Potential Pathogenic Genes and Mechanism of Ankylosing Spondylitis: A Study Based on WGCNA and Bioinformatics Analysis

Bo Wu

Jilin University First Hospital

Jing Yu

Jilin University First Hospital

Yibing Liu

Jilin University

Gaojing Dou

Jilin University First Hospital

Zhiyun Zhang

Jilin University

Xuepeng Pan

Jilin University First Hospital

Hongyu Wang

Jilin University

Pengcheng Zhou

Jilin University First Hospital

Dong Zhu (✉ zhu_dong@mail.jlu.edu.cn)

Jilin University First Hospital <https://orcid.org/0000-0001-9422-2103>

Research

Keywords: Ankylosing spondylitis, bioinformatics, WGCNA, pathogenic genes, mechanism

Posted Date: May 4th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-438711/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at World Neurosurgery on November 1st, 2021. See the published version at <https://doi.org/10.1016/j.wneu.2021.11.014>.

Abstract

Ankylosing spondylitis (AS) is immunologic arthritis that causes high disability and low quality of life. The pathogenesis of mandatory spondylitis is uncertain. The purpose of this study is to explore the high-risk pathogenic driver genes for the occurrence and development of AS based on the bioinformatics method at the molecular level, further elaborate the molecular mechanism of the pathogenesis of AS, and provide potential biological targets for the diagnosis and treatment of clinical AS. The Gene expression profile data GSE16879 was downloaded from the GEO database, and weighted gene co-expression network analysis (WGCNA) was performed. Highly correlated genes were divided into 14 modules, and 582 genes contained in the yellow (Classical Module) and 59 genes contained in grey60 (Hematological Module) modules had the strongest correlation with AS. After Protein-protein interaction (PPI) analysis, the top 20 genes with the highest scores were obtained from Classical Module and Hematological Module respectively. DAVID database was used for GO analysis and KEGG analysis to analyze the biological functions of high-risk genes related to AS. The results showed that the process of SRP-dependent cotranslational protein targeting to membrane, ribosome, NADH dehydrogenase (ubiquinone) activity, platelet activation, integrin complex, and extracellular matrix binding were enriched. In this paper, WGCNA, an efficient system biology algorithm, was used to analyze the high-risk pathogenic driver gene of ankylosing spondylitis. We provide new targets for the diagnosis and treatment of clinical AS and new ideas for further study.

Introduction

Ankylosing spondylitis (AS) is an insidious, progressive, and debilitating form of autoimmune arthritis with its primary target being the axial skeleton [1]. It can cause characteristic inflammatory, back pain, and reduce spinal mobility [2], leading to structural and functional impairments and a decrease in quality of life [3]. It is reported that in the United States, such a disease is more likely to happen in men before the age of forty, with a prevalence of about 0.5% [4]. Besides, early symptoms of AS are often not apparent, and most of the patients can't be diagnosed at the early stage. That usually leads to a delay in AS treatment and loss of the best time for treatment.

The current treatment for AS is often through drugs. Several medicines are applied to stop or delay patients' spinal problems and relieve pain. NSAIDs, which usually is the first drug given to patients, can stop the body from making inflammation-causing chemicals called prostaglandins. Corticosteroids, TNF inhibitors, and Interleukin inhibitors (IL-17 inhibitors) are also available to relieve the suffering of patients. Currently, the prognosis of AS has an individual difference. There is no radical cure at present. We can only control the disease progression through targeted treatment.

AS is characterized by a complex pathological process. The basic pathology is primary, chronic, pannus destructive inflammation, and Ligament ossification is a secondary repair process. The lesions typically begin in the sacroiliac joint and slowly travel up the spine, involving the synovium and capsule of the intervertebral facet joints and the soft tissue around the spine. At the late stage, it can make the soft tissue around the entire spine calcify, ossify, and lead to severe hunchback. The disease can also spread downwards at the same time, affecting both hips and, in a few cases, knees. Most patients have a gene that can make a protein called HLA-B27, which is widely believed to have a strong connection with AS. Currently, only one genetic biomarker for diagnosing ankylosing spondylitis — HLA-B27 was reported. The rest biomarkers are still not reported [5], and the uses of the resulting biomarkers for AS are widespread. Also, some researchers have suggested that eukaryotic translation elongation factor 1 epsilon 1 (EEF1E1) may be a potential genetic biomarker for diagnosing AS. In patients with AS, the expression levels of these genes were significantly up-regulated, and the study showed that EEF1E1 participates in AS by influencing the aminoacyl-tRNA biosynthesis [5]. Another research suggested that another two genes named MAPK7 and NDUFS4 played important roles in the pathogenesis of AS via the cAMP-mediated signaling pathway. Both of them could be targeted by Indomethacin [6]. Nevertheless, the study of AS is relatively superficial, and its mechanism is still unclear. So, we attempt to find new biomarkers and increase the mechanistic understanding by using an analysis of gene expression microarray profiling, providing early prevention and management of AS [7].

Weighted gene co-expression network analysis (WGCNA) is a new system biology method and is increasingly used in bioinformatics for analyzing gene expression chip mapping data [8-10]. By constructing a gene co-expression network, highly correlated genes were clustered into several modules. When a module is associated with external information, biologically interesting modules are detected. Key genes that play a key role in disease development have been identified in interesting modules related to important biological functions or pathways. Also, WGCNA can be used to screen candidate biomarkers or therapeutic targets. Therefore, we conducted this study to look for new biomarkers, related genes, or potential mechanisms related to AS.

Materials And Methods

Data Collection and Preprocessing

The gene expression profile (GSE25101) of AS was downloaded from the Gene Expression Omnibus (GEO) database. GSE25101 contains 16 normal samples and 16 AS samples [45].

Co-expression Network Construction

The "WGCNA" package [10] in R software (version 4.0.0) was used for the network construction. Samples conforming to $Z.K < 2.5$ were considered peripheral and deleted from the expression and trait data. Pearson correlation coefficients were calculated for all gene comparisons, and the soft thresholding power was evaluated. Then, the weighted network adjacency matrix was calculated based on $a_{ij} = |\text{cor}(x_i, x_j)|^c$. x_i and x_j are nodes i and j of the network, which is determined by scale-free topology criteria [11]. The adjacency matrix was transformed into a topological overlap matrix to identify highly interconnected gene modules and clusters. Topological overlap measure (TOM) was used to determine the interconnectivity of the network. Gene module detection USES hierarchical clustering method was based on different measures (1-TOM) [12]. The dynamic branch cutting method was used to identify the branches of the hierarchical cluster tree [13].

Identification and functional annotation of clinically important modules

Modules closely related to biological or clinical information were selected from hierarchical clustering as interesting modules for subsequent analysis. In the process, gene significance (GS), module significance (MS), and module eigengene (ME) was calculated. GS was defined as the negative logarithm of the p-value, and MS was the average genetic meaning of the entire module gene. ME was the first principal component of a given module. Significance between ME and clinical traits was also calculated through a correlation analysis because modules with high salience were path-related and could be candidates [14-15]. Gene ontology (GO) functional annotation and Kyoto Encyclopedia of Genes and Genomics (KEGG) enrichment by Annotation, Visualization, and Integrated discovery database [16] were performed to further validate the selected modules by looking for potential mechanisms and biological pathways.

Identification and validation of hub genes

The hub genes of interesting modules were determined by the absolute value of the members of the gene module >0.9 and the significance of the gene characteristics >0.22 [17]. Then, a PPI (protein-protein interaction) network of DEGs was constructed through the Search Tool for the Retrieval of Interacting Genes (STRING) database. A medium confidence (0.4) was applied to filter the interactions. Detection was performed based on Cytoscape (version 3.4.0) software to reveal modules of the PPI network [18-19]. In addition, the miRNA targeting those hub genes were predicted by FunRich software (version 3.1.3). The miRNA-mRNA interaction networks were also plotted using Cytoscape.

Results

Calculate soft threshold, construct co-expression matrix and partition module

According to the distribution of the scale-free network, similarity and anisotropy coefficients among genes were calculated, and the cluster tree of the system between genes was constructed. Using the function of the WGCNA software package in R software, the optimal soft threshold = 8 is calculated to divide the co-expression module. After the soft threshold was determined, the dynamic shear tree method was used to preliminarily identify and merge similar modules. Cluster analysis was carried out on the modules, and the modules close to each other were merged into new modules. The minimum number of genes in each gene co-expression network module was set as 30, and 14 modules were obtained, among which the gray module was the gene set that could not be aggregated to other modules (Figure 1).

Screening of high-risk pathogenic gene modules for AS

The WGCNA software package was used to calculate the correlation between these modules and AS according to the feature vectors of each module, to build the cluster tree of each module and disease phenotype, and to calculate the Pearson correlation coefficient between different modules and AS. The results showed that 582 genes contained in the yellow (Classical Module) ($r=0.43$, $P=1.4e-27$) and 59 genes contained in grey60 (Hematological Module) ($r=0.2$, $P=0.13$) modules had the strongest correlation with AS. See Figure 2 for the results.

Screening of hub genes for AS

The genes in the Classical Module were selected and analyzed by 11 different methods. The top 20 genes with the highest scores were obtained, including LOC653773, MRPL32, DPM1, DPY30, MRPL1, MRPS33, CWC15, RWDD1, COMMD6, LSM1, CETN3, SNRPG, C15orf15, ITGB3BP, HINT1, NDUFB2, LSM5, RPL23, TMEM126B, NDUFA4. AS shown in Table 1. In the same way, the top 20 genes with the highest scores in the Hematological Module were obtained, including TSPAN9, MGLL, ABLIM3, ITGB3, ITGB5, SH3BGRL2, TREML1, SAMD14, CTTN, NAT8B, C6orf21, RBPMS2, ACRBP, GUCY1A3, AQP10, CDKN1A, GP9, ESAM, Septin 5, MYL9. AS shown in Table 2.

Biological function notes for AS hub genes

The genes of the Classical Module and Hematological Module were uploaded to DAVID's website for GO analysis and KEGG analysis to explore the biological functions of differentially expressed genes. The GO analysis results showed that the process of SRP-dependent cotranslational protein targeting to membrane, ribosome, and NADH dehydrogenase (ubiquinone) activity in the Classical Module was significantly abnormal, and the process of platelet activation, integrin complex, and extracellular matrix binding in Hematological Module was abnormal. And the KEGG analysis results showed that genes in the Classical Module are associated with Parkinson's disease, Huntington's disease, Alzheimer's disease; while genes in the Hematological Module have an association with Platelet activation, Tight junction, and ECM-receptor interaction. As shown in Figure 3, 4 and Table 3, 4.

miRNA-mRNA interaction networks

The miRNA analysis was conducted by FunRich software. There are 38 miRNAs predicted targeting genes in the Classical Module, including hsa-miR-22-3p, hsa-miR-32-5p, hsa-miR-320c and et. al. A total of 64 miRNAs were identified targeting Hematological Modules, including hsa-let-7b-5p, hsa-let-7c-5p, hsa-let-7e-5p, and et. al. (Supplement table 1 and 2) The miRNA-mRNA interaction network of genes in selected modules also were plotted (Figure 5A and Figure 5B).

Discussion

Ankylosing spondylitis is a chronic nonspecific inflammation caused by a variety of factors. Due to its high rate of disability, it seriously affects patients' quality of life [20]. The pathogenesis of AS is complex, involving genetic factors, environmental factors, immune response, and other aspects. No specific molecular mechanism has been reported so far. The therapeutic effect is not satisfactory and the economic burden on patients is increased [21]. In this study, the WGCNA algorithm was used to explore the molecular mechanism of AS and provide a basis for new therapeutic strategies

In this study, after constructing a co-expression matrix and Partition Module, and screening of high-risk pathogenic gene modules for AS, 582 genes in the Classical Module and 59 genes in the Hematological Module were identified related to AS. GO and KEGG analyses were performed to gain further insights into the molecular mechanism of the AS.

The GO and KEGG analysis results showed that the process of SRP-dependent cotranslational protein targeting to membrane, ribosome, and NADH dehydrogenase (ubiquinone) activity was significantly related to the genes of the Classical Module, while the genes in the Hematological Module enriched in the process of platelet activation, integrin complex, and leukocyte migration. The SRP pathway is closely related to the level of membrane protein and secreted protein in eukaryotic cells [22], which may promote AS by regulating the secretion of inflammatory mediators and the occurrence of inflammation. According to a recent report, ubiquinone is not only a mobile component of the mitochondrial electron transport chain but also a membrane antioxidant [23]. Besides, it was found that the reaction products of the antioxidant activity of ubiquinol stimulated the formation of inorganic and organic oxygen radicals [24], which can lead to inflammation. The ribosome was reported that anti-ribosomal P protein IgG can be a diagnostic indicator for systemic lupus erythematosus and rheumatoid arthritis (autoimmune diseases) [25]. That means anti-ribosomal P protein IgG may also be a potential diagnostic indicator for AS. At present, platelet activation was pointed as a sign of worsened AS [26], which supports the results of this study. Besides, some drugs have been shown to relieve the symptoms of ankylosing spondylitis by inhibiting platelet activation [27]. That suggests a novel treatment strategy for AS. For the integrin complex, as the key protein of the integrin signaling pathway, the ILK protein was identified with higher expression levels in AS patients than in healthy people [28], which may be a potential marker of AS.

To obtain the hub genes among the 582 genes contained in the Classical Module and 59 genes contained in the Hematological Module, they were analyzed by using the PPI network according to the STRING database. Finally, we got the top20 hub genes in each module, including DPY30, MRPS33, RPL23 in the Classical Module, and TSPAN9, TRELM1, CDKN1A in the Hematological Module. In addition, the miRNA-mRNA interaction network was plotted and 38 miRNAs were predicted targeting genes in Classical Module, and 64 miRNAs were identified targeting Hematological Module.

DPY30, the core subunit of the SET1/MLL family histone H3K4 methyltransferase complex, can affect glucose homeostasis [29]. DPY30 was reported to control proliferation by regulating the inhibitor of differentiation (ID) protein expression, leading to reactive oxygen species (ROS) levels rise and aging bypasses [30]. Cell aging and increased ROS levels may be one reason for inflammation in AS. MRPS33, one of the genes that encode the small subunit of the mitochondrial ribosome, is reported to show hypermethylation in older muscle tissue [31], which leads to the loss of muscle mass and strength in the elderly. That may support the myophagism in AS patients. Besides, the adult flies with the MRPS33 gene knocked out are more likely to develop cardiomyopathy [32], and the symptoms of cardiovascular disease also appeared as one of the accompanying symptoms of AS patients [33]. RPL23, a negative regulator of cellular apoptosis, is proved that it can encode autoreactive proteins, reacting with T cells and autoantibodies [34]. This reaction may be one of the components of the inflammatory response in patients with AS. Besides, the development of AS occurs due to excessive proliferation of fibroblasts [35], while RPL23 negatively regulates cellular apoptosis via the RPL23/Miz-1/c-Myc circuit [36], which may expose the mechanism of the development of AS.

TSPAN9, a gene associated with platelet activation [37], can be used as a target for platelet activation in patients with AS, and platelet activation plays a role in heart disease secondary to AS [39]. TSPAN9 has also been linked to decreased muscle tone and joint relaxation [40], which also associates with AS. TRELM1, a gene associated with platelet dysfunction, can also lead to heart disease secondary to AS [39]. Moreover, TRELM1 binds to fibrinogen and plays a role in bleeding caused by inflammatory damage. CDKN1A, cyclin-dependent kinase inhibitors involved in cell cycle regulation, can regulate the proliferation of skeletal muscle cells. Besides, the CDKN1A gene is an inflammatory response gene in the central nervous system, which may be related to the nervous system secondary to AS.

There are several highlights in the study. WGCNA, an efficient systems biology algorithm, was used for the first time to analyze the drivers of AS. Secondly, WGCNA has a unique advantage in dealing with gene expression datasets. The results of this study not only confirmed the results of previous studies but also provided new biomarkers for future studies of AS. However, this study also has some limitations and needs independent cohort and/or in-vitro experiments for validation. Whether these biomarkers can be applied in clinical practice remains to be further verified.

Conclusion

In this study, WGCNA, an efficient system biology algorithm, was used to analyze the high-risk pathogenic driver genes of ankylosing spondylitis, further elucidating the molecular mechanism of the occurrence and development of AS, and providing candidate biomarkers for the clinical diagnosis and treatment of AS, providing a new idea for the further study of ankylosing spondylitis.

Abbreviations

Ankylosing Spondylitis , AS;

Weighted Gene Co-expression Network Analysis, WGCNA;

Gene Ontology, GO;

Kyoto Encyclopedia of Genes and Genomes, KEGG;

Protein-protein Interaction, PPI;

Database for Annotation, Visualization, and Integrated Discovery, DAVID;

Gene Expression Omnibus, GEO;

Signal Recognition Particle, SRP;

Nicotinamide Adenine Diphosphate Hydride, NADH;

Non-steroidal Anti-inflammatory drugs, NSAIDs;

Tumor Necrosis Factor, TNF;

Interleukin, IL;

Human Leukocyte Antigen-B27, HLA-B27;

Eukaryotic Translation Elongation Factor 1 Epsilon 1, EEF1E1;

Mitogen Activated Protein Kinase 7, MAPK7;

NADH Dehydrogenase-Ubiquinone-FeS 4, NDUFS4;

Extracellular Matrix-receptor, ECM-receptor;

Integrin-Linked Kinase, ILK;

Search Tool for the Retrieval of Interacting Genes, STRING;

Inhibitor of Differentiation, ID;

Reactive Oxygen Species, ROS.

Declarations

Ethical Approval: Not applicable

Consent to Participate: Not applicable

Consent to Publish: consent

Availability of data and material: consent

Competing Interests:

There is no competing interests.

Funding:

The present study was supported by grants from National Natural Science Foundation of China (grant number12072129).

Author contributions

Bo Wu, Sheng Zhong and Wenzhuo Yang designed experiments; Zhiyun Zhang wrote manuscript; Xiaye Lv, Gaojing Dou, Xinhui Wang and Junliang Ge carried out experiments; Xuefeng Pan and Hongyu Wang analyzed experiments results.

Acknowledgements:Not applicable

References

1. Smith JA. Update on ankylosing spondylitis: current concepts in pathogenesis. *Curr Allergy Asthma Rep*, 15(1): p. 489. doi:10.1007/s11882-014-0489-6.
2. Golder V, Schachna L. Ankylosing spondylitis: an update. *Aust Fam Physician*, 42(11): p. 780–784.
3. Braun J, Sieper J. Ankylosing spondylitis. *Lancet*, 369(9570): p. 1379–1390.
4. Reveille JD, Weisman MH. The epidemiology of back pain, axial spondyloarthritis, and HLA-B27 in the United States. *Am J Med Sci*, 345(6): p. 431–6.
5. Fan X, Qi B, Ma L, Ma F. Screening of underlying genetic biomarkers for ankylosing spondylitis. *Mol Med Rep*, 19(6): p. 5263–5274. doi:10.3892/mmr.2019.10188.
6. Ni Y, Jiang C. Identification of potential target genes for ankylosing spondylitis treatment. *Medicine (Baltimore)*, 97(8): p. e9760.
7. Vukusic K, Jonsson M, Brantsing C, Dellgren G, Jeppsson A, Lindahl A, et al. High density sphere culture of adult cardiac cells increases the levels of cardiac and progenitor markers and shows signs of vasculogenesis. *Biomed Res*, 2013: p. 696837.
8. Stuart JM, Segal E, Koller D, Kim SK. A gene-coexpression network for global discovery of conserved genetic modules. *Science*. 2003;302(5643):249–55.

9. Zhang B, Horvath S. A general framework for weighted gene coexpression network analysis. *Stat. Appl. Genet. Mol. Biol.* 4: p. Article17.
10. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.
11. Zhang B, Horvath S. A general framework for weighted gene coexpression network analysis. *Stat Appl Genet Mol Biol.* 2005;4:Article17.
12. Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabasi AL. Hierarchical organization of modularity in metabolic networks. *Science.* 2002;297(5586):1551–5.
13. Langfelder P, Zhang B, Horvath S (2008). Defining clusters from a hierarchical cluster tree: the dynamic tree cut package for R. *Bioinformatics.* 24: p. 719–720.
14. Ghazalpour A, Doss S, Zhang B, Wang S, Plaisier C, Castellanos R, et al. Integrating genetic and network analysis to characterize genes related to mouse weight. *PLoS Genetics*, 2(8): p. e130.
15. Fuller TF, Ghazalpour A, Aten JE, Drake TA, Lusis AJ, Horvath S. Weighted gene coexpression network analysis strategies applied to mouse weight. *Mamm Genome.* 2007;18:463–72.
16. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* 2003;4:3.
17. Horvath S, Dong J. Geometric interpretation of gene coexpression network analysis. *PLoS Comput. Biol.* 4: p. e1000117.
18. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, HuertaCepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43: p. D447–D452.
19. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13:2498–504.
20. Kotsis K, Voulgari PV, Drosos AA, Carvalho AF, Hyphantis T. Health-related quality of life in patients with ankylosing spondylitis: a comprehensive review. *Expert Rev Pharmacoecon Outcomes Res*, 14(6): p. 857–872.
21. Smith JA. Update on ankylosing spondylitis: current concepts in pathogenesis. *Curr Allergy Asthma Rep*, 15(1): p. 489.
22. Cui YY, Zhang SB, Xie NZ, Zhang TC, Zhang DW. Signal recognition particles regulate protein transport system [J]. *Chinese journal of microbiology*, 43(06): p. 1358–1365.
23. Wang Y, Hekimi S. Understanding Ubiquinone. *Trends Cell Biol.* 2016 May, 26(5): p.367–378.
24. Nohl H, Gille L, Staniek K. The biochemical, pathophysiological, and medical aspects of ubiquinone function. *Ann N Y Acad Sci.* 854: p. 394–409.
25. Carmona-Fernandes D, Santos MJ, Canhão H, Fonseca JE. Anti-ribosomal P protein IgG autoantibodies in patients with systemic lupus erythematosus: diagnostic performance and clinical profile. *BMC Med.* 11: p. 98.
26. Wang F, Yan CG, Xiang HY, Xing T, Wang NS. The significance of platelet activation in ankylosing spondylitis. *Clin Rheumatol*, 27(6): p. 767–9.
27. Fang YY, Wan L, Dong WZ, Wen JT, Liu J. [Effect of triptolide in improving platelet activation in patients with ankylosing spondylitis by regulating VEGFA,SDF-1,CXCR4 pathway]. *Zhongguo Zhong Yao Za Zhi*, 44(16): p. 3520–3525.
28. Li T, Zheng B, Huang Z, Lu H, Lin Q, Liao Z, Lin Z, Zhao L, Wang X, Gu J. Over-expression of talin 1 and integrin-linked kinase in PBMCs of patients with ankylosing spondylitis: a proteomic study. *Clin Exp Rheumatol*, 28(6): p. 828–35.
29. Liu B. DPY30 functions in glucose homeostasis via integrating activated histone epigenetic modifications. *Biochem Biophys Res Commun*, 507(1–4): p. 286–290.
30. Simboeck E, Gutierrez A, Cozzuto L, Beringer M, Caizzi L, Keyes WM, Di Croce L. DPY30 regulates pathways in cellular senescence through ID protein expression. *EMBO J*, 32(16): p. 2217–30.
31. Turner DC, Gorski PP, Maasar MF, Seaborne RA, Baumert P, Brown AD, Kitchen MO, Erskine RM, Dos-Remedios I, Voisin S, Eynon N, Sultanov RI, Borisov OV, Larin AK, Semenova EA, Popov DV, Generozov EV, Stewart CE, Drust B, Owens DJ, Ahmetov II, Sharples AP. DNA methylation across the genome in aged human skeletal muscle tissue and muscle-derived cells: the role of HOX genes and physical activity. *Sci Rep*, 10(1): p. 15360.
32. Casad ME, Abraham D, Kim IM, Frangakis S, Dong B, Lin N, Wolf MJ, Rockman HA. Cardiomyopathy is associated with ribosomal protein gene haplo-insufficiency in *Drosophila melanogaster*. *Genetics*, 189(3): p. 861–70.
33. Han XH, Lu LQ. Ankyloss spondylitis [J]. *Chinese Journal of Tissue Engineering*, 6(001): p. 22–24.
34. Ito Y, Hashimoto M, Hirota K, Ohkura N, Morikawa H, Nishikawa H, Tanaka A, Furu M, Ito H, Fujii T, Nomura T, Yamazaki S, Morita A, Vignali DA, Kappler JW, Matsuda S, Mimori T, Sakaguchi N, Sakaguchi S. Detection of T cell responses to a ubiquitous cellular protein in autoimmune disease. *Science*, 346(6207): p.363–8.
35. Ma C, Wen B, Zhang Q, Shao P, Gu W, Qu K, Shi Y, Wang B. Polydatin Regulates the Apoptosis and Autophagy of Fibroblasts Obtained from Patients with Ankylosing Spondylitis. *Biol Pharm Bull*, 42(1): p.50–56.
36. Qi Y, Li X, Chang C, Xu F, He Q, Zhao Y, Wu L. Ribosomal protein L23 negatively regulates cellular apoptosis via the RPL23/Miz-1/c-Myc circuit in higher-risk myelodysplastic syndrome. *Sci Rep*, 7(1): p. 2323.
37. Prottly MB, Watkins NA, Colombo D, Thomas SG, Heath VL, Herbert JM, Bicknell R, Senis YA, Ashman LK, Berditchevski F, Ouwehand WH, Watson SP, Tomlinson MG. Identification of Tspan9 as a novel platelet tetraspanin and the collagen receptor GPVI as a component of tetraspanin microdomains. *Biochem J*, 417(1): p. 391–400.
38. Fang YY, Wan L, Dong WZ, Wen JT, Liu J. Effect of triptolide in improving platelet activation in patients with ankylosing spondylitis by regulating VEGFA, SDF-1, CXCR4 pathway. *Zhongguo Zhong Yao Za Zhi*, 44(16): p. 3520–3525.

39. Bergfeldt L, Edhag O, Holm G, Norberg R. Platelet aggregating activity in serum from patients with HLA-B27 associated rheumatic and cardiac disorders: a possible link to the proliferative vascular changes. *Br Heart J*, 65(4): p. 184–7.

40. Fanizza I, Bertuzzo S, Beri S, Scalera E, Massagli A, Sali ME, Giorda R, Bonaglia MC. Genotype-phenotype relationship in a child with 2.3 Mb de novo interstitial 12p13.33-p13.32 deletion. *Eur J Med Genet*, 57(7): p. 334–8.

Tables

Table 1. The top 20 genes with the highest scores in the Classical Module after 11 different methods

Gene symbol	MCC	DMNC	MNC	Degree	EPC	BottleNeck	EcCentricity	Closeness	Radiality	Betweenness	Stress	ClusteringCoef
LOC653773	901	0.15	91	96	150	3	0.25	318	4.03	9519	1E+05	0.07
MRPL32	691	0.15	80	81	146	2	0.25	308	3.97	5146	60076	0.08
DPM1	680	0.19	66	66	141	1	0.25	294	3.88	1852	26334	0.11
DPY30	638	0.15	77	79	144	4	0.25	308	3.98	4180	50988	0.08
MRPL1	621	0.18	65	66	137	2	0.25	299	3.93	2368	32676	0.1
MRPS33	574	0.18	62	64	138	1	0.25	299	3.93	2756	36566	0.1
CWC15	569	0.18	64	65	136	2	0.25	293	3.87	1908	24736	0.1
RWDD1	539	0.15	75	80	138	1	0.25	306	3.95	5228	60242	0.07
COMMD6	534	0.16	67	67	136	1	0.25	300	3.93	2940	38910	0.09
LSM1	512	0.15	69	71	139	2	0.25	304	3.96	3356	43170	0.08
CETN3	476	0.2	56	56	124	1	0.25	293	3.9	1620	25494	0.12
SNRPG	461	0.15	67	68	138	2	0.25	300	3.93	3037	40036	0.08
C15orf15	457	0.14	68	69	135	3	0.25	300	3.93	3043	38272	0.08
ITGB3BP	439	0.19	55	58	131	6	0.25	295	3.92	1802	26086	0.1
HINT1	428	0.12	75	79	136	3	0.25	308	3.98	7418	76634	0.06
NDUFB2	423	0.16	60	61	131	4	0.25	294	3.89	2256	30004	0.09
LSM5	421	0.14	65	66	132	1	0.25	303	3.98	4033	50380	0.08
RPL23	420	0.15	65	65	133	1	0.25	301	3.95	3010	39272	0.09
TMEM126B	404	0.14	66	70	133	1	0.25	298	3.91	3258	38920	0.07
NDUFA4	403	0.12	71	74	137	1	0.33	307	3.99	4723	54542	0.06

Table 2. The top 20 genes with the highest scores in the Hematological Module after 11 different methods

Gene Symbol	MCC	DMNC	MNC	Degree	EPC	BottleNeck	EcCentricity	Closeness	Radiality	Betweenness	Stress	ClusteringCc
TSPAN9	13	0.3241	5	8	13.261	14	0.1862	18.833	4.5119	267.17	468	0.1786
MGLL	11	0.3241	5	6	12.678	6	0.1862	17	4.3367	104.43	270	0.3333
ABLIM3	9	0.4635	3	6	12.407	18	0.1862	17.583	4.4243	322.13	530	0.2
ITGB3	9	0.3789	4	5	11.766	2	0.1862	15.75	4.2053	74.333	130	0.4
ITGB5	8	0.309	3	7	12.321	6	0.1489	17.617	4.3586	241.77	430	0.0952
SH3BGRL2	8	0.2842	4	6	11.776	3	0.1489	16.45	4.2053	155.53	338	0.2
TREML1	6	0	1	6	9.271	5	0.1862	16	4.1834	265.57	392	0
SAMD14	5	0	1	5	8.946	8	0.1862	15.5	4.1615	179.47	250	0
CTTN	4	0.3078	2	4	11.422	1	0.1489	14.733	4.03	46.7	118	0.1667
NAT8B	4	0	1	4	10.458	4	0.1489	15.533	4.1834	104.87	174	0
C6orf21	4	0	1	4	10.158	4	0.1489	14.483	4.0081	93.8	174	0
RBPMS2	4	0	1	4	9.728	2	0.1862	14.667	4.0738	93.833	178	0
ACRBP	3	0	1	3	9.392	13	0.1489	15.367	4.2272	151.73	224	0
GUCY1A3	3	0	1	3	6.715	5	0.1489	13.017	3.7891	81.833	138	0
AQP10	3	0	1	3	4.919	2	0.1241	11.5	3.4387	90.733	140	0
CDKN1A	2	0.3078	2	2	8.49	1	0.1241	12.783	3.811	0	0	1
GP9	2	0	1	2	8.056	1	0.1241	13.033	3.8767	6	22	0
ESAM	2	0	1	2	7.717	2	0.1489	13.117	3.9205	25.833	38	0
Septin 5	2	0	1	2	6.959	1	0.1489	12.583	3.7672	36.133	76	0
MYL9	2	0	1	2	6.739	3	0.1489	12.9	3.8767	113.73	174	0

Table 3. Gene functions in the Classical Module after GO and KEGG analysis

Term	Count	%	PValue
GO:0006614~SRP-dependent cotranslational protein targeting to membrane	54	13.30049261	2.44E-63
GO:0006412~translation	74	18.22660099	4.04E-61
GO:0019083~viral transcription	52	12.80788177	1.45E-54
GO:0000184~nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	52	12.80788177	7.30E-53
GO:0006413~translational initiation	53	13.05418719	1.55E-50
GO:0005840~ribosome	64	15.7635468	3.03E-62
GO:0022625~cytosolic large ribosomal subunit	34	8.374384236	3.08E-37
GO:0005743~mitochondrial inner membrane	63	15.51724138	6.93E-33
GO:0022627~cytosolic small ribosomal subunit	22	5.418719212	6.88E-23
GO:0005925~focal adhesion	40	9.852216749	1.25E-15
GO:0003735~structural constituent of ribosome	73	17.98029557	1.81E-64
GO:0044822~poly(A) RNA binding	102	25.12315271	3.56E-35
GO:0003723~RNA binding	53	13.05418719	4.48E-19
GO:0005515~protein binding	256	63.05418719	2.53E-10
GO:0008137~NADH dehydrogenase (ubiquinone) activity	13	3.201970443	4.24E-10
hsa03010:Ribosome	66	16.25615764	9.05E-63
hsa00190:Oxidative phosphorylation	27	6.650246305	7.18E-14
hsa05012:Parkinson's disease	24	5.911330049	1.24E-10
hsa05016:Huntington's disease	26	6.403940887	2.22E-09
hsa05010:Alzheimer's disease	24	5.911330049	3.80E-09

Table 4. Gene functions in the Hematological Module after GO and KEGG analysis

Term	Count	%	PValue
GO:0030168~platelet activation	6	13.95349	5.05E-06
GO:0070527~platelet aggregation	4	9.302326	9.91E-05
GO:0050900~leukocyte migration	5	11.62791	1.46E-04
GO:0030198~extracellular matrix organization	5	11.62791	8.79E-04
GO:0007160~cell-matrix adhesion	4	9.302326	0.001014
GO:0008305~integrin complex	3	6.976744	0.001673
GO:0005925~focal adhesion	6	13.95349	0.001759
GO:0005886~plasma membrane	19	44.18605	0.001986
GO:0005887~integral component of plasma membrane	10	23.25581	0.00361
GO:0009986~cell surface	6	13.95349	0.007076
GO:0005178~integrin binding	3	6.976744	0.018888
GO:0050840~extracellular matrix binding	2	4.651163	0.051106
GO:0004872~receptor activity	3	6.976744	0.070553
hsa04611: Platelet activation	6	13.95349	5.71E-05
hsa04530: Tight junction	5	11.62791	1.76E-04
hsa04512: ECM-receptor interaction	5	11.62791	1.76E-04
hsa04510: Focal adhesion	6	13.95349	4.97E-04
hsa05205: Proteoglycans in cancer	5	11.62791	0.003978

Figures

Figure 1

(A) Analysis of network topology for various soft-thresholding powers. The left panel shows the scale-free fit index, signed R^2 (y-axis), and the soft threshold power (x-axis). $\beta = 8$ was chosen for subsequent analysis. The right panel shows the mean connectivity (y-axis) is a strictly decreasing function of the power β (x-axis). (B) Clustering dendrogram of genes. The color bands provide a simple visual comparison of module assignments (branch cuttings) based on the dynamic tree cutting method. (C) Construction of gene expression cluster trees. (D) Sample dendrogram and trait heatmap. The leaves of the tree correspond to ankylosing spondylitis(AS) samples.

Figure 2

(A) and (B) Calculation of correlation between different gene co-expression modules and AS. Figure (A) shows the correlation between different gene co-expression modules and AS. Figure (B) shows the differential expression of these gene modules between AS samples and normal samples. Figure (C) Scatter diagram for module membership vs. gene significance of stage (AS or non-AS) in the Classical Module. Figure (D) Scatter diagram for module membership vs. gene significance of stage (AS or non-AS) in the Hematological Module.

Figure 3

Figure (A) Selection of the first 20 genes with the strongest correlation with AS in the Classical Module. The closer to the center of the figure, the stronger the correlation. Figure (B) A loop of the top 20 genes in the correlation in the Classical Module. Figure (C) shows the genes functions in the y Classical Module after GO and KEGG analysis.

Figure 4

Figure (A) Selection of the first 20 genes with the strongest correlation with AS in the Hematological Module. The closer to the center of the figure, the stronger the correlation. Figure (B) A loop of the top 20 genes in the correlation in the Hematological Module. Figure (C) shows the functions of the genes in the Hematological Module after GO and KEGG analysis.

A

B

Figure 5

Figure (A) The miRNA-mRNA interaction network of Classical Module (B) The miRNA-mRNA interaction network of Hematological Module.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementtable1.docx](#)
- [Supplementtable2.docx](#)