1074 (2014).
9. Yang, Q., Qiu, C., Yang, J., Wu, Q. & Cui, Q. miREnvironment database: providing a bridge for microRNAs, environmental factors and phenotypes. *Bioinformatics* **27**, 3329–3330 (2011).
10. Liu, X. *et al.* SM2miR: a database of the experimentally validated small molecules' effects on microRNA expression. *Bioinformatics* **29**, 409–411 (2013).
11. Jiang, W. *et al.* Identification of links between small molecules and miRNAs in human cancers based on transcriptional responses. *Sci Rep* **2**, 282 (2012).
12. Meng, F. *et al.* Constructing and characterizing a bioactive small molecule and microRNA association network for Alzheimer's disease. *J R Soc Interface* **11**, 20131057 (2014).
13. Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
14. Mattiske, S., Suetani, R. J., Neilsen, P. M. & Callen, D. F. The oncogenic role of miR-155 in breast cancer. *Cancer Epidemiol Biomarkers Prev* **21**, 1236–1243 (2012).
15. Hongliang, C., Shaohun, H., Aihua, L. & Hua, J. Correlation between expression of miR-155 in colon cancer and serum carcinoembryonic antigen level and its contribution to recurrence and metastasis forecast. *Saudi Med J* **35**, 547–553 (2014).
16. Yanaihara, N. *et al.* Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* **9**, 189–198 (2006).
17. Kelly, W. K. *et al.* Phase I study of an oral histone deacetylase inhibitor, suberoylanilide hydroxamic acid, in patients with advanced cancer. *J Clin Oncol* **23**, 3923–3931 (2005).
18. Olsen, E. A. *et al.* Phase IIB multicenter trial of vorinostat in patients with persistent, progressive, or treatment refractory cutaneous T-cell lymphoma. *J Clin Oncol* **25**, 3109–3115 (2007).
19. Volinia, S. *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci USA* **103**, 2257–2261 (2006).
20. Lee, E. M. *et al.* Suberoylanilide hydroxamic acid (SAHA) changes microRNA expression profiles in A549 human non-small cell lung cancer cells. *Int J Mol Med* **24**, 45–50 (2009).
21. Jagavelu, K. *et al.* Endothelial cell toll-like receptor 4 regulates fibrosis-associated angiogenesis in the liver. *Hepatology* **52**, 590–601 (2010).
22. Herr, D. *et al.* Regulation of endothelial proliferation by the renin-angiotensin system in human umbilical vein endothelial cells. *Reproduction* **136**, 125–130 (2008).
23. Craig, R. W. MCL1 provides a window on the role of the BCL2 family in cell proliferation, differentiation and tumorigenesis. *Leukemia* **16**, 444–454 (2002).
24. Jeon, Y. M. *et al.* Fibroblast growth factor-7 facilitates osteogenic differentiation of embryonic stem cells through the activation of ERK/Runx2 signaling. *Mol Cell Biochem* **382**, 37–45 (2013).
25. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**, 15545–15550 (2005).

## Acknowledgements

This work was supported by the Foundation for The National Natural Science Foundation of China [31401134, 31501078, 61571169] and the Natural Science Foundation of Heilongjiang Province of China [QC2014C017], and the University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province [UNPYSCT-2015037].

## Author Contributions

W.J. and F.M. conceived and designed the study. J.W. designed the database framework and built the website. F.Y., X.C., D.L., S.W., E.D. and X.Y. collected, organized and compiled the data and revised the manuscript. All of the authors read and approved the final manuscript.

## Additional Information

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

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

**How to cite this article:** Meng, F. *et al.* Psmir: a database of potential associations between small molecules and miRNAs. *Sci. Rep.* **6**, 19264; doi: 10.1038/srep19264 (2016).



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