ISSN: 2664-9640

# Application of bioweight gene co-expression network analysis in biomedicine

Xiaoqian Wu

Tianjin Polytechnic University, China

## Abstract

**Although high-throughput biomonitrometry has achieved excellent results in the field of biomedical application, it is still the focus of the medical industry at the present stage to discover valuable data in the system with more and more information. Especially in network biology, people have a new understanding of biological research. Therefore, on the basis of understanding the co-expression network analysis of weighted genes, this paper analyzes its application direction according to the current development trend of biomedicine, and makes in-depth exploration combining with specific cases, so as to lay a foundation for future biomedical research.**

## Keywords

**Weight gene co-expression network analysis; Biological; Medicine; Identification of cardiac remodeling; Myocardial infarction.**

## 1. Overview of co-expression network analysis of weighted genes

### 1.1. Define

WGCNA uses the expression correlation coefficients between molecules to evaluate the "co-expression" relationship between molecules. If it is in the same module, then the molecular expression will be similar and very different from other module molecules. At the same time, molecules with similar expression patterns are likely to participate in the same biological processes. Therefore, by analyzing the information correlation between modules and phenotypes, we can find the biological implications.[1-2]

### 1.2. The advantages and disadvantages

The advantage of this method is that it can preserve the continuity of network nodes and has strong analytical capability. Generally speaking, 15 samples are needed to choose this method. The disadvantage is that other data must be combined to obtain information on gene regulation, such as protein-protein interaction and methylation. Assuming that the information is composed of multiple modules, it is likely to reduce the authenticity of the analysis and affect the final identification results.[3-5]

### 1.3. Steps

According to the analysis in Figure 1, the specific steps are as follows: First, in order to build an effective gene co-expression network, data normalization should be done to ensure that the gene expression profile between samples has a comparative type; Secondly, the correlation matrix of gene expression was calculated carefully. This refers to the correlation coefficient between all genes, the specific formula is $S_{ij} = |cor(i, j)|$ , So the corresponding matrix is $S = \lfloor S_{ij} \rfloor$. Combined power exponent $A = \lfloor a_{ij} \rfloor$ The weights can be converted to $\Omega = \lfloor \omega_{ij} \rfloor$ ;Thirdly, A is transformed into topological matrix and has been applied in biomedicine. Among them, $1 - \omega_{ij}$ It can define the node dissimilarity and be used to detect the network module. The more

ISSN: 2664-9640

connected the computing module genes are, the more likely they are to become the key genes in the module. Fourth, the link between module genes and external information, such as clinical information, can find more key genes of biological significance.

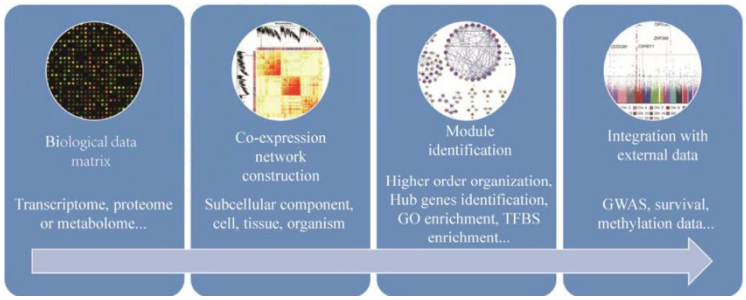


Figure 1. Flow chart of WGCNA

## 2. Analysis of the application of WGCNA in biomedicine

First, disease. Because the node molecules contained in the network module belong to important molecules in the module, it plays an active role in the research of disease category, prediction and mechanism, etc. For example, By using WGCNA and miRNA differential expression analysis, Wang et al. found that two differential expressions in human prostate cancer can regulate three important nodes related to cell cycle. Combined with experimental inquiry analysis, these two mirnas not only accelerate apoptosis, but also restrict growth. Thus, a cell cycle may be a major cause of prostate cancer.[6-8]

Second, normal tissue. In order to give full play to the functions of various human tissues, all kinds of cells, molecules and organelles must cooperate and coordinate with each other. Combined with WGCNA, the dimension of high-throughput omics data is reduced to 10 functional modules, and the relationship between the two can be studied to accurately break the relationship between individuals, cells and tissues.

Third, drugs. Drug understanding, as the basic content to solve human diseases, is also the core content of biomedical research. Using WGCNA, Iskar et al. studied data on rat liver and human cell lines treated with the drug. Seventy percent of the modules were shared by the cell lines, and 15 percent were conserved in both the rat and the human body. This discovery has changed the traditional understanding of the mechanism of action of drugs and provided a new basis for future medical research.

Fourth, evolution. In order to get a comprehensive understanding of the genetic differences between human and mice, Miller and others by using WGCNA research collection of thousands of cases of brain chip data, found that the cerebral gene expression between the two networks from the overall or biased towards the conservative, gene in mice module can be identified in human brain, and identification of specific module contains the brain microglia module of alzheimer's disease, also can undertake research in mice, the biomedical analysis brought a new way of thinking about the future.

Fifth, gene function. In most species, there are limits to the degree to which the genome can be annotated, especially for function. Through the understanding of multiple gene expression data sets, we can use WGCNA to build a network model, and we can find the hidden functions. Childs et al. identified a total of 71 modules in their study of the rice transcriptome data, and functionally annotated none of them. This study indicates that WGCNA has a positive role in gene function research and can provide functional data information for species with incomplete annotations. In addition, Walley et al. also studied the developmental gene changes of maize based on WGCNA and analyzed the proteome as the control group. Although the overlap rate of the two was not high, they were complementary to each other. Therefore, protein, mRNA and other data will affect the gene regulatory network.

ISSN: 2664-9640

# 3. Case study

## 3.1. Identify key node genes in cardiac remodeling

Combined with the analysis in figure 2 below, NCBI GEO was used to obtain the whole set of gene data GSE7487 and GSE738 of cardiac reconstruction after myocardial infarction. After preliminary processing, WGCNA was combined to rebuild the gene co-expression network, which could not only correctly identify the correlation between the data, but also the relationship between the key node genes and the two data. At the same time, the relationship between cardiac remodeling after infarction can be verified in rat models. In this study, WGCNA, as an efficient and excellent biomedical research method, could not only discover the key node genes of heart reconstruction, as shown in Figure 3 below, but also clarify the effect of RCAN1 on ACE1-ACF2 in renin angiotensin system.

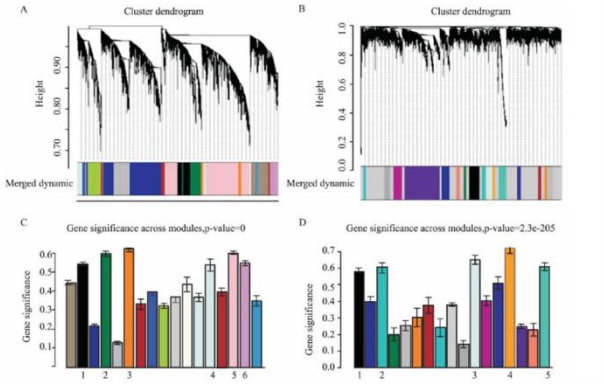


Figure 2. Co-expression network module of weighted genes in cardiac reconstruction after myocardial infarction

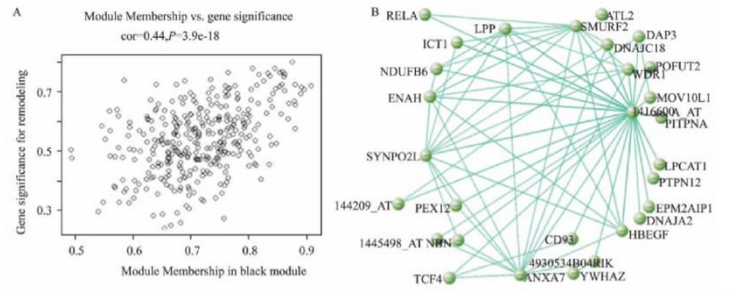


Figure 3. Key node genes in heart reconstruction

## 3.2. Myocardial infarction hub gene

Acute myocardial infarction occurs in human body because of rapid decrease or termination of coronary artery blood supply, resulting in acute myocardial ischemia, thereby causing myocardial necrosis. It is important to use WGCNA to study the gene at the heart of acute myocardial infarction (AMI), because it is not only a serious threat to human health, but also a growing mortality rate. The chip data related to it was found from the gene database and then combined with WGCNA to build the network model, as shown in Figure 3 below. On the basis of careful identification of the key modules of the disease state, gene screening and pathway analysis were carried out, and it was found that the regulation of inflammation and apoptosis were the main factors affecting myocardial infarction, and also the potential targets for future treatment of this disease.

# 4. Conclusion

To sum up, with the continuous improvement of China's economic detection technology, a large number of low-cost and high-efficiency application technologies have appeared on the market,

ISSN: 2664-9640

such as DWAS, high-throughput mass spectrometer, epigenetics, etc. However, with the continuous rise of genetic data, the data analysis methods previously proposed are unable to meet the higher analysis requirements. Because the biomedical research process is not a simple pathway or molecular addition, it is important to make full use of WGCNA technology in conjunction with the field trends and continue to optimize. On the one hand, this technique can simplify the data into multiple modules based on dimensionality reduction. On the other hand, it can analyse how multiple genes function in modules and parse more complex diseases. It can be seen that WGCNA will be widely used in biomedical research in the future. With the continuous improvement of technical level, It is believed that WGCNA will give full play to its application advantages and help biomedical researchers better solve medical problems.

## References

[1] Li Y , Zheng J N , Wang E H , et al. Application of weighted gene co-expression network analysis to reveal key modules and hub genes in generalized aggressive periodontitis[J]. Archives of oral biology, 2020, 119:104895.

[2] Yan, Sun, Jinghan, et al. The application of weighted gene co-expression network analysis in identifying key modules and hub genes associated with disease status in Alzheimer's disease.[J]. Annals of translational medicine, 2019, 7(24):800-800.

[3] Liang J W , Fang Z Y , Huang Y , et al. Application of Weighted Gene Co-Expression Network Analysis to Explore the Key Genes in Alzheimer's Disease[J]. IOS Press Open Library, 2018, 65(4).

[4] Liang J W , Fang Z Y , Huang Y , et al. Application of Weighted Gene Co-Expression Network Analysis to Explore the Key Genes in Alzheimer's Disease[J]. Journal of Alzheimer's disease: JAD, 2018, 65(Suppl 4):1-12.

[5] Yang J , Li C , Zhou J , et al. Identification of Prognostic Genes in Leiomyosarcoma by Gene Co-Expression Network Analysis[J]. Frontiers in Genetics, 2020, 10.

[6] Leonel N. Leal, Josue M. Romao, Guido J. Hooiveld , and so on . Nutrient supply alters transcriptome regulation in adipose tissue of pre-weaning Holstein calves[J]. Biomedicine & Pharmacotherapy, 2018, 106(8):134-141.

[7] Rahimi R , Karimi J , Khodadadi I , et al. Silymarin ameliorates expression of urotensin II (U-II) and its receptor (UTR) and attenuates toxic oxidative stress in the heart of rats with type 2 diabetes[J]. Biomedicine & Pharmacotherapy, 2018, 101:244–250.

[8] Wang J , Xu J , Li K , et al. Identification of WTAP-related genes by weighted gene co-expression network analysis in ovarian cancer[J]. Journal of Ovarian Research, 2020, 13(1):119.