Towards the understanding of microRNA and environmental factor interactions and their relationships to human diseases

Chengxiang Qiu1, Geng Chen1 & Qinghua Cui1,2,3

1Department of Biomedical Informatics, School of Basic Medical Sciences, Peking University, 38 Xueyuan Rd, Beijing, 100191, China, 2MOE Key Lab of Molecular Cardiovascular Science, Peking University, 38 Xueyuan Rd, Beijing, 100191, China, 3Institute of Systems Biomedicine, Peking University Health Science Center, 38 Xueyuan Rd, Beijing, 100191, China.

Increasing studies have shown that the interactions between microRNAs (miRNAs) and environmental factors (EFs) play critical roles in determining phenotypes and diseases. In this study, we revealed a number of important biological insights by analyzing and modeling of miRNA-EF interactions and their relationships with human diseases. We demonstrated that the miRNA signatures of EFs could provide new information on EFs. More importantly, we quantitatively showed that the miRNA signatures of drug/radiation could be used as indicators for evaluating the results of cancer treatments. Finally, we developed a computational model that could efficiently identify the possible relationship between EF and human diseases. Meanwhile, we provided a website (http://cmbi.hsc.pku.edu.cn/miren) for the main results of this study. This study elucidates the mechanisms of EFs, presents a framework for predicting the results of cancer treatments, and develops a model that illustrates the relationships between EFs and human diseases.

The phenotypes of an organism are determined by the complex interactions between genetic factors (GFs) and environmental factors (EFs). The interactions between GFs and EFs link them together to form complex networks and thus work together on the network level to influence phenotypes and diseases, especially complex ones such as cancer and cardiovascular diseases. Recently, the theories and methods of network medicine have shed light on the network-level study of GF–EF interactions1. The computational analysis and modeling of the GF–EF interactions greatly improved the understanding of the mechanisms of EFs and explored new GF–EF interactions. For example, the analysis and modeling of one class of important environmental factors, drugs, and their interactions with their targets (GFs) has revealed a number of important insights2 and identified new drug-target interactions3. Moreover, the analysis of the interactions between drugs and genetic factors has successfully identified new indications for approved drugs that not only generate revenues for pharmaceutical companies but also benefit patients, which is the real incentive4.

microRNAs (miRNAs) are a class of newly identified GFs, which mainly repress the expression of target mRNAs at the post-transcriptional level5. Studies have reported that miRNAs are critical in various important biological processes, such as cell growth, proliferation, differentiation, development, and apoptosis6. In terms of their importance, a miRNA dysfunction is thus associated with a broad range of diseases7. More recently, increasing studies have shown that miRNAs can functionally interact with a wide spectrum of EFs, such as drugs8, virus8, alcohol10, cigarette11, stress12, diet13, and radiation14. Moreover, miRNA–EF interactions critically affect and determine phenotypes. Hence, the dysfunction of miRNA–EF interactions is associated with abnormal phenotypes and even diseases15. Given their biological importance, the computational analysis and modeling of miRNA–EF interactions is becoming increasingly essential in elucidating the mechanisms of EFs, identifying the miRNA signatures of EFs, and exploring new indications of approved drugs. However, such an analysis and modeling has not been widely performed because of the lack of a large-scale miRNA-EF interaction dataset.

In a recent study, we gathered the experimentally supported miRNA–EF interactions and released the miREnvironment database, which contains more than 2500 entries, including ~800 miRNAs, ~260 environmental factors, ~180 phenotypes, and 17 species. To obtain a comprehensive understanding of miRNA–EF interactions and their relationships to human diseases, we analyzed and modeled human miRNA–EF interactions based on the miREnvironment database. As a result, a number of important biological patterns were revealed. Moreover, we presented quantitative evidence on the potential of miRNA signatures as biomarkers of the results of cancer treatment. Finally, we developed a model to predict the potential relationship between EFs and human
Results

miRNA–EF interaction is correlated with miRNA characteristics. It is interesting to investigate whether miRNA characteristics have roles in miRNA–EF interactions. Hence, we analyzed the correlation between the number of miRNA-interacting EFs with the expression level, tissue specificity, conservation, and disease spectrum width (DSW) of the miRNAs. The miRNA–EF interactions showed significantly regular patterns. The number of EFs interacting with a miRNA has a significant positive correlation with the expression level of the miRNA ($R=0.50$, $P=1.784\times10^{-14}$, Spearman’s correlation, Figure 1A). This result suggests that highly expressed miRNAs tend to interact with more EFs. Moreover, the number of EFs interacting with a miRNA is negatively correlated with the tissue specificity index of the miRNA ($R=-0.32$, $P=1.850\times10^{-6}$, Spearman’s correlation, Figure 1B). This finding indicates that miRNAs expressed in a broader tissue profile normally operate under more diverse cellular conditions and thus interact with more EF. The number of EFs interacting with a miRNA was also found to be significantly correlated with the miRNA conservation ($R=-0.47$, $P=3.117\times10^{-14}$, Spearman’s correlation, Figure 1C). More conserved miRNAs tend to have greater possibilities of interaction with EFs than less conserved miRNAs (Figure S1). For example, more than 60% of the miRNAs conserved in vertebrates or invertebrates interact with EFs, whereas only 4.8% of the human-specific miRNAs interact with EFs. In this study, we defined and calculated the DSW for each miRNA (see Methods). Interestingly, the DSW of a miRNA is significantly correlated with the number of interacting EFs ($R=0.60$, $P<2.2\times10^{-16}$, Spearman’s correlation, Figure 1D), suggesting that the miRNAs associated with a wider range of diseases tend to interact with more EFs. For example, miR-21, which has the widest DSW (0.31), has the most number (34) of interacting EFs (Figure 1D).

An EF–EF interaction network linked by miRNAs. Next, we evaluated the relationship between any two EFs through their miRNA signatures. Two EFs are connected if they share significantly enriched interacting miRNAs. Using this procedure, we constructed an EF–EF interaction network (Figure 2). As shown in Figure 2, the EF–EF interaction network shows a modularity structure. Related EFs tend to cluster together in the network. For example, insulin, exercise, and anabolic stimulus are linked with each other and thus grouped together. Indeed, all of the three EFs are related to metabolism. Meanwhile, HCV (hepatitis C virus) and alcohol are related. Among the four miRNAs (miR-122, miR-126, miR-212, and miR-375) that interact with alcohol and the 11 miRNAs (miR-122, miR-130a, miR-130a, miR-181b, miR-192, miR-196a, miR-196b, miR-199a, miR-21, miR-375, and miR-99b)
that interact with HCV, two are overlapped (miR-122 and miR-135). The miR-122 is liver-specific and known to be critical to liver function, whereas miR-373 is highly expressed in pancreas and has downstream effects on the liver\(^{1,2}\). Both EFs have effects on the liver. These results suggest that the miRNA-level signatures of EFs can be used as metrics for evaluating the relationships among EFs.

The cancer treatment drugs are mainly grouped in two clusters, as shown by the circled regions in Figure 2. Adriamycin is the hub of the cluster 1 (denoted by the red circle), suggesting that adriamycin may be functionally related to the other drugs. Indeed, adriamycin is a general drug for cancer treatment. Meanwhile, the drugs in the green circle (cluster 2) are connected by hydrogen peroxide, which is the hub of the cluster. The American Cancer Society has stated that "there is no scientific evidence that hydrogen peroxide is a safe, effective or useful cancer treatment\(^{3}\). However, studies have shown that hydrogen peroxide is an important signaling molecule in the regulation of various biological processes\(^{4}\) and a potential drug for cancer chemoprevention and therapy\(^{5}\). In this study, we present quantitative miRNA-level evidence that hydrogen peroxide is functionally related to a number of anti-cancer drugs, indicating that it is a potential cancer treatment. Moreover, function enrichment analysis shows that the miRNAs in the two clusters have common and cluster-specific functions. Both clusters are rich in cancer-related functions (i.e. tumor suppressor, onco-miRNAs, cell cycle, apoptosis, immune response, and epithelial-mesenchymal transition). Cluster 1 has specific functions such as Akt pathway and carbohydrate metabolism; whereas cluster 2 has specific functions such as bone regeneration, cardiogenesis, cell fate determination, DNA repair, folliculogenesis, and granulopoiesis (Figure 2).

**miRNA signatures of cancer treatment drug/radiation can be indicators for therapeutic result evaluation**. Cancer is one of the most common fatal diseases worldwide. Therefore, a molecular signature that can evaluate and predict the therapeutic results of a cancer treatment drug or radiation is critical for patient survival. Such a signature is also helpful in finding better therapeutic strategies for cancer. At the molecular level, the formation and development of cancer is normally accompanied by the gain- and loss-of-function changes of oncogenes and tumor suppressors, respectively. As a class of newly identified genes, some miRNAs also show strong oncogenic or tumor-suppressing abilities\(^{6}\). We previously arranged a comprehensive collection of human miRNA oncogenes and miRNA tumor suppressors and showed that the two classes of miRNAs have various biological differences, including functions and targets\(^{7}\). Given that cancer samples normally show gain- and loss-of-function changes in miRNA oncogenes and tumor suppressors, respectively\(^{8}\), a successful (drug/radiation sensitive) cancer treatment should inhibit the activity of miRNA oncogenes while amplifying the activity of miRNA tumor suppressors. In contrast, a failed (drug/radiation resistant) treatment can result from failure in the inhibition and amplification of miRNA oncogenes and tumor suppressors, respectively. Thus, to evaluate and predict the results of cancer treatment, we analyzed the enrichment of the differentially expressed miRNAs after cancer treatment in the two groups of miRNAs (miRNA oncogenes and

---

**Figure 2** | EF–EF interaction network and the enriched functions of the miRNAs that interact with the two main clusters of cancer drugs. The blue box represents the common functions of miRNAs that interact with two clusters of drugs. The red box represents the functions of miRNAs that only interact with drugs of cluster 1. The green box represents the functions of miRNAs that only interact with drugs of cluster 2.
tumor suppressors) (see Methods, Figure 3A). We performed three independent studies with the following subjects: (1) resistant leukemia treated by adriamycin; (2) sensitive lung cancer treated by irradiation, and (3) resistant pancreatic cancer treated with gemcitabine (Table 1). Results showed that although the differentially expressed miRNAs from the three studies were quite different at the single-miRNA level, all miRNA signatures showed significant and regular patterns in terms of the two miRNA sets, namely, miRNA oncogenes and tumor suppressors (Table 1, Figure 3B). For the treatment-resistant cancer, the miRNA oncogenes were significantly up-regulated, whereas the miRNA tumor suppressors were significantly down-regulated. On the other hand, for the treatment-sensitive cancer, the miRNA tumor suppressors were significantly up-regulated, whereas the miRNA oncogenes were significantly down-regulated. These results indicate that a successful cancer treatment inhibits the activity of miRNA while amplifying the activity of miRNA tumor suppressors. The resistance of patients to treatments resulted from the unsuccessful inhibition and amplification of miRNA oncogene and tumor suppressor functions, respectively. This observation also indicates that the enrichment analysis of the miRNA signatures of the cancer treatment drugs or radiation in the two pre-defined miRNA sets efficiently evaluated and predicted the results of cancer treatment. This finding further implies that a method for improving the results of cancer treatment is to look for strategies that inhibit and amplify the activities of miRNA oncogenes and tumor suppressors, respectively.

miRNA signatures can connect EFs with human diseases. The signatures of the protein-coding genes of a drug and a disease can connect the drug with the disease if they are associated with each other. In this study, we hypothesize that the miRNA signatures of EFs and diseases should be intrinsically connected if the EFs and diseases are associated by nature. The miRNA–EF or miRNA–drug interactions are a new dimension of information regarding EFs, particularly drugs. Hence, new and valuable relationships between EFs and human diseases can be found based on their miRNA signatures. In this study, we developed a model to predict potential associations between EFs and human diseases (Figure 4) by evaluating the significance of the similarity of two sets of independent miRNA signatures (see Methods). As a result, we identified a number of potential associations between EFs and human diseases. To evaluate the accuracy and reliability of the prediction, we validated the top 30 predictions with previous literature. The results showed that 80.0% (24) of the top 30
Figure 4 | Workflow for inferring potential associations between EFs and human diseases.
The miRNA tissue specificity index values were obtained from the study of Lu et al.\textsuperscript{7}. Human miRNAs are divided into five categories according to their conservation, namely, a miRNA is conserved only in humans (G5), primates (G4), mammals (G3), vertebrates (G2), or other more distant species (G1, which is the most conserved group).

Human miRNA and disease association data. We downloaded the human miRNA and disease association data from the Human microRNA Disease Database (HMDD, http://cmbi.bjmu.edu.cn/hmdd, September 2011 version), which integrated ~6000 entries, including ~500 miRNA genes and ~300 diseases from ~1500 publications. Since its release on December 2007, HMDD has been updated almost every month and is one of the most comprehensive databases of human miRNA-disease associations. We then extracted the distinct miRNA-disease association data from the entire HMDD dataset. Because some of the HMDD data is also integrated into the miRNA Environment database, the overlapping datasets in the HMDD database were removed to ensure that the miRNA signatures of the human diseases are independent from the miRNA signatures of the EFs. We standardized the names of miRNAs according to those in miRBase.

Disease spectrum width of a miRNA. In this paper, the "disease spectrum width" (DSW) of a given miRNA (\( m_i \)) was defined as

\[
DSW_i = \frac{n_i}{N}
\]

(1)

where \( n_i \) is the number of diseases associated to \( m_i \) and \( N \) is the total number of diseases associated to all miRNAs. DSW is a parameter that evaluates the effect of a miRNA in human diseases.

Enrichment analysis of the miRNA signatures of cancer treatment drug/radiation in miRNA oncogenes and tumor suppressors. We used TAM\textsuperscript{28}, a tool for miRNA set enrichment analysis, to explore the regular patterns of the miRNA signatures of a cancer treatment drug/radiation. The miRNA set was defined as a group of miRNAs with related meanings according to some rules (i.e. function)\textsuperscript{2}. We evaluated the significance of the miRNA signature enrichment of cancer treatments in terms of two miRNA sets highly related with cancer, namely, the miRNA oncogenes and tumor suppressors\textsuperscript{28}. For each miRNA signature of drug/radiation treatment, we first curated the up-regulated and down-regulated miRNAs, as shown in Figure 3A. Then, we evaluated the significance of the up-regulated miRNAs in the miRNA oncogenes and tumor suppressors, respectively. This procedure was repeated for the down-regulated miRNAs. Combining the results of the up-regulated and down-regulated miRNAs, we can now examine the cancer treated by the given drug or radiation is either sensitive or resistant to the treatment.

**EF-disease association identification model.** The theoretical basis for the inference of drug and disease association based on molecular activity is the intrinsic connection between the molecular signatures of the drug and disease if there are indeed associated. This theoretical basis was confirmed by successful studies on drug-repositioning\textsuperscript{19,20}. In this study, we hypothesized that the miRNA signatures of the EFs, including drugs, and human diseases should be intrinsically connected. We obtained the miRNA signatures of human diseases from the HMDD database and the miRNA signatures of environmental factors, as described above. To ensure the independence of the two sets of miRNA signatures, we removed the overlapping data from the HMDD database. Most of the miRNA in both sets of signatures do not have information on differential expressions. Thus, we used Fisher’s exact test to evaluate the significance of the overlap between the miRNA signatures of the given EF and disease.

**Statistical analysis.** Statistical computations and tests were performed using R, a statistical computing language (http://www.r-project.org/), and the correlations were calculated using Spearman’s correlation, a non-parametric method. Moreover, we evaluated the significance of the enrichment between the miRNA signatures of two EFs using Fisher’s exact test. The same test was used to evaluate the significance of the enrichment among the miRNA signatures of EFs and human diseases.

1. Barabasi, A. L., Gulbahce, N. & Loscalzo, J. Network medicine: a network-based approach to human disease. Nat Rev Genet 12 (1), 56 (2011).
2. Yildirim, M. A. et al. Drug-target network. Nat Biotechnol 25 (10), 1119 (2007).
3. Campillos, M. et al. Drug target identification using side-effect similarity. Science 321 (5886), 263 (2008).
4. Dudley, J. T., Deshpande, T. & Butte, A. J. Exploiting drug-disease relationships for computational drug repositioning. Brief Bioinform 12 (4), 303 (2011).
5. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116 (2), 281 (2004).
6. Esquela-Kerscher, A. & Slack, F. J. Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer 6 (4), 259 (2006).
7. Lu, M. et al. An analysis of human microRNA and disease associations. *PLoS ONE* 3(10)e3420. doi:10.1371/journal.pone.0003420 (2008).
8. Lima, R. T. et al. MicroRNA regulation of core apoptosis pathways in cancer. *Eur J Cancer* 47 (10), e3420. doi:10.1371/journal.pone.0003420 (2008).
9. Lin, Z. & Flemington, E. K. miRNAs in the pathogenesis of oncogenic human viruses. *Cancer Lett* 305 (2), 186 (2011).
10. Ladeiro, Y. et al. MicroRNA profiling in hepatocellular tumors is associated with clinical features and oncogene/tumor suppressor gene mutations. *Hepatology* 47 (6), 1955 (2008).
11. Izzotti, A. et al. Modulation of microRNA expression by budesonide, phenethyl isothiocyanate and cigarette smoke in mouse liver and lung. *Carcinogenesis* 31 (5), 894 (2010).
12. Gidron, Y. De Zwaan, M., Quint, K. & Ocker, M. Influence of stress and health-behaviour on miRNA expression. *Mol Med Report* 3 (3), 455 (2010).
13. Alisi, A. et al. Mirnome analysis reveals novel molecular determinants in the pathogenesis of diet-induced nonalcoholic fatty liver disease. *Lab Invest* 91 (2), 283 (2011).
14. Niemoeller, O. M. et al. MicroRNA expression profiles in human cancer cells after ionizing radiation. *Radiat Oncol* 6, 29 (2011).
15. Yang, Q. et al. miREnvironment Database: providing a bridge for microRNAs, environmental factors, and phenotypes. *Bioinformatics* 27(23), 3329 (2011).
16. Poy, M. N. et al. miR-375 maintains normal pancreatic alpha- and beta-cell mass. *Proc Natl Acad Sci U S A* 106 (14), 5813 (2009).
17. Questionable methods of cancer management: hydrogen peroxide and other ‘hyperoxygenation’ therapies. *CA Cancer J Clin* 43 (1), 47 (1993).
18. Veal, E. A., Day, A. M. & Morgan, B. A. Hydrogen peroxide sensing and signaling. *Mol Cell* 26 (1), 1 (2007).
19. Lopez-Lazaro, M. Dual role of hydrogen peroxide in cancer: possible relevance to cancer chemoprevention and therapy. *Cancer Lett* 252 (1), 1 (2007).
20. Wang, D. et al. Human microRNA oncogenes and tumor suppressors show significantly different biological patterns: from functions to targets. *PLoS One* 5 (9) (2010).
21. Cao, Y. X., Dai, C. W. & Zhang, G. S. [Screening for drug resistance related microRNAs in K562 and K562/A02 cell lines.]. *Zhonghua Xue Ye Xue Za Zhi* 31 (6), 361 (2010).
22. Wang, X. C. et al. Expression and function of miRNA in postoperative radiotherapy sensitive and resistant patients of non-small cell lung cancer. *Lung Cancer* 72 (1), 92 (2011).
23. Li, Y. et al. Up-regulation of miR-200 and let-7 by natural agents leads to the reversal of epithelial-to-mesenchymal transition in gemcitabine-resistant pancreatic cancer cells. *Cancer Res* 69 (16), 6704 (2009).
24. Lamb, J. et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313 (5795), 1929 (2006).
25. Sirota, M. et al. Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci Transl Med* 3 (96), 96ra77 (2011).
26. Smilke, M. F. et al. Autoantibodies, human T lymphotropic virus type 1 and type 1 diabetes mellitus in Jamaicans. *West Indian Med J* 51 (3), 153 (2002).
27. Liang, Y., Ridzon, D., Wong, L. & Chen, C. Characterization of microRNA expression profiles in normal human tissues. *BMC Genomics* 8, 166 (2007).
28. Lu, M. et al. TAN: a method for enrichment and depletion analysis of a microRNA category in a list of microRNAs. *BMC Bioinformatics* 11, 419 (2010).

**Acknowledgements**

This work was supported by the Natural Science Foundation of China (Grant No. 30900829).

**Author contributions**

QC designed the study and wrote the paper. CQ and GC performed the analysis and modeling. CQ and GC contributed equally to this study.

**Additional information**

**Supplementary information** accompanies this paper at http://www.nature.com/scientificreports

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

**License:** This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/

**How to cite this article:** Qiu, C., Chen, G. & Cui, Q. Towards the understanding of microRNA and environmental factor interactions and their relationships to human diseases. *Sci. Rep.* 2, 318; DOI:10.1038/srep00318 (2012).