Netw Model Anal Health Inform Bioinforma (2015) 4:1
DOI 10.1007/s13721-014-0074-x

ORIGINAL ARTICLE

MicroRNA–mRNA interaction analysis to detect potential dysregulation in complex diseases

Wenlong Tang · Chao Xu · Yu-Ping Wang · Hong-Wen Deng · Ji-Gang Zhang

Received: 11 September 2014 / Revised: 25 November 2014 / Accepted: 5 December 2014 / Published online: 10 January 2015

Abstract It is now recognized that genetic interactions (epistasis) are important sources of the hidden genetic variations and may play an important role in complex diseases. Identifying genetic interactions not only helps to explain part of the heritability of complex diseases, but also provides the clue to understand the underlying pathogenesis of complex diseases. Advances in high-throughput technologies enable simultaneous measurements of multiple genomic features from the same samples on a genome-wide scale, and different omics features are not acting in isolation but interact/crosstalk at multiple (within and across individual omics features) levels in complex networks. Therefore, genetic interaction needs to be accounted for across different omics features, potentially allowing an explanation of phenotype variation that single omics data cannot capture. In this study, we propose an analysis framework to detect the miRNA–mRNA interaction enrichment by incorporating principal components analysis and canonical correlation analysis. We demonstrate the advantages of our method by applying to miRNA and mRNA data on glioblastoma (GBM) generated by The Cancer Genome Atlas project. The results show that there are enrichments of the interactions between co-expressed miRNAs and gene pathways which are associated with GBM status. The biological functions of those identified genes and miRNAs have been confirmed to be associated with glioblastoma by independent studies. The proposed approach provides new insights in the regulatory mechanisms and an example for detecting interactions of multi-omics data on complex diseases.

Keywords Interaction analysis · Multi-omics · MicroRNAs · Glioblastoma

W. Tang · C. Xu · Y.-P. Wang · H.-W. Deng · J.-G. Zhang
Center for Bioinformatics and Genomics, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA 70112, USA

W. Tang · Y.-P. Wang
Department of Biomedical Engineering, Tulane University, New Orleans, LA 70118, USA

W. Tang · C. Xu · H.-W. Deng · J.-G. Zhang (✉)
Department of Biostatistics and Bioinformatics, Tulane University, 1440 Canal Street, Suite 2001, New Orleans, LA 70112, USA
e-mail: jzhang9@tulane.edu

1 Introduction

The phenotypic variations of complex diseases are highly complex due to its polygenic inheritance and environmental influence. With high-throughput technologies, omics data analyses, e.g. genome-wide association studies (GWAS), have been carried out with notable success in identifying genetic variants underlying complex diseases and have tremendously improved our understanding of the genetic mechanisms of many common complex diseases (Burton et al. 2007). A primary interest in omics data analyses was typically the identification of a set of SNPs or genes that are marginally associated with a target disease. An increasing number of researches have highlighted that gene interactions (epistasis) are important sources of the hidden genetic variations for complex diseases, which can be explained by joint effects of multiple SNPs/genes but not by their main effects (Cordell 2009; Kallberg et al. 2007; Nelson et al. 2001; Ritchie et al. 2001; Wan et al. 2010). Identifying genetic interactions not only helps to explain part of the heritability of complex diseases, but also provides the clue to

understand the underlying pathogenesis of complex diseases and improves predictions of individual disease risk in humans (Cordell 2009; Mackay 2014; Phillips 2008; Visweswaran et al. 2009).

So far, most studies of gene interaction analyses focused on single omics data such as GWASs and gene expression data. Many statistical methods have been proposed for detecting gene interactions (Cordell 2009; Mackay 2014), including multifactor dimensionality reduction method (Ritchie et al. 2001), PLINK (Purcell et al. 2007), Random Jungle (Schwarz et al. 2010), BEAM (Zhang and Liu 2007), the combinatorial partitioning method (Nelson et al. 2001), the restricted partition method (Culverhouse et al. 2004), and the combinatorial searching method (Sha et al. 2006). Some of them made an exhaustive search for pairwise gene interactions at the genome-wide level; some made a selection of a subset of SNPs or genes for interaction tests on the basis of existing biological knowledge or statistical features; some adopt Machine-learning and data-mining algorithms for data reduction and/or feature selection to reduce the computational burden in gene interaction analysis (Mackay 2014).

Recently, advances in high-throughput technologies enable simultaneous measurements of multiple genomic features (e.g., RNA transcription data, genotype variation data and proteomic data) from the same samples on a genome-wide scale, and provide us a tremendous amount of information to uncover a number of risk factors for human complex diseases (Guan et al. 2010; Juran and Lazaridis 2011). Importantly, different omics features are not acting in isolation but interact/crosstalk at multiple levels in complex networks. Gene interaction analysis of individual omics studies fall short of providing a comprehensive view of the genetic factors and their functions in the form of complex function/regulatory networks for complex diseases (Farber and Lusis 2009). Therefore, genetic interaction analysis needs to be accounted for across different omics features by integrating multi-omics data, powerfully and comprehensively identifying molecular and genomic factors/mechanisms underlying the pathogenesis of complex diseases (Farber 2010).

In this study, we propose an analysis framework to detect the miRNA–mRNA interaction enrichment by incorporating principal components analysis and canonical correlation analysis, which are associated with the phenotypes under study. It is well known that genes often cooperate with each other to perform various cellular functions and are organized into functional modules with densely connected genes within gene pathways (Newman 2006). Similarly, miRNAs also show cooperative effects on complex diseases (Chhabra et al. 2010; Feederle et al. 2011; Mavrakis et al. 2011). Thus, it is likely that there exists the enrichment of miRNA–mRNA interactions between co-expressed miRNAs and gene pathways. We demonstrate the advantages of our method by applying to miRNA and mRNA data on glioblastoma (GBM) generated by The Cancer Genome Atlas (TCGA) project. The results show that there are enrichments of the interactions between co-expressed miRNAs and gene pathways which are associated with GBM status and our method can provide functional information about individual genes and miRNAs to uncover the dysregulation of functional modules and their mutual interactions. Our study may set an example for detecting interactions of multi-omics data on complex diseases.

2 Methods

In this study, we focus on samples in two classes (e.g., the two subtypes of a disease). Suppose that in class k ($k = 1$ or 2) each sample consists of gene expression profile and miRNA expression profile. MiRNA data and gene expression data consist of X miRNAs and Z mRNAs, respectively. Let Y ($y_i = k$, $i = 1, 2,\ldots, n$) be a vector of the phenotypes for samples, where n is the number of samples.

According to previous studies (Liu et al. 2012a; Zhang et al. 2009), miRNA–mRNA interactions can be represented by the correlation differences of miRNA–mRNA pairs between the two different classes. Thus, the goal of our method is to identify the correlation differences of co-expressed miRNAs and gene pathways between the two classes. The procedure of the analysis we performed in this study is shown in Fig. 1.

2.1 Method description

The details of our method are described in the following:

1. Preprocessing of miRNA and mRNA data. X miRNAs are clustered based on the correlations to identify the miRNAs with potential cooperative effects, and X miRNAs are assigned to p clusters. Similarly, Z mRNAs are divided into corresponding gene pathways by prior gene pathway knowledge (e.g. MSigDB database).
2. Principal component analysis (PCA). We adopt PCA to reduce the dimensionality of each data set of variables (e.g. miRNA clusters and mRNA pathways), respectively. Let us take miRNA data as an example to explain how PCA is performed. Let $A \in R^{n\times N}$ be the miRNA data set in cluster d ($d = 1, 2,\ldots, p$), where N is the number of miRNAs in cluster d. PCA is applied to A, and the first m principal components (PCs) are used as variables for further analysis. The number of m is determined by the method used in

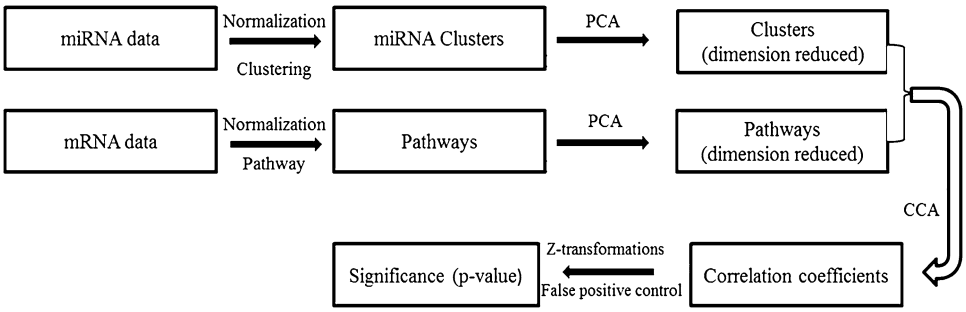

Fig. 1 The flow chart for the data analysis procedure

Soneson et al. (2010). The same procedure is applied to mRNA data in each gene pathway.

3. Identification of the correlation differences. The canonical correlation analysis (CCA) is a way of measuring the correlation between two multidimensional variables. In this study, CCA is performed to measure the correlation between clustered miRNAs and gene pathways. For a given miRNA cluster d and a pathway s, CCA is applied to these two data sets across all samples; then we take the first pair of canonical variables to calculate the correlation coefficient ρ_{kds} ($k = 1$ or 2) in each class. To test the null hypothesis H0:

$\rho_{1ds} = \rho_{2ds}$ versus alternative hypothesis H_1: $\rho_{1ds} \neq \rho_{2ds}$, we adopt a metric developed by Fisher (Fisher 1936):

$$D_{ds} = \frac{z_{1ds} - z_{2ds}}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}} \quad (1)$$

in which z_{kds} ($k = 1$ or 2) is

$$z_{kds} = 0.5 \log_e \left| \frac{1 + \rho_{kds}}{1 - \rho_{kds}} \right| \quad (2)$$

In Eq. (1) n_1 and n_2 are numbers of samples in class 1 and class 2, and the D value represents the correlation coefficient difference of ρ_{1ds} and ρ_{2ds}, which can be examined using a critical value of the standard normal distribution (Kraemer 2006). Then, we adjust the estimated significance level to account for multiple hypotheses testing by q values (Storey 2002).

Additionally for a given pair of a miRNA cluster and a gene pathway with significant interaction effect, the correlation tests for individual microRNA–mRNA pairs in the miRNA cluster and gene pathway are performed by Eqs. (1) and (2) to select the individual miRNA–mRNA pairs with significant interaction. For the identified miRNA–mRNA pairs after multiple testing adjustment (Storey 2002), we make the target prediction analysis by using "microRNA Target Filter" from Ingenuity Pathway Analysis [IPA, (Ingenuity System)] and find the dysregulation patterns of miRNA–mRNA or mRNA–mRNA.

2.2 Data and gene pathway information

The GBM data used in this study are publicly available from the website of The Cancer Genome Atlas (TCGA). The patients used in this study include two subtypes of GBM: Pro-Neural, Neural (Verhaak et al. 2010). We performed a data preprocessing according to the following criteria: (1) select the individuals which have both miRNA and mRNA data; (2) eliminate the outliers using the method proposed by Filzmoser et al. (2008); (3) to reduce the bias in clustering analysis by differential co-expressed miRNAs across different groups (Bhattacharyya and Bandyopadhyay 2013; Chia and Karuturi 2010), we chose the similar sample sizes for Pro-Neural and Neural subtypes. Thus, the final data set we used has 30 samples with 15 patients each in the subtypes of Pro-Neural and Neural. Each included sample has miRNA (1,510 probes) and mRNA expression data (22,277 probes). In addition, we gather publicly available pathway information from a database: The Molecular Signatures Database (MSigDB), in which there are a total of 880 pathways.

3 Results

All the 1,510 probes in miRNA data are clustered based on their correlations. The dendrogram is shown in Fig. 2. It yields 13 clusters with a cut-off value at 0.850, and in Table 1, the number of probes is shown for each cluster. The correlation between each pair of miRNA cluster and pathway in the two subtypes of GBM is calculated by the proposed approach described above. Three pairs with

significant correlation difference between Pro-Neural and Neural subtypes are shown in Table 2.

As shown in Table 2, we observe that in the subtype Neural, the miRNA cluster 8 and gene pathway 773 show little correlation with the correlation coefficient 0.117. However, they are strongly correlated in the subtype of Pro-Neural with the correlation coefficient 0.963. There is a significant difference of correlation of the cluster 8 and the gene pathway 773 between subtypes Pro-Neural and Neural. This difference indicates that there are interactions between some miRNAs and mRNAs across subtypes Pro-Neural and Neural, and these interactions between miRNAs and mRNAs might imply that there exists change of miRNA–mRNA regulation patterns between the two subtypes. Similarly, other two miRNA Cluster-Gene Pathway pairs show significant correlation differences across the subtypes Pro-Neural and Neural.

It is noted that among the three pairs of miRNA clusters and pathways in Table 2, miRNA cluster 1 is involved in two pairs. It suggests that miRNAs in cluster 1 may play an important role in miRNA–mRNA interaction effects in the subtypes Pro-Neural and Neural. Moreover, there are many overlap genes between the pathway 138 and 658. Thus, we take the miRNA cluster 1 and pathway 138 (KEGG_LONG_TERM_DEPRESSION) as an example for further identification of miRNA–mRNA pairs with interaction effects. After multiple testing adjustments, 72 miRNA–mRNA pairs are identified with interaction effects at false discovery rate (FDR) ≤ 0.05. These pairs include 13 genes and 16 miRNAs.

From the view of gene functions, previous studies have indicated that some of these 13 genes identified in our analysis are related to glioma. For example, IGF1R is a transmembrane tyrosine kinase that is frequently overexpressed by tumors. IGF1R has been shown to be abnormally active in gliomas (Gammeltoft et al. 1988), and its inhibition prevents tumor growth in preclinical models (Kiaris et al. 2000). IGF1R expression in glioblastoma is shown to elicit a host response leading to protection from unmodified tumor cells (Bielen et al. 2011; Riedemann and Macaulay 2006). The protein encoded by GNAS is a key component of many signal transduction pathways. It is reported that there is an association between GNAS genotype and survival among patients suffering from GBM (El Hindy et al. 2011). In addition, some miRNAs identified in our analysis have been found to be associated with gliomas. For example, hsa-miR-218 plays a critical role in the progression of many human cancers as a tumor suppressor (Liu et al. 2012b). Hsa-miR-218 is involved in preventing the invasiveness of glioma cells (Song et al. 2010). Hsa-miR-128 is a brain-enriched miRNA, and is identified as a tumor suppressive miRNA that has been shown to regulate neuronal differentiation, maturation, and/or survival. Hsa-miR-128 can repress human

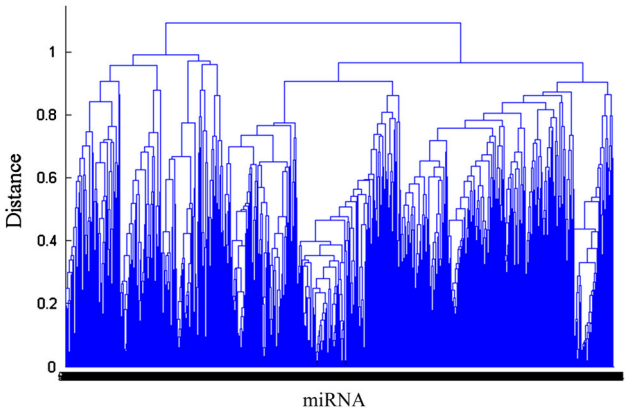

Fig. 2 Clustering chart for all the 1,510 miRNA probes

Table 1 The number of miRNA probes in the 13 clusters

Cluster no.	1	2	3	4	5	6	7	8	9	10	11	12	13
No. of probes	95	144	204	23	4	110	6	454	286	16	41	108	19

Table 2 A summary of miRNA cluster-gene pathway pairs with significant interaction effects

Pairs	Correlation coefficient		Correlation difference (p value)	FDR
	Pro-Neural	Neural		
Cluster 8 v.s. Gene Pathway 773	0.963	0.117	5.01E−06	0.024
Cluster 1 v.s. Gene Pathway 138	0.290	0.972	6.81E−06	0.024
Cluster 1 v.s. Gene Pathway 658	0.551	0.985	7.88E−06	0.024

Pathway 773: REACTOME_ETHANOL_OXIDATION; Pathway 138: KEGG_LONG_TERM_DEPRESSION; Pathway 658: REACTOME_REGULATION_OF_INSULIN_SECRETION_BY_GLUCAGON_LIKE_PEPTIDE_1

glioblastoma cell growth and invasiveness (Bielen et al. 2011; Papagiannakopoulos et al. 2012).

These miRNA–mRNA interactions indicate the change of regulation pattern across the Pro-Neural and Neural groups: (1) between miRNAs and mRNAs; (2) between mRNAs and mRNAs. For the following analysis, we will focus on hsa-miR-218 as an example to investigate the change of regulation pattern. There are two genes, ITPR1 and CRH, which have interaction effects with hsa-miR-218. For ITPR1, in the Pro-Neural group, there is no evidence that the expression of gene ITPR1 is correlated to hsa-miR-218, but in the Neural group, they show a very strong correlation, with range from 0.820 to 0.899. Based on target prediction analysis for hsa-miR-218 by "microRNA Target Filter" from ingenuity pathway analysis (IPA), which contains the databases: TarBase, TargetScan, and miRecords, it is found that ITPR1 is not the target gene of hsa-miR-218. Meanwhile, it is known that miRNAs usually result in translational repression of mRNAs. Thus, one possibility is that there is a target gene of hsa-miR-218, which results in the interaction between hsa-miR-218 and ITPR1. By correlation analysis and target prediction analysis for hsa-miR-218, it is observed that GNAI3, as a target gene of hsa-miR-218, shows the correlation difference with hsa-miR-218 between Pro-Neural and Neural groups, and shows a strong negative correlation with ITPR1, as shown in Table 3. The result implies that a dysregulation of the hsa-miR-218-ITPR1 pair across the Pro-Neural and Neural groups can be due to the dysregulation of hsa-miR-218–GNAI3 pair. In addition, in miRNA cluster 1, we also identify several miRNAs (e.g. hsa-miR-128), for which GNAI3 is not their target gene. However, there are also similar relationships as GNAI3 and ITPR1. This is mainly due to strong correlations between these miRNAs and hsa-miR-218, and they may carry out cooperative function in regulating mRNA expression. In this study, we clustered the miRNAs with similar co-expression profiles across all samples. For the identified miRNA clusters which showed interaction effects with gene pathways, some miRNAs in the same clusters have been reported to have co-expressed patterns. For example, in miRNA cluster 1, hsa-mir-103,128 and 218, were identified with co-expressed patterns (Sengupta and Bandyopadhyay 2011).

4 Discussion and conclusion

Interaction analysis of miRNA and mRNA data is helpful to better understand how miRNAs regulate mRNA expressions and how the regulation changes affect phenotypes. However, it is a challenge to identify the miRNA–mRNA interactions related to complex diseases due to the high-dimensional data. For example, with 1,000 miRNAs and 20,000 mRNAs, there will be 2.0×10^7 combinations of miRNAs and mRNAs. Recently, some studies indicated that miRNAs work in clusters to accomplish their function throughout many biological processes (Leung et al. 2008). MiRNAs located in the same cluster are usually co-regulated and co-expressed (Tanzer and Stadler 2004, 2006). In addition, it is well known that genes often cooperate with each other to perform various cellular functions and are organized into functional modules with densely connected genes within gene pathways (Newman 2006). Thus, we proposed a method to detect the enrichment of miRNA and mRNA interactions between co-expressed miRNAs clusters and gene pathways.

A major challenge in our analysis is the high dimensions of the miRNAs and mRNAs in each miRNA cluster and gene pathway. A natural solution would be to reduce the dimensions. PCA is adopted for dimension reduction to capture properties of miRNA and mRNA expressions. In the GBM data analysis, we consider two PCs for each miRNA cluster and each gene pathway, respectively, according to Soneson et al. (2010). The first two PCs averagely represented 68 % variation of the miRNA data and 51 % variation of the mRNA data. The scree plots of variance distribution of PCs (data not shown) become flat from the third components, indicating that the rest of the components mostly contain noise. It is noted that it is not our intention to suggest that two PCs will be sufficient for all practical data analysis. Rather, we intend to raise the awareness of the extra information brought by PCs beyond the first one or two. In practical data analysis, we suggest that researchers explore different numbers of PCs, and select the proper number based on the characteristics of specific data set, such as the biological implications and predictive power of the identified differential pathways. The elapsed time for the computation is $\sim$16 s in the described case, running on a computer with 64-bit

Table 3 The correlations of ITPR1and hsa-miR-218 with GNAI3 in Pro-Neural and Neural classes

		ITPR1			hsa-miR-218	
		203710_at	211323_s_at	216944_s_at	201179_s_at	201180_s_at
GNAI3 201180_s_at	Neural	−0.749	−0.671	−0.696	−0.743	−0.766
	Pro-Neural	−0.744	−0.589	−0.727	−0.053	−0.065

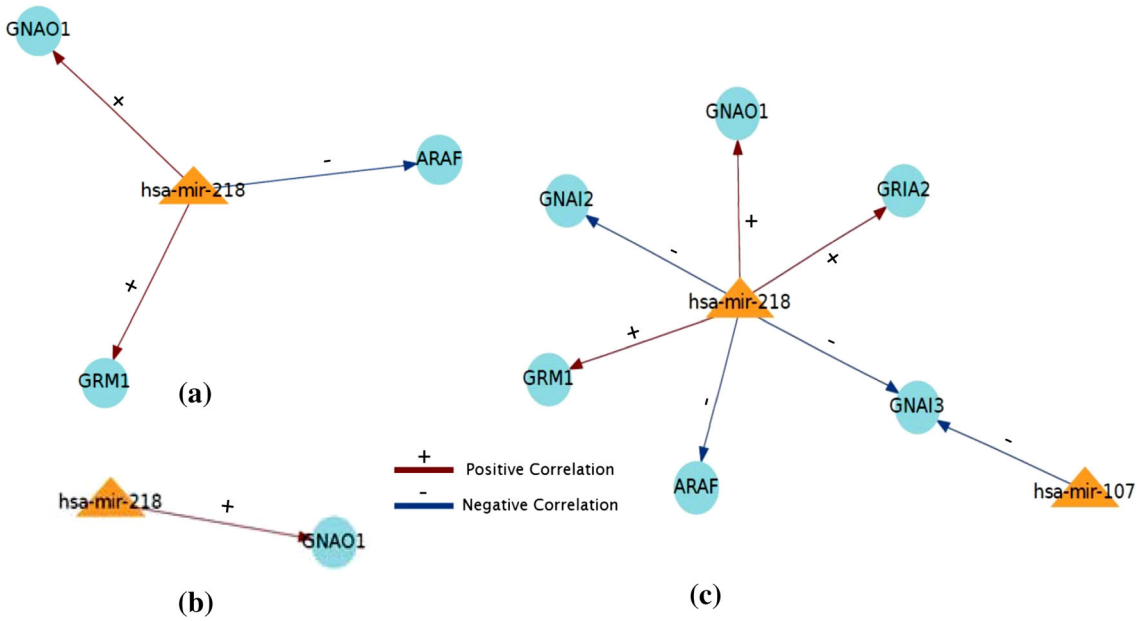

Fig. 3 MAGIA2 results using mRNA and miRNA expression data in pathway 138 and miRNA cluster 1. **a** Using the data from both Pro-Neural and Neural groups, **b** using the data from only Pro-Neural group, **c** using the data from only Neural group

Windows 7, Intel (R) Core (TM) i7-2600K CPU @3.4-GHz, and 8 GB RAM. The elapsed time may vary upon different data sizes and computer conditions.

Applications on the real data suggest that, with the proposed approach, we are able to identify the pairs of miRNA clusters and gene pathways, which show interaction effects enrichment related to GBM status. MAGIA2 (a popular miRNA–mRNA integrated analysis tool) also yields the similar results for the pathway 138 and the miRNA cluster 1 as shown in Fig. 3. In Fig. 3b, c, it can be seen that the correlation coefficients of hsa-miR-218 and GNAI3 show a difference across the Pro-Neural and Neural groups. In Fig. 3a, if using the data from both Pro-Neural and Neural groups, the dysregulation between hsa-miR-218 and GNAI3 will not be observed. These results indicate that there is dysregulation related to the trait under study, between hsa-miR-218 and GNAI3. Meanwhile, our method can give more information about regulation. For example, our method identifies the interaction effect between hsa-miR-218 and ITPR1. Although ITPR1 is not a target gene of hsa-miR-218, it is observed that GNAI3, as a target gene of hsa-miR-218, shows a high correlation with ITPR1. It is possible that the interaction effect between hsa-miR-218 and ITPR1 pair across the Pro-Neural and Neural groups can be due to the interaction effect between hsa-miR-218 and GNAI3. In addition, our method indicates the co-expressed miRNAs which are associated with the trait under study. So our method can give more regulation information and clues of the complex regulation network.

In summary, the novelty of our method is that we incorporate prior biological knowledge to detect the miRNA–mRNA interactions, which indicate the existence of dysregulation associated with phenotypes of interest. It enables us to get an insight into the roles of causal miRNAs and mRNAs in disease diagnosis and therapy development. Additionally, our method provides a way of integrating multi-omics data to delineate the knowledge of relevant molecular pathways of disease pathogenesis.

Web sources

TCGA GBM data: http://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp.
MSigDB: http://www.broadinstitute.org/gsea/downloads.jsp.
IPA: http://www.ingenuity.com/products/ipa.

References

Bhattacharyya M, Bandyopadhyay S (2013) Studying the differential co-expression of microRNAs reveals significant role of white matter in early Alzheimer's progression. Mol BioSyst 9:457–466

Bielen A et al (2011) Enhanced efficacy of IGF1R inhibition in pediatric glioblastoma by combinatorial targeting of PDGFRalpha/beta. Mol Cancer Ther 10:1407–1418

Burton PR et al (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661–678

Chhabra R, Dubey R, Saini N (2010) Cooperative and individualistic functions of the microRNAs in the miR-23a ~ 27a ~ 24-2 cluster and its implication in human diseases. Mol Cancer 9:232

Chia BK, Karuturi RK (2010) Differential co-expression framework to quantify goodness of biclusters and compare biclustering algorithms. Algorithms Mol Biol 5:23

Cordell HJ (2009) Detecting gene–gene interactions that underlie human diseases. Nat Rev Genet 10:392–404

Culverhouse R, Klein T, Shannon W (2004) Detecting epistatic interactions contributing to quantitative traits. Genet Epidemiol 27:141–152

El Hindy N et al (2011) Role of the GNAS1 T393C polymorphism in patients with glioblastoma multiforme. J Clin Neurosci 18:1495–1499

Farber CR (2010) Identification of a gene module associated with BMD through the integration of network analysis and genome-wide association data. J Bone Miner Res 25:2359–2367

Farber CR, Lusis AJ (2009) Future of osteoporosis genetics: enhancing genome-wide association studies. J Bone Miner Res 24:1937–1942

Feederle R et al (2011) The members of an Epstein–Barr virus microRNA cluster cooperate to transform B lymphocytes. J Virol 85:9801–9810

Filzmoser P, Maronna R, Werner M (2008) Outlier identification in high dimensions. Comput Stat Data Anal 52:1694–1711

Fisher R (1936) The use of multiple measurements in taxonomic problems. Ann Eugenics 7:179–188

Gammeltoft S et al (1988) Expression of two types of receptor for insulin-like growth factors in human malignant glioma. Cancer Res 48:1233–1237

Guan YF et al (2010) Functional genomics complements quantitative genetics in identifying disease–gene associations. Plos Comput Biol 6:e1000991

Juran BD, Lazaridis KN (2011) Genomics in the Post-GWAS Era. Semin Liver Dis 31:215–222

Kallberg H et al (2007) Gene–gene and gene–environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. Am J Hum Genet 80:867–875

Kiaris H, Schally AV, Varga JL (2000) Antagonists of growth hormone-releasing hormone inhibit the growth of U-87MG human glioblastoma in nude mice. Neoplasia 2:242–250

Kraemer HC (2006) Correlation coefficients in medical research: from product moment correlation to the odds ratio. Stat Methods Med Res 15:525–545

Leung WS et al (2008) Filtering of false positive microRNA candidates by a clustering-based approach. BMC Bioinformatics 9(Suppl 12):S3

Liu Y et al (2012a) Gene interaction enrichment and network analysis to identify dysregulated pathways and their interactions in complex diseases. BMC Syst Biol 6:65

Liu Y et al (2012b) MiR-218 reverses high invasiveness of glioblastoma cells by targeting the oncogenic transcription factor LEF1. Oncol Rep 28:1013–1021

Mackay TF (2014) Epistasis and quantitative traits: using model organisms to study gene–gene interactions. Nat Rev Genet 15:22–33

Mavrakis KJ et al (2011) A cooperative microRNA-tumor suppressor gene network in acute T-cell lymphoblastic leukemia (T-ALL). Nat Genet 43:673–678

Nelson MR et al (2001) A combinatorial partitioning method to identify multilocus genotypic partitions that predict quantitative trait variation. Genome Res 11:458–470

Newman ME (2006) Modularity and community structure in networks. Proc Natl Acad Sci 103:8577–8582

Papagiannakopoulos T et al (2012) Pro-neural miR-128 is a glioma tumor suppressor that targets mitogenic kinases. Oncogene 31:1884–1895

Phillips PC (2008) Epistasis–the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9:855–867

Purcell S et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575

Riedemann J, Macaulay VM (2006) IGF1R signalling and its inhibition. Endocr Relat Cancer 13(Suppl 1):S33–S43

Ritchie MD et al (2001) Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. Am J Hum Genet 69:138–147

Schwarz DF, Konig IR, Ziegler A (2010) On safari to Random Jungle: a fast implementation of Random Forests for high-dimensional data. Bioinformatics 26:1752–1758

Sengupta D, Bandyopadhyay S (2011) Participation of microRNAs in human interactome: extraction of microRNA–microRNA regulations. Mol BioSyst 7:1966–1973

Sha Q et al (2006) A combinatorial searching method for detecting a set of interacting loci associated with complex traits. Ann Hum Genet 70:677–692

Soneson C et al (2010) Integrative analysis of gene expression and copy number alterations using canonical correlation analysis. BMC Bioinform 11:191

Song L et al (2010) miR-218 inhibits the invasive ability of glioma cells by direct downregulation of IKK-beta. Biochem Biophys Res Commun 402:135–140

Storey JD (2002) A direct approach to false discovery rates. J Roy Stat Soc 64:478–498

Tanzer A, Stadler PF (2004) Molecular evolution of a microRNA cluster. J Mol Biol 339:327–335

Tanzer A, Stadler PF (2006) Evolution of microRNAs. Methods Mol Biol 342:335–350

Verhaak RG et al (2010) Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 17:98–110

Visweswaran S, Wong AK, Barmada MM (2009) A Bayesian method for identifying genetic interactions. AMIA Annu Symp Proc 2009:673–677

Wan X et al (2010) BOOST: a fast approach to detecting gene–gene interactions in genome-wide case–control studies. Am J Hum Genet 87:325–340

Zhang Y, Liu JS (2007) Bayesian inference of epistatic interactions in case-control studies. Nat Genet 39:1167–1173

Zhang J, Li J, Deng HW (2009) Identifying gene interaction enrichment for gene expression data. PLoS One 4:e8064