Gene 518 (2013) 209–214

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

# DiffCorr: An R package to analyze and visualize differential correlations in biological networks

Atsushi Fukushima *

*RIKEN Plant Science Center, Kanagawa 230-0045, Japan*

ARTICLE INFO

Available online 13 December 2012

*Keywords:*
Correlation network
Differential correlation
Network modules
R program

ABSTRACT

Large-scale "omics" data, such as microarrays, can be used to infer underlying cellular regulatory networks in organisms, enabling us to better understand the molecular basis of disease and important traits. Correlation approaches, such as a hierarchical cluster analysis, have been widely used to analyze omics data. In addition to the changes in the mean levels of molecules in the omics data, it is important to know about the changes in the correlation relationship among molecules between 2 experimental conditions. The development of a tool to identify differential correlation patterns in omics data in an efficient and unbiased manner is therefore desirable. We developed the DiffCorr package, a simple method for identifying pattern changes between 2 experimental conditions in correlation networks, which builds on a commonly used association measure, such as Pearson's correlation coefficient. DiffCorr calculates correlation matrices for each dataset, identifies the first principal component-based "eigen-molecules" in the correlation networks, and tests differential correlation between the 2 groups based on Fisher's z-test. We illustrated its utility by demonstrating biologically relevant, differentially correlated molecules in transcriptome coexpression and metabolite-to-metabolite correlation networks. DiffCorr can explore differential correlations between 2 conditions in the context of post-genomics data types, namely transcriptomics and metabolomics. DiffCorr is simple to use in calculating differential correlations and is suitable for the first step towards inferring causal relationships and detecting biomarker candidates. The package can be downloaded from the following website: http://diffcorr.sourceforge.net/.



## 1. Introduction

Large-scale "omics" data obtained from high-throughput technology, such as microarrays, can be used to infer underlying cellular regulatory networks in organisms. Typically, quantitative data analyses for such omics data are carried out to measure levels of cellular components or pathways with significantly altered mean levels of abundance between 2 experimental conditions, e.g., diseased and normal cells. In transcript profiling analysis, such genes are called "differentially expressed genes (DEGs)" (Pan, 2002; Reiner et al., 2003). As another approach to characterize cellular behavior, correlation approaches have been widely used for omics data, e.g., a hierarchical cluster analysis (Eisen et al., 1998). In addition to mean levels of abundance and detecting clustered molecules with similar profile patterns, changes in correlation patterns between molecules, referred to as "differential correlations," are also informative (Choi et al., 2005; de la Fuente, 2010).

Correlation network-based approaches are often used in the analysis of omics data including transcriptomics and metabolomics. In correlation networks (also referred to as gene coexpression networks in the case of transcriptomic data), molecules are represented by nodes in a graph and pairs of molecules are linked by undirected edges (Usadel et al., 2009). Such correlation networks have been successfully applied to a variety of problems in molecular biology. For example, coexpression networks have provided clues about the function of unknown genes associated with biosynthesis pathways, including those of glucosinolates and flavonoids in plant science (Saito et al., 2008). Graph-based clustering methods (Wang et al., 2010) have also been used to characterize condition- and genotype-dependent patterns in molecular abundances as a cluster or a module, which is a densely connected group in a correlation network; for example, see references Fukushima et al. (2011) and Fukushima et al. (2012).

Recently, differential network approaches, which are based on significant connectivity differences between 2 networks, have gained increasing attention (de la Fuente, 2010; Ideker and Krogan, 2012). In the context, a number of approaches exist, which identify differential correlations for large-scale omics datasets. The typical approaches for detecting differential correlations include topological overlap in a graph (Altay et al., 2011; Ray and Zhang, 2010; Tesson et al., 2010; Yu et al., 2011), extension of the traditional F-statistic (Lai et al., 2004), an additive model (Kostka and Spang, 2004), Fisher's z-test (Choi et al., 2005), an

*Abbreviations:* HCA, hierarchical cluster analysis; FDR, false discovery rate; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; PC, principal component; PCA, principal component analysis; GC–MS, gas chromatography coupled with mass spectrometry; CHS, chalcone synthase; *tt4*, *transparent testa 4*.

* RIKEN Plant Science Center, Yokohama, Kanagawa 230-0045, Japan. Tel.: +81 45 503 9491; fax: +81 45 503 9489.
*E-mail address:* a-fukush@psc.riken.jp.

0378-1119 
http://dx.doi.org/10.1016/j.gene.2012.11.028

interaction score based on the Renyi relative entropy (Cho et al., 2009), the Haar basis (Gillis and Pavlidis, 2009), combination of graphical Gaussian model and posterior odds ratio (Chu et al., 2011), the liquid association concept (Li, 2002; Valcarcel et al., 2011), a combination of robust correlations and hypothetical testing (called as ROS-DET) (Kayano et al., 2011), random re-sampling methods (Watson, 2006), graph-theoretic statistics (Odibat and Reddy, 2012), and an empirical Bayesian approach (Dawson and Kendziorski, 2012; Dawson et al., 2012). Of these, Liu and co-workers have implemented several methods to identify differential coexpressions in their R package (http://www.r-project.org), DCGL (Liu et al., 2010). The simplest technique, based on Fisher's z-test of correlation coefficient to identify differential correlations, has however not been widely used yet and, to the best of our knowledge, is not implemented for omics data in the available R packages.

We present the R package DiffCorr as a means to identify differential correlations between 2 conditions based on Fisher's z-test. The program also contains a function that provides a comprehensive list of differentially correlated pairs in a dataset as a text file. We illustrate its utility by demonstrating biologically relevant, differentially correlated clusters in correlation networks derived from transcriptomic and metabolomic data.

## 2. Materials and methods

### 2.1. Implementation

In this section, we describe the features, functionalities, and the structure of the DiffCorr package. The software package implemented in R (http://www.r-project.org) is available as a zip file (Supplementary data 1) and the source file (Supplementary data 2). It runs under major operating systems, such as Microsoft Windows.

### 2.2. Main functions

The R program DiffCorr contains a set of functions for identifying differential correlations in a correlation network derived from large-scale omics data. Functions in DiffCorr package can be divided into 3 main categories: (1) module detection, constructing correlation network, and calculating the eigen-molecules for each condition; (2) visualization of eigen-molecule networks; and (3) export of the results of testing based on Fisher's z-test (Fig. 1 and Table 1).

(1) *get.eigen.molecule*: extracts conditional modules derived from a hierarchical cluster analysis (HCA) using the *cluster.molecule* function. For visualization of modules, *get.eigen.molecule.graph* also provides a graph object of eigengene (Langfelder and Horvath, 2007) using the igraph package (Csardi and Nepusz, 2006) (Fig. 1).
(2) *plot.DiffCorr.group*: draws module members for each condition. This function is based on the plot function using the igraph package (Csardi and Nepusz, 2006). This provides profile patterns of module members for each module.
(3) *comp.2.cc.fdr*: exports a list of significantly differential correlations as a text file. This function uses the fdrtool package (Strimmer, 2008) to control the false discovery rate (FDR). The resulting file contains molecule IDs (e.g., probe-set ID and metabolite name), conditional correlation coefficients, the p-values of the correlation test, the difference of the 2 correlations, the corresponding p-values, and the result of Fisher's z-test with controlling FDR.

More detailed statistical descriptions for identifying differentially correlated molecules are in the next subsection.

### 2.3. Calculation of differential correlations

Fisher's z-test was used to identify significant differences between 2 correlations, based on its stringency test and its provision of conservative estimates of true differential correlations among molecules between 2 experimental conditions in the omics data. To test whether the 2 correlation coefficients were significantly different, we first transformed correlation coefficients for each of the 2 conditions, $r_A$ and $r_B$, into $Z_A$ and $Z_B$, respectively. The Fisher's transformation of coefficient $r_A$ is defined by: $Z_A = \frac{1}{2}\log\frac{1+r_A}{1-r_A}$. Similarly,

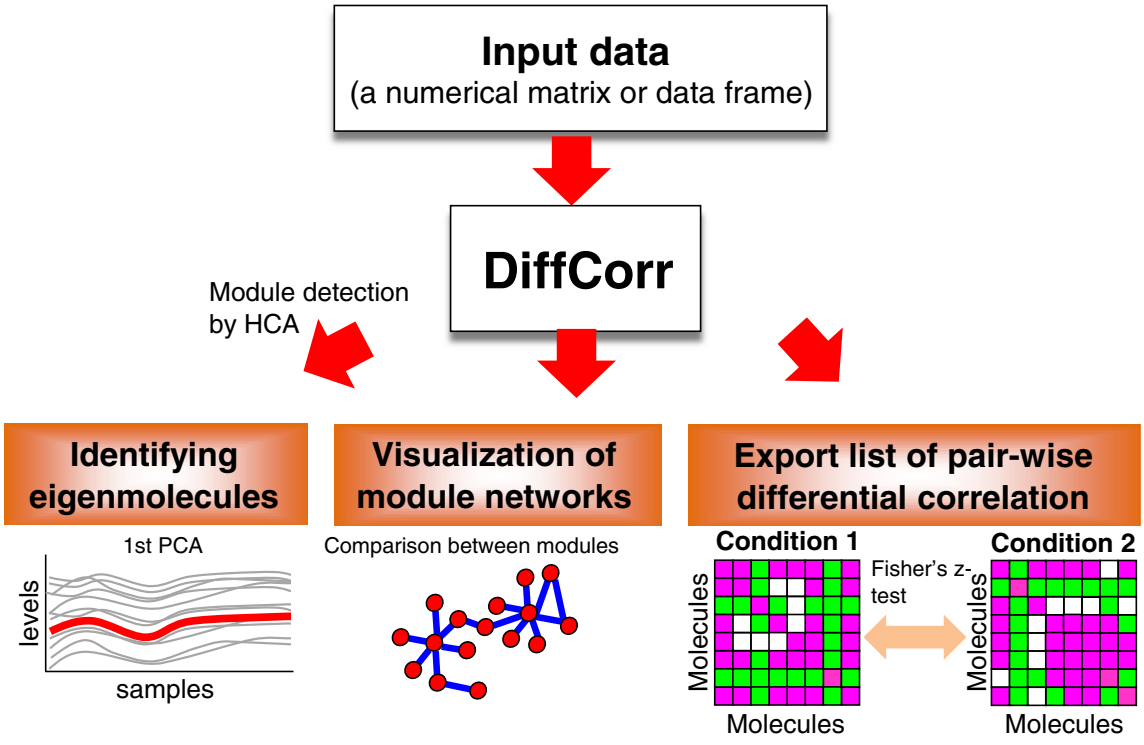


**Fig. 1.** An overview of DiffCorr analysis steps and main functions in DiffCorr. An outline of the DiffCorr approach with the 3 main processes. HCA, hierarchical cluster analysis.

**Table 1**
Main functions on DiffCorr package.

| Name | Description |
|---|---|
| *get.eigen.molecule* | Returns eigen-molecules or eigengenes. |
| *plot.DiffCorr.group* | Draws conditional modules. |
| *comp.2.cc.fdr* | Provides a list of significantly differential correlations as a text file. |

we transform coefficient $r_B$ into $Z_B$. Differences between the two correlations can be tested using the following Eq. (1)

$$Z = \frac{Z_A - Z_B}{\sqrt{\frac{1}{n_A - 3} + \frac{1}{n_B - 3}}} \quad (1)$$

$n_A$ and $n_B$ represent the sample size for each of the conditions for each biomolecule pair (Fukushima et al., 2011; Fukushima et al., 2012; Morgenthal et al., 2006). The $Z$ value has an approximately Gaussian distribution under the null hypothesis that the population correlations are equal. Controlling the FDR described by Benjamini and Hochberg (1995) is a stringent and practical method in multiple testing problems. However, while it assumes all tests to be independent, this is not the case for correlation tests. We therefore used the local false-discovery rate (fdr) derived from the fdrtool package (Strimmer, 2008).

### 2.4. Identifying eigen-molecules

To test whether 2 correlated modules in correlation networks are significantly different, we first calculate the eigen-molecule or "eigengene" (Langfelder and Horvath, 2007) in the network as a representative correlation pattern within each module. The eigen-molecule is based on the first principal component (PC) of a data matrix of a module extracted from a hierarchical cluster analysis using the *hclust* function in R. The *get.eigen.molecule* function uses the pcaMethods package (Stacklies et al., 2007) to perform principal component analysis (PCA) and returns the top 10 PCs (default). Using these eigen-molecule modules we can also test differential correlation between modules in addition to pair-wise differential correlations between molecules.

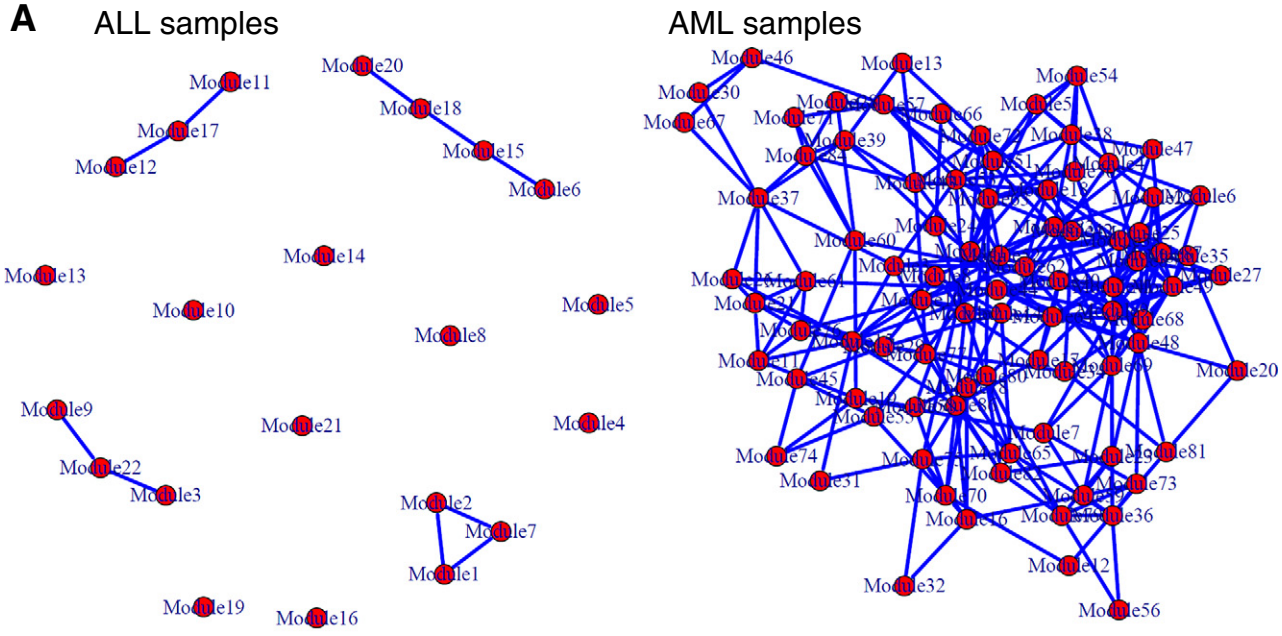


**Fig. 2.** Representations of the module network and differential coexpressions. (A) Images of the module networks from the Golub dataset, including ALL and AML samples. (B) A typical result of pair-wise differential correlations from DiffCorr.

**Table 2**
Top 10 list of differentially coexpressed genes from the Golub dataset.

| Molecule X | Description of molecule X | Molecule Y | Description of molecule Y | r1 (ALL) | r2 (AML) | lfdr[a] (difference) |
|---|---|---|---|---|---|---|
| D43949_at | KIAA0082 | HG4185-HT4455_at | Estrogen Sulfotransferase, Ste | −0.09 | 0.98 | 1.1E−03 |
| HG2873-HT3017_at | Ribosomal protein L30 (RPL30) | X03689_s_at | Eukaryotic translation elongation factor 1 alpha 1 | 0.31 | 1.00 | 1.1E−03 |
| HG3214-HT3391_at | Metallopanstimulin 1 | D49824_s_at | HLA-B null allele mRNA | 0.33 | 1.00 | 1.1E−03 |
| HG3214-HT3391_at | Metallopanstimulin 1 | X03689_s_at | Eukaryotic translation elongation factor 1 alpha 1 | 0.31 | 1.00 | 1.1E−03 |
| HG3364-HT3541_at | Ribosomal protein L37 | D49824_s_at | HLA-B null allele mRNA | 0.31 | 1.00 | 1.1E−03 |
| HG3364-HT3541_at | Ribosomal protein L37 | X03689_s_at | Eukaryotic translation elongation factor 1 alpha 1 | 0.30 | 1.00 | 1.1E−03 |
| HG4319-HT4589_at | Ribosomal protein L5 | D49824_s_at | HLA-B null allele mRNA | 0.31 | 0.99 | 1.1E−03 |
| HG4319-HT4589_at | Ribosomal protein L5 | Z49148_s_at | H.ribosomal protein L29 | 0.29 | 1.00 | 1.1E−03 |
| L06499_at | Ribosomal protein L37a (RPL37A) | D49824_s_at | HLA-B null allele mRNA | 0.31 | 1.00 | 1.1E−03 |
| L06499_at | Ribosomal protein L37a (RPL37A) | X03689_s_at | Eukaryotic translation elongation factor 1 alpha 1 | 0.27 | 1.00 | 1.1E−03 |

[a] Local FDR.

### 2.5. Scaling and clustering

When applying cluster analysis to omics data, data pre-treatment is important to the resulting clusters. For example, the outcome of PCA, which identifies the directions capturing greatest variance in dataset, is greatly affected by different pre-treatment methods. DiffCorr therefore implements the *scalingMethods* function to integrate different pre-treatment methods with down-stream correlation analyses. The function includes auto-scaling (unit-variance scaling), range scaling, Pareto scaling, vast scaling, level scaling, and power transformation (van den Berg et al., 2006).

### 2.6. Availability and requirements

- Project name: DiffCorr
- Project home page: http://diffcorr.sourceforge.net/
- Operating system(s): Platform independent
- Programming language: R
- Other requirements: R≥2.14.1
- License: GNU GPL
- Any restrictions to use by non-academics: None.

The package itself is available as online Supplementary materials. It will be submitted to CRAN repository of R packages.

## 3. Results and discussion

To demonstrate the usefulness of DiffCorr package, we describe and discuss the results from analysis of transcriptomic and metabolomic datasets.

### 3.1. The Golub data (ALL/AML leukemia dataset)

This dataset consist of gene expression profiles from 38 tumor samples including 2 different leukemia subtypes: 27 acute lymphoblastic leukemia (ALL) and 11 acute myeloid leukemia (AML) samples (Golub et al., 1999). The microarray platform used, Affymetrix GeneChip HuGeneFL (known as HU6800), contains 6800 probe-sets. We filtered out all the probe-sets with negative values in any samples, resulting in 2568 genes. Using the DiffCorr package, the genes were grouped according to their expression patterns in each subtypes (ALL or AML) using the *cluster.molecule* function. We used (1 − correlation coefficient) as a distance measure (the cutoff value was a coefficient of 0.6) based on the *cutree* function. We then visualized the module network using the *get.eigen.molecule* and *get.eigen.molecule.graph* functions (Fig. 2A). The *comp.2.cc.fdr* function provides the resulting pair-wise differential correlations from a dataset (Fig. 2B). Table 2 shows the top 10 significantly differential coexpressions (FDR<0.05), which were all AML-specific correlations. For example, a correlation between D43949_at (KIAA0082) and HG4185-HT4455_at (Estrogen Sulfotransferase, Ste) was −0.09 in ALL, and 0.98 in AML. As can be seen, the list includes genes encoding the ribosomal proteins L5, L29, L30, L37, and L37a. The list also contained eukaryotic translation elongation factor 1 alpha 1 (eEF1A1), which is associated with translation elongation factor activity and has oncogenic potency. The DiffCorr package also detects oppositely correlated pairs where, for example, 2 molecules exhibit positive correlation in one condition and negative correlation in the other condition, a condition referred to as a "switching mechanism" (Kayano et al., 2011). Table 3 shows the list of switching mechanisms of gene expression between ALL and AML samples. Several oncogenes are present, including the *IQ motif containing GTPase activating protein 1* (*IQGAP1*), which is often over-expressed in cancer (see review White et al., 2009).

### 3.2. The metabolome data of flavonoid-deficient Arabidopsis

Kusano et al. investigated flavonoid-deficient *Arabidopsis thaliana* (the model plant) and the wild-type by using gas chromatography coupled with mass spectrometry (GC–MS)-based metabolite profiling

**Table 3**
Correlated gene pairs changed to the opposite direction between ALL/AML samples in the Golub dataset.

| Molecule X | Description of molecule X | Molecule Y | Description of molecule Y | r1 (ALL) | r2 (AML) | lfdr[a] (difference) |
|---|---|---|---|---|---|---|
| X05130_s_at | Prolyl 4-hydoxylase beta subunit (EC 1.14.11.2) | X03689_s_at | Fragment elongation factor TU (N-terminus) | 0.84 | −0.79 | 3.1E−03 |
| D49824_s_at | HLA-B null allele mRNA | M94880_f_at | MHC class I (HLA-A*8001) mRNA. | 0.83 | −0.82 | 1.1E−03 |
| D17716_at | N-acetylglucosaminyltransferase V | X01677_f_at | Liver glyceraldehyde-3-phosphate dehydrogenase (G3PD EC 1.2.1.12) | 0.81 | −0.74 | 1.3E−02 |
| U57721_at | L-kynurenine hydrolase | U72512_at | B-cell receptor associated protein (hBAP) alternatively spliced 3′UTR | 0.73 | −0.85 | 9.7E−03 |
| U34343_at | 13 kDa differentiation-associated protein | U23852_s_at | T-lymphocyte specific protein tyrosine kinase p56lck (lck) abberant | 0.72 | −0.86 | 4.5E−03 |
| L77730_at | A3 adenosine receptor (ADORA3) exon 2 | L33075_at | Ras GTPase-activating-like protein (IQGAP1) | −0.59 | 0.94 | 1.1E−03 |
| U32849_at | Nmi | S82297_at | Beta 2-microglobulin | −0.71 | 0.90 | 1.1E−03 |
| X78627_at | H.translin | Z26876_at | H.ribosomal protein L38 | −0.71 | 0.88 | 3.1E−03 |
| X78627_at | H.translin | M36072_at | Ribosomal protein L7a (surf 3) large subunit | −0.73 | 0.88 | 1.1E−03 |
| M64716_at | Ribosomal protein S25 | X78627_at | H.translin | −0.74 | 0.88 | 1.1E−03 |

[a] Local FDR.

**Table 4**
Correlated metabolite pairs changed between wild-type/*tt4* plants.

| Molecule X | Molecule Y | r1 (WT) | r2 (*tt4*) | *p* (difference) | lfdr[a] (difference) |
|---|---|---|---|---|---|
| Malate | Phe | 0.45 | 0.89 | 8.6E−03 | 6.1E−01 |
| Malate | Tyr | 0.44 | 0.94 | 6.8E−04 | 6.1E−01 |
| Malate | Sinapate | 0.77 | 0.89 | 2.5E−01 | 8.6E−01 |
| Malate | Sinapate | 0.50 | 0.89 | 1.8E−02 | 7.2E−01 |
| Phe | Shikimate | 0.69 | 0.84 | 2.9E−01 | 8.6E−01 |

[a] Local FDR.

(Fukushima et al., 2011; Kusano et al., 2007). This mutant lacks gene encoding chalcone synthase (CHS) and cannot synthesize any flavonoids, which are plant secondary metabolites that function as protectants against ultraviolet B (UV-B) irradiation. This dataset consists of the metabolite profiles of 37 samples, including 2 genotypes: 17 Columbia-0 wild-type and 20 *transparent testa 4* (*tt4*, flavonoid deficient mutant) plants. The data also contain a wide-range of primary metabolites including amino acids, organic acids, fatty acids, sugars and sugar alcohols. The DiffCorr package detected significant differential correlations between sinapate and aromatic metabolites in *tt4* and wild-type plants (Table 4). As reported previously (Kusano et al., 2007), aromatic metabolites in the shikimate pathway, namely sinapate, phenylalanine (Phe), and tyrosine (Tyr), were significantly correlated in *tt4*, but not in wild-type plants. This implies a linkage with the role of sinapoyl-malate against UV-B irradiation in the flavonoid-less *tt4* mutant. We next showed that *Arabidopsis* attempts to compensate for deficiency in either flavonoid or sinapoyl-malate production by over-accumulating the alternative protectants (Kusano et al., 2011). These results suggest that DiffCorr can be applied to not only transcriptomic data, but to other post-genomics data type, including metabolomic data.

### 3.3. Assessment of the results from DiffCorr package and their interpretation

To assess the results of the DiffCorr program, we performed gene ontology (GO) enrichment analysis for pair-wise differential coexpressions from DiffCorr and ROS-DET (Kayano et al., 2011) with the GOstats package (Falcon and Gentleman, 2007). Table 5 lists the top 10 pairs with significantly over-represented GO terms involved in differential coexpressions between the ALL and AML samples in the Golub dataset referred to above (Golub et al., 1999). We have shown the best biological terms using their best *p*-values (hyper geometric test). We also compared the top 100 ranked gene-pairs with the previously reported results in the original papers from Golub et al. (1999). In our list, we detected 4 differential genes, encoding RbAp48 (NCBI ID: X74262), Dynein light chain (U32944), SRP9 (U20998), and LTC4 synthase (U50136), which were reported in Golub et al. (1999) among the most highly correlated genes with AML-ALL classification. In the top 100 ranked gene-pairs from ROS-DET, there were 2 genes of the AML-ALL class distinction. The gene encoding LTC4 synthase (U50136) was commonly identified in both the programs, while the gene encoding Cyclin D3 (M92287) was specifically identified by ROS-DET. These results suggest that simple differential coexpressions can be used for identifying novel biomarker candidates.

**Table 5**
Top 10 pairs with significant overrepresented GO terms involved in "switching mechanism" coexpressions between the ALL and AML samples in the Golub dataset. We have listed the results obtained using DiffCorr and ROS-DET (Kayano et al., 2011).

| Method | Category | GO-ID | Functional term | *p*-Value | Count/size |
|---|---|---|---|---|---|
| DiffCorr | BP | GO: 0050690 | Regulation of defense response to virus by virus | 1.3E−04 | 3/13 |
| | MF | GO: 0001609 | Adenosine receptor activity, G-protein coupled | 8.4E−03 | 1/1 |
| | CC | GO: 0042612 | MHC class I protein complex | 1.3E−03 | 2/7 |
| ROS-DET | BP | GO: 0006406 | mRNA export from nucleus | 6.5E−03 | 2/17 |
| | MF | GO: 0032403 | Protein complex binding | 7.2E−03 | 3/60 |
| | CC | GO: 0000932 | Cytoplasmic mRNA processing body | 1.4E−02 | 1/2 |

Generally, correlation or differential correlation does not necessarily reflect causal relationships (Markowetz and Spang, 2007; Steuer, 2006). Although they should include causal relationships, it should be noted that correlation networks based on undirected graphs cannot easily distinguish between direct and indirect dependencies in biomolecular networks. DiffCorr package can provide the first step towards inferring causal relationships between molecules in regulatory networks. The examples presented here highlight the potential of the differential network approach using direct measurements based on Fisher's z-test, as well as some of the intrinsic limitations of the method. The results suggest that the DiffCorr package might prove useful for identifying significant switching mechanism and for detecting biomarker candidates. For data matrices containing over 10,000 molecules, a large amount of memory (e.g., several dozen of gigabytes) is required to calculate correlations. These analyses were performed using the R statistical package with the pcaMethods, igraph, and fdrtool, on a 64-bit Windows computer with 24 GB physical memory.

## 4. Conclusions

The R package DiffCorr affords users a simple and effective framework to detect differential correlations between 2 conditions in omics data. The package is based on Fisher's z-test and is simple to calculate differential correlations. The approach is useful for the first step towards inferring causal relationships and detecting biomarker candidates. The DiffCorr base on the concept of "differential network biology" (de la Fuente, 2010; Ideker and Krogan, 2012) is suitable not only for transcriptomic and metabolomic data, but also for proteomic data, genome-wide association studies and integrated omics data (Fukushima et al., 2009; Kim et al., 2010).

## Conflict of interest

The author declares that they have no conflict of interest.

## Acknowledgments

I would like to thank Miyako Kusano for the critical reading of this paper. This work was partly supported by a Grant-in-Aid for Young Scientists (B; grant no. 23700355) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

Altay, G., Asim, M., Markowetz, F., Neal, D.E., 2011. Differential C3NET reveals disease networks of direct physical interactions. BMC Bioinforma. 12, 296.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B. 57, 289–300.
Cho, S.B., Kim, J., Kim, J.H., 2009. Identifying set-wise differential co-expression in gene expression microarray data. BMC Bioinforma. 10, 109.
Choi, J.K., Yu, U., Yoo, O.J., Kim, S., 2005. Differential coexpression analysis using microarray data and its application to human cancer. Bioinformatics 21, 4348–4355.
Chu, J.H., Lazarus, R., Carey, V.J., Raby, B.A., 2011. Quantifying differential gene connectivity between disease states for objective identification of disease-relevant genes. BMC Syst. Biol. 5, 89.
Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. Int. J. Complex Syst. (p. http://necsi.org/events/iccs6/viewabstract.php?id=88).
Dawson, J.A., Kendziorski, C., 2012. An empirical Bayesian approach for identifying differential coexpression in high-throughput experiments. Biometrics 68, 455–465.

Dawson, J.A., Ye, S., Kendziorski, C., 2012. R/EBcoexpress: an empirical Bayesian framework for discovering differential co-expression. Bioinformatics 28, 1939–1940.
de la Fuente, A., 2010. From 'differential expression' to 'differential networking' — identification of dysfunctional regulatory networks in diseases. Trends Genet. 26, 326–333.
Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D., 1998. Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. U. S. A. 95, 14863–14868.
Falcon, S., Gentleman, R., 2007. Using GOstats to test gene lists for GO term association. Bioinformatics 23, 257–258.
Fukushima, A., Kusano, M., Redestig, H., Arita, M., Saito, K., 2009. Integrated omics approaches in plant systems biology. Curr. Opin. Chem. Biol. 13, 532–538.
Fukushima, A., Kusano, M., Redestig, H., Arita, M., Saito, K., 2011. Metabolomic correlation-network modules in *Arabidopsis* based on a graph-clustering approach. BMC Syst. Biol. 5, 1.
Fukushima, A., et al., 2012. Exploring tomato gene functions based on coexpression modules using graph clustering and differential coexpression approaches. Plant Physiol. 158, 1487–1502.
Gillis, J., Pavlidis, P., 2009. A methodology for the analysis of differential coexpression across the human lifespan. BMC Bioinforma. 10, 306.
Golub, T.R., et al., 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286, 531–537.
Ideker, T., Krogan, N.J., 2012. Differential network biology. Mol. Syst. Biol. 8, 565.
Kayano, M., Takigawa, I., Shiga, M., Tsuda, K., Mamitsuka, H., 2011. ROS-DET: robust detector of switching mechanisms in gene expression. Nucleic Acids Res. 39, e74.
Kim, T.Y., Kim, H.U., Lee, S.Y., 2010. Data integration and analysis of biological networks. Curr. Opin. Biotechnol. 21, 78–84.
Kostka, D., Spang, R., 2004. Finding disease specific alterations in the co-expression of genes. Bioinformatics 20 (Suppl. 1), i194–i199.
Kusano, M., et al., 2007. Unbiased characterization of genotype-dependent metabolic regulations by metabolomic approach in *Arabidopsis thaliana*. BMC Syst. Biol. 1, 53.
Kusano, M., et al., 2011. Metabolomics reveals comprehensive reprogramming involving two independent metabolic responses of *Arabidopsis* to UV-B light. Plant J. 67, 354–369.
Lai, Y., Wu, B., Chen, L., Zhao, H., 2004. A statistical method for identifying differential gene–gene co-expression patterns. Bioinformatics 20, 3146–3155.
Langfelder, P., Horvath, S., 2007. Eigengene networks for studying the relationships between co-expression modules. BMC Syst. Biol. 1, 54.
Li, K.C., 2002. Genome-wide coexpression dynamics: theory and application. Proc. Natl. Acad. Sci. U. S. A. 99, 16875–16880.
Liu, B.H., Yu, H., Tu, K., Li, C., Li, Y.X., Li, Y.Y., 2010. DCGL: an R package for identifying differentially coexpressed genes and links from gene expression microarray data. Bioinformatics 26, 2637–2638.
Markowetz, F., Spang, R., 2007. Inferring cellular networks—a review. BMC Bioinforma. 8 (Suppl. 6), S5.
Morgenthal, K., Weckwerth, W., Steuer, R., 2006. Metabolomic networks in plants: transitions from pattern recognition to biological interpretation. Biosystems 83, 108–117.
Odibat, O., Reddy, C., 2012. Ranking differential hubs in gene co-expression networks. J. Bioinformatics Comput. Biol. 10, 1240002.
Pan, W., 2002. A comparative review of statistical methods for discovering differentially expressed genes in replicated microarray experiments. Bioinformatics 18, 546–554.
Ray, M., Zhang, W., 2010. Analysis of Alzheimer's disease severity across brain regions by topological analysis of gene co-expression networks. BMC Syst. Biol. 4, 136.
Reiner, A., Yekutieli, D., Benjamini, Y., 2003. Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics 19, 368–375.
Saito, K., Hirai, M.Y., Yonekura-Sakakibara, K., 2008. Decoding genes with coexpression networks and metabolomics — 'majority report by precogs'. Trends Plant Sci. 13, 36–43.
Stacklies, W., Redestig, H., Scholz, M., Walther, D., Selbig, J., 2007. pcaMethods—a bioconductor package providing PCA methods for incomplete data. Bioinformatics 23, 1164–1167.
Steuer, R., 2006. Review: on the analysis and interpretation of correlations in metabolomic data. Brief. Bioinform. 7, 151–158.
Strimmer, K., 2008. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. Bioinformatics 24, 1461–1462.
Tesson, B.M., Breitling, R., Jansen, R.C., 2010. DiffCoEx: a simple and sensitive method to find differentially coexpressed gene modules. BMC Bioinforma. 11, 497.
Usadel, B., et al., 2009. Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. Plant Cell Environ. 32, 1633–1651.
Valcarcel, B., et al., 2011. A differential network approach to exploring differences between biological states: an application to prediabetes. PLoS One 6, e24702.
van den Berg, R.A., Hoefsloot, H.C., Westerhuis, J.A., Smilde, A.K., van der Werf, M.J., 2006. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics 7, 142.
Wang, J., Li, M., Deng, Y., Pan, Y., 2010. Recent advances in clustering methods for protein interaction networks. BMC Genomics 11 (Suppl. 3), S10.
Watson, M., 2006. CoXpress: differential co-expression in gene expression data. BMC Bioinforma. 7, 509.
White, C.D., Brown, M.D., Sacks, D.B., 2009. IQGAPs in cancer: a family of scaffold proteins underlying tumorigenesis. FEBS Lett. 583, 1817–1824.
Yu, H., Liu, B.H., Ye, Z.Q., Li, C., Li, Y.X., Li, Y.Y., 2011. Link-based quantitative methods to identify differentially coexpressed genes and gene pairs. BMC Bioinforma. 12, 315.