NR4A3 and CCL20 Clusters Dominate the Dynamic Gene Network of Peripheral CD146+ Blood Cells in the Early Stage of Acute Myocardial Infarction in Human.

Yanhui Wang
Shandong University of Science and Technology
Chenxin Li
Shandong University of Science and Technology
Jessica M Stephenson
University of Texas Health Science Center at Houston
Sean P Marrelli
University of Texas health science center at houston
Yanming Kou
Shandong University of Science and Technology
Dazhi Meng
Beijing University of Technology
Ting Wu (ting.wu@uth.tmc.edu)
University of Texas Health Science Center at Houston https://orcid.org/0000-0001-8152-1112

Research

Keywords: acute myocardial infarction (AMI), CD146, Pearson network, clustering coefficient, differential connectivity genes (DCGs)

Posted Date: December 28th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-133154/v1

NR4A3 and *CCL20* clusters dominate the dynamic gene network of peripheral CD146^{+} blood cells in the early stage of acute myocardial infarction in human.

Yan-hui Wang[1, *], PhD; Chen-xin Li[1], MS; Jessica M Stephenson[2], BS; Sean P Marrelli[2], PhD; Yan-ming Kou[1], MS; Da-zhi Meng[3, *], MS; Ting Wu[2, *], PhD;

[1]College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao, China

[2]Department of neurology, University of Texas health science center at Houston, Houston, Tx, USA

[3]College of Applied Science, Beijing University of Technology, Beijing, China

*Corresponding authors

Yanhui Wang, tel: +86-15192542635, yanhuiwang2014@163.com, 579 Qianwangang Road, Huangdao district, Qingdao, Shandong, 266590, China.

Dazhi Meng, tel: +86-13701377108, dzhmeng07@163.com, 100 Pingleyuan chaoyang district, Beijing, 10024, China.

Ting Wu, tel: +1 5022998114, ting.wu@uth.tmc.edu, 6431 Fannin street, Houston, 77031, USA.

Short title: CCL20 and NR4A3 lead CD146-mediated AMI pathology.

Abstract

Background: CD146 is a tight junction associated molecule involved in maintaining endothelial barrier and balancing immune-inflammation response in cardiovascular disease. Notably, the peripheral $CD146^+$ cells significantly upsurge under vessel dyshomeostasis like acute myocardial injury (AMI), appearing to be promising therapeutic targets. In this study, in a new view of gene correlation, we aim at deciphering the underlying complex mechanism of $CD146^+$ cells in the development of AMI.

Methods: Transcription dataset GSE 66360 of $CD146^+$ blood cells from clinical subjects were downloaded from NCBI. Pearson networks were constructed and the clustering coefficients were calculated to disclose the differential connectivity genes (DCGs). Analysis of gene connectivity and gene expression was performed to reveal the hub genes and hub genes clusters followed by gene enrichment analysis.

Results and conclusions: Among the total 23520 genes, 27 genes out of 126 differential expression genes are identified as DCGs. Those DCGs normally stay in the peripheral of networks while transfer to the functional central position under AMI situation. Moreover, it is revealed that DCGs spontaneously crowd together into two functional models, *CCL20* cluster and *NR4A3* cluster, influencing the CD146-mediated signaling pathways during the pathology of AMI for the first time.

Keywords: acute myocardial infarction (AMI), CD146, Pearson network, clustering coefficient, differential connectivity genes (DCGs)

1. Introduction

Cluster of differentiation 146 (CD146) / melanoma cell associated molecule is an essential immunoglobulin-like protein initially discovered in metastatic melanoma (1). It locates at endothelial tight junctions across all vessel beds, mediating physiological and pathological events under vascular dyshomeostasis (2, 3). Pioneering researchers regard CD146 as a historical marker for isolating circulation endothelial cells that sloughed off the inflamed vasculature (4). Over several decades, CD146 has also been discovered in other cell types including mesenchymal stem cells (5), endothelial progenitor cells (6), macrophages (7), T helper 17 cells (8), B lymphocytes (9), T lymphocytes (9, 10), and natural killer cells (9). The $CD146^{+}$ circulating cells occupy about 2 % of peripheral mononuclear cells in healthy individuals (9) and most notably, this percentage increases in certain conditions associated with vascular dysfunction for instance myocardial infarction, connective tissue diseases, and cancers (6, 11-13). Moreover, CD146 activated T cells have shown an enhanced ability to interact with endothelium in adhesion, rolling, and transmigration, evidenced by human and murine studies (14, 15). Given its multi-function in vessel structure, angiogenesis, and lymphocyte activation and its enabled detection in the bloodstream, CD146 appears to be a potential target for vascular disorders (16-18).

Complex networks are of great interest to researchers in the fields of computational biology and bioinformatics (19-21). It has been gradually extended from initial gene comparison to protein-protein network modeling, to protein-genetic investigation, and up to the disease-

disease association exploration (22). Most of the successful bioinformatics approaches that identify the initial key genes, however, have based on sole gene expression comparison and accordingly the top differential expression genes (DEGs) forward to mechanism validation without paying attention to the gene interaction rearrangement (23, 24). Instead, the hub-structured network is an important motif that is, to our best knowledge, leading the genome-wide association characterization in complex networks (25, 26). It generates the structure view angle to present the innermost gene-gene interaction, thus giving a comprehensive understanding of underlying mechanisms of disorders.

Acute myocardial injury (AMI) dataset GSE 66360 rests on the performance of $CD146^+$ populations during the AMI early development (12, 27-29). In this paper, we try to decipher gene reassemble with the correlation network structure parameter analysis (30, 31), and extract optimal genes collection, the differential connectivity genes (DCGs), and reveal functional gene clusters which likely leading the pathogenesis of peripheral $CD146^+$ blood cells during the AMI development in human for the first time.

2. Materials and methods

2.1 Data

The GSE66360 (12) gene transcription profile data of human AMI in the NCBI database was selected as the primary interest. Clinical subjects including fifty healthy individuals and forty-nine AMI patient subjects were recruited in the original investment by the Topol group. To

gather the data, $CD146^+$ cells were obtained by CD146-based magnetic immunoisolation from the subjects' blood samples. RNA samples were isolated from the $CD146^+$ cells and processed microarray by Affymetrix human U133 Plus 2.0 array. In this study, two cohorts were formed, a discovery cohort, consisting of twenty-two healthy subjects (control group) and twenty-one AMI patients (AMI group), which were used for the discovery of genes and appropriate testing methods, along with a validation cohort, consisting of twenty-eight healthy subjects and twenty-eight AMI patients, which were used for the validation of the genes and methods discovered in the other cohort. No data was excluded from the original databases used during this study.

2.2 Study design

Firstly, using the hypothesis test, we distinguished DEGs based on the gene expression profile in the discovery cohort and then verified in the validation cohort. Secondly, the gene networks of DEGs were constructed based on the Pearson coefficients, followed by the network separation assessment. Thirdly, the clustering coefficient, which is a parameter indicating gene connectivity, was calculated for each DEG under each gene network (32). Accordingly, genes with a clustering coefficient that represent a consistent increase in the AMI group among different cohorts were labeled as DCGs. Finally, two-dimensional analysis of gene connectivity and expression was employed, for identifying the hub gene clusters, alongside performing the gene enrichment analysis (**Figure 1**).

2.3 Identify DEGs

A total of 23520 genes were screened in each sample. The hypothesis test was used to screen the DEGs between control and AMI group (33). The method primarily gave weight for the distribution shape of the expression spectrum. If the distribution shape was different between the two groups, then the gene expression was different and the significance level was α_1. If not, a normal distribution test (significance level α_2) and homogeneity test of variance (significance level α_3) would be carried out. t-test or welch's t-test (significance level α_4) was used for normal distribution; Rank sum test (significance level α_4) was used for abnormal distribution with a similar distribution of expression spectrum (33). We defined $\alpha_1 = 0.00001$, $\alpha_2=0.00002$, $\alpha_3 = 0.00001$, and $\alpha_4= 0.00001$ as the significant level of the hypothesis test.

2.4 Clustering coefficient

A local clustering coefficient was introduced to measure the compactness, or the connectivity of genes within a suspected cluster, of a complete array formed by the adjacent nodes within a network (34). To clarify, assume that a node i in a network was connected to k_i nodes. The k_i nodes were called neighbors of node i. The ratio of the actual number E_i of edges and the total number $k_i(k_i-1)/2$ of possible edges between k_i nodes were defined as the clustering coefficient, C_i, of node i, that is, $C_i = 2E_i/(k_i(k_i-1))$.

2.5 Pearson network construction and assessment

Pearson correlation networks of DEGs were constructed according to the absolute value of Pearson coefficients. Two genes were considered correlated if the absolute value of the Pearson

coefficient was greater than the threshold x ($0< x <1$), then a line could be drawn between the two genes. In the cases when genes were not correlated, there would be no link in the network and thus no line could be drawn. Gene clusters were determined by examining the clustering coefficients, and those with a non-zero value could be labeled clusters. Gene clusters represent a functional module as a whole with varying degrees of connectivity; while a degree describes the number of genes connected to one another. The average clustering coefficients of DEGs were calculated to evaluate the overall separation of the control and AMI networks. The method was implemented in R i386 3.6.2.

Natural biological networks are scale-free networks and the degree distributions follow the power-law exponential distribution index range $2 \sim 3$ (35, 36). We indicated the gene networks under threshold 0.5 and 0.7 since the power-law indexes of degree distribution in discovery cohort were in the range of $2 \sim 3$ and presented the corresponding networks in validation cohort in this study (**Supplement Table 1**).

2.6 Identify DCGs

In the analysis of network connection parameters, the greater the difference between the control and AMI group, the higher the correlation with AMI. The following is describing our unique identifying method. Assume that the average clustering coefficient of the control group and AMI group could be separated at the threshold [0.1, 0.9]. First, the clustering coefficient of each gene in the Pearson correlation networks of the control group and AMI group, under the threshold $0.1 \sim 0.9$, were calculated with step length 0.1. Secondly, the average clustering

coefficient of each gene cross threshold 0.4 ~ 0.8 was calculated to compare changes in connectivity between the two groups within the validation and discovery cohorts. Finally, if clustering coefficient differences in discovery cohort and validation cohorts were consistently greater than 0.1 between the AMI group and the control group, the genes were identified as candidates for DCGs.

To test the reliability of the proposed candidate genes across different datasets, we expanded our method to a combination cohort, which included all subjects in both discovery and validation cohorts. Increasing the number of subjects, but also introducing some variation in the data due to the less categorized subject population. The overall network between the control and the AMI groups were still separable through threshold 0.4 ~ 0.8 (data not shown). While having 27 out of 39 candidate genes still showing clustering coefficient differences were greater than 0.1 in this combination cohort were define as DCGs.

2.7 Gene set enrichment

Gene set enrichment was performed by the STRING server. Biological process, Reactome pathways and protein-protein association networks were generated for *CCL20* cluster, *NR4A3* cluster and DCGs.

2.8 Graphs

Heatmaps of DEGs and DCGs were generated by using heatmap.2 function in the gplots package. Networks were computed by using igraph::graph.data.frame function. Layout

algorithm of layout.kamada.kawai was used for visualizing the overall DEGs networks and the connections for individual genes. Layout algorithm of layout.circle was used to visualize the gene connections within DEGs and DCGs in circle view. Cytoscape network function was used to generate the clustered DCGs networks.

3. Results

3.1 DEGs identification

In our initial analysis, 126 out of 23520 genes are significantly altered in the AMI group compare to the control group in discovery cohort, defined as DEGs, with the majority (79 of 126) demonstrate an up-regulation feature (**Figure 2A**). And those genes show a similar expression pattern in the validation cohort (**Figure 2B**).

3.2 Assessment of DEGs' networks

The overall gene networks of DEGs in the control group and the AMI group are distinctly independent through a large range of thresholds in discovery cohort and validation cohort (**Figure 3A**). The networks in discovery cohort are separable through threshold 0.1 ~ 0.9 and in validation cohort are separable through 0.1 ~ 0.8. The average separable widths of the discovery cohort and validation cohort are 0.218 and 0.0518, respectively. The validation cohort shows a narrower split range possibly attributed to the variations between two cohorts, for instance, the differential sample size, age, and co-morbid disorders.

In addition, the gene connections within DEGs' networks in the AMI group are more complex than those in the control group in both cohorts (**Figure 3C**). In discovery cohort, the number of gene clusters within AMI network gradually decreases from 125 to 67, when threshold increases from 0.4 to 0.8, while it more sharply decreases from 123 to 17 in the control group (**Figure 3B**). Similarly, the clusters decline with a lower slope in the AMI group compared to the control group in validation cohort (**Figure 3B**).

The data described above suggest that gene networks of DEGs are largely and consistently disturbed by AMI stimulation as is seen in two independent cohorts, verifying our findings. Thus, DEGs and DEGs-networks are mathematically reliable and hereafter can be set as the foundation for in-depth gene interaction data mining.

3.3 DCGs identification and connectivity analysis

Beyond DEGs, we identified 27 genes as DCGs whose clustering coefficients difference is greater than 0.1 in the discovery cohort, the validation cohort, and the extended combination cohort (**Figure 4A, Supplement Table 2**). The sub-networks of DCGs present obvious tighter connections in the AMI group compare to the control group, in both discovery and validation cohorts (**Figure 4C**). When threshold increased from 0.4 to 0.8, the average degree of DCGs progressively decreases from 16.0 to 4.30 in the AMI group, while it decreases from 7.93 to 0 in the control group in discovery cohort (**Figure 4D**). Similarly, in validation cohort, this number decreases from 19.9 to 1.41 in the AMI group while from 3.26 to 0 in the control group

197 (**Figure 4D**). Besides clustering coefficients, gene expression of those DCGs is showing a
198 steady increase in the AMI group in discovery and validation cohorts (**Figure 4B**).

199 Therefore, we propose that the networks' differential of all DEGs largely attributes to the
200 connection changes within DCGs. As visualized in kamada-kawai layout, the DCGs randomly
201 participate in the DEGs' network and connect to a few genes under a normal steady state.
202 However, they appear to interact with more functional genes and shift into central positions
203 after AMI in both discovery and validation cohorts (**Figure 5**).

204 *3.4 Two-dimensional analysis of gene connectivity and gene expression*

205 Since the power-law indexes of degree distribution in the discovery are in the range of natural
206 network, we regard the discovery cohort as a more precise dataset and it is selected for the
207 following analysis. The average clustering coefficient and gene expression of DCGs are plotted
208 into a scatterplot for the two-dimensional analysis (**Figure 6A**). *NR4A3* and *CCL20* present
209 high levels of clustering coefficient and gene expression changes, defined as $CC^{high}GeExp^{high}$
210 genes. *SOCS3*, *FOSL2*, *PLIN2* are the genes that were found to have high clustering coefficient
211 changes (fold change > 2) with low expression change, defined as $CC^{high}GeExp^{low}$ genes; while
212 *IL1R2*, *NLRP3*, *ANXA3* and *AC079305.10* are the genes that were found to have high
213 expression changes (fold change > 0.4) with low clustering coefficient changes, defined as
214 $CC^{low}GeExp^{high}$ genes. All information on these genes is shown in **Table 1**. Subgraphs of
215 *NR4A3*, *CCL20,* and other DCGs provide pieces of evidence that support their increased gene
216 connectivity after AMI in the discovery cohort (**Figure 6C**, **Supplement Figure 1A and 1B**).

3.5 NR4A3 and CCL20 clusters identification

Zooming into the subgraphs of individual genes, we reveal that DCGs stay "non-activated" in the control group (**Figure 7A**). Interestingly, they appear to be spontaneously gathering together as two separate clusters after AMI stimulation (**Figure 7B, Supplement Figure 1A and 1B**). *CCL20* connects with *SKIL*, *MMP9*, *ITPRIP*, *ANXA3*, *GLUL*, *CXCL16*, *IL1R2*, *TMCC3*, *NLRP3*, *PYGL*, *RNF144B*, *BCL6*, *LILRB2*, *CLEC4E*, *FCER1G*, and *AC079305.10*, identify as the *CCL20* cluster. *NR4A3* connects with *NR4A2*, *FOSL2*, *CDKN1A*, *SOCS3*, *GABARAPL1*, *ITPRIP*, *SYTL3*, *PELI1*, *MAP3K8*, and *PLIN2*, identify as the *NR4A3* cluster. While there are overlapping genes between clusters, *ITPRIP*, *SKIL* and *MAPK38* are the intermediate genes that connect both clusters according to their subgraphs (**Supplement Figure 1B**). The $CC^{high}GeExp^{high}$ gene, *NR4A3* or *CCL20*, serve as leading-like hub gene in each cluster. The clustering coefficient fold changes of *NR4A3* and *CCL20* are 15.7 and 10.2, respectively; and the gene expression fold changes are 0.379 and 0.422, respectively.

3.6 Gene enrichment

Biological process analysis shows that DCGs are involving in response to organic substrates, positive regulation of leukocyte activation, immune response, immune system process, response to cytokine and regulation of cytokine production. The *CCL20* cluster is essential to the immune response, immune system process, and regulation of localization while the *NR4A3* cluster is essential to cellular response to corticotropin-releasing hormone stimulus, positive regulation of leukocyte activation, and regulation of apoptotic process (**Table 2**).

Reactome pathway analysis revealed that DCGs are related to the immune system with regards to tasks such as signaling by interleukins, namely interleukin-1, interleukin-4 and interleukin-13 signaling, the innate immune system and the dectin-2 family. The *CCL20* cluster is essential to immune system, innate immune system, the dectin-2 family, and neutrophil degranulation while the *NR4A3* cluster is essential to RNA Polymerase II Transcription, Generic Transcription Pathway, and MyD88 cascade initiated on plasma membrane (**Table 2**).

4. Discussion

CD146 is a junction-associated adhesion molecule that participates in immune and inflammatory pathological processes in the initiation and development of vascular diseases (2). CD146 activated leukocytes are recruited to the inflamed endothelial to induce the expression of chemokines and cytokines and, in doing so, progressively destroys the blood vessel barrier. Our study found that following AMI stimulation, in $CD146^{+}$ human blood cells, 126 out of total 23,520 genes show significant differential expression ($P<0.0001$) and among those, 27 genes show consistent connectivity changes and serve as DCGs. Unlike DEGs, DCGs are able to not only aggregate gene expression but also encompass gene connectivity properties, internally coupling into functional gene clusters—*NR4A3* cluster and *CCL20* cluster, orchestrating the gene networks' entire dynamics in CD146 associated AMI pathophysiology development. Meanwhile, *NR4A3* and *CCL20* are revealed as hub genes since they experienced both connectivity and expression experienced significant changes after AMI stimuli. Furthermore, gene enrichment analysis shows that the DCGs are involved in inflammation-

immune response, with *CCL20* being principal to the immune response and regulation of localization; while, the *NR4A3* cluster is principal to leukocyte activation, apoptotic process, and cellular response to corticotropin-releasing hormone stimulus; such findings align with the well-known hypothesis that CD146 mediated inflammation plays an important role in the pathogenesis of AMI.

The network structural parameter analysis method is applied to weave the gene-gene correlation network. We identify DCGs which present steadily elevated connectivity under AMI conditions, in both the discovery and validation cohorts, further confirming the upregulation seen in the combination cohort (**Supplement Table 2**). As expected, the gene expression of DCGs was increased after AMI, but was not distinguishable from DEGs solely by expression signature (data not shown). *NR4A3* and *CCL20* as highlight hub genes were also defined as AMI biomarkers after pre-filtering the co-morbidity relevant genes by the original Topol group (12). *SOCS3* tends to be the only "shared" AMI biomarker candidate revealed by other groups in which the same GSE66360 dataset is included as one of their study objects (27, 28). Recognizing the *CCL20*, *NR4A3*, and *SOCS3* as top DCGs substantiate previous outputs and in turn, the validity of our method is enhanced. Therefore, we recommend the gene connectivity analysis, along with gene expression signature, to be used as a powerful and unbiased way for researchers to rank the importance of candidate DEGs.

NR4A3 belongs to the *NR4A* orphan nuclear receptor family (with *NR4A2* and *NR4A1*), playing a protective role in AMI development. The JM Penninger group reports that *NR4A3* is the

277 highest-ranking gene in circulating human endothelial cells under atherosclerosis (37).
278 Transcription analysis of human left ventricular myocardium shows that *NR4A3* up-regulated
279 during ischemia and reperfusion in normal and chronic ischemic myocardium (38). Similarly,
280 *NR4A3* is found to be elevated 10-days post left anterior descending artery ligation ischemia
281 surgery in mice (39). Overexpression of *NR4A3* significantly reduces infarct size, preventing
282 deterioration of left ventricular function and repression of neutrophil infiltration in the heart of
283 mice after coronary artery ligation, relate to the activation of *JAK2/STAT3* and the inhibition
284 of *STAT3* dependent *NF-κB* signaling pathways (40). Additionally, it has to point out that the
285 *NR4A* subgroup including *NR4A3* is an immediate early response gene induced by diverse
286 physiological, i.e., mechanical agitation, calcium, and inflammation cytokines (41). This
287 reinforces our data that, in the very early-stage AMI, *NR4A3* has a significant 16-fold clustering
288 coefficient climb and 42% gene expression increase. Yet the nuclear factor *NR4A3* implications
289 in $CD146^{+}$ related myocardial disorders remain a mystery.

290 *CCL20*, a C-C motif chemokine, is a chemoattractant for recruiting leukocytes to sites of injury
291 and inflammation (**Figure 6B**). *CCL20* secretion is induced by pro-inflammatory chemokines
292 and cytokines, such as *CXCL12*, *IL17*, *IL1β*, *IL6*, and is in part related to *JAK/STAT* pathway
293 signaling in multiple cells (42-44). *IL6* and soluble *IL6* receptor stimulate *STAT3* binding to
294 the *CCL20* promotor and *IL17* stimulate the phosphorylated *NF-κB* binding to the *CCL20*
295 promoter in murine astrocytes, facilitate the neuroinflammation within central nervous system
296 (42). In addition, the co-expression of *CCL20* receptor *CCR6* and CD146 is a marker of effector

297 memory Th17 cells, which mediate migration and is thought to be essential for inflammation
298 in human psoriasis (8). Moreover, it is reported that *CCL20* level elevated in clinical patients'
299 serum with ischemic myocardial infarction (45, 46). In vitro study shows that *CCL20*
300 expression increase in $CD146^{+}$ human mesenchymal stromal cells at the early pro-
301 inflammatory phase in fracture healing (47). Thus, we hypothesis that *CCL20* binding its
302 receptor *CCR6* is what drives the CD146-mediated vessel inflammation progress in early AMI
303 phase.

304 In terms of functional models, DCGs are self-organized into two clusters, the *NR4A3* and
305 *CCL20* clusters, with 18 genes and 12 genes in each cluster, respectively. All genes directly
306 link to its hub gene and partly link to adjacent genes as shown in **Figure 7**. Protein-protein
307 connection analyzed by STRING database produced a structure that is greatly similar to our
308 network pattern in which *CCL20* connects with *CXCL16*, *IL1R2*, *MMP9*, *NLRP3*, *BCL6*
309 *LILRB2*, *PELI1*, *CLEC4E*, *FCER1G*, and *NR4A3* connect with *NR4A2*, *FOSL2*, *RNF144B*,
310 *CDKN1A*, *SOCS3* (**Supplement Figure** 2). A few of gene-gene correlations within clusters are
311 stated in inflammatory diseases. *MMP9* activation correlates with *CCL20* expression in
312 astrocytes via *Notch-1/Akt/NF-κB* pathway promoting leukocyte migration cross blood-brain
313 barrier (48). *NR4A2* and *NR4A3* as orphan nuclear receptors mediate neutrophil number and
314 survival in chronic inflammatory signals multiple hematologic disorders (49-51). *FOSL2* acts
315 as an activating protein-1 transcription factor promoting hematopoietic progenitor cell to

316 macrophage and neutrophils in a *SOCS3* dependent manner is reported (52). Nevertheless, most
317 of the cluster functions are rarely reported in AMI pathogenesis.

318 Taken together, we reveal that *NR4A3* and *CCL20* clusters are novel functional modules in
319 $CD146^+$ cells-mediated immuno-inflammatory balance, triggering increased susceptibility to
320 vascular deterioration and accelerating myocardial injury. *NR4A3 and CCL20* as hub genes
321 largely impact the early AMI development and can be promising targets for clinical diagnosis.
322 In-depth studies are necessary for understanding the mechanisms of peripheral $CD146^+$ cells
323 in cardiovascular disease.

324 **Abbreviations list**

325 CD146: cluster of differentiation 146; DEGs: differential expression genes; DCGs: differential
326 connectivity genes; AMI: Acute myocardial injury; CC: clustering coefficient; GeExp: gene
327 expression.

328 **Declaration**

329 **Ethics approval and consent to participate**

330 Not required. The data has been available in the NCBI database.

331 **Availability of data and materials**

332 All data are included in the manuscript.

Consent for publication

Not required.

Authors' contributions

Yanhui, Dazhi Meng and Ting Wu did the overall design of the study. Jessica M Stephenson did the language editor for the manuscript. Chenxin Li contributed to the igraph figures. All authors involved in the analysis of the data. All authors read and approved the final manuscript.

Competing interests

No competing interests.

Funding

This research was supported by the Cultivation Project of Young and Innovative Talents in Universities of Shandong Province and by NSFC (Grant No: 11501331).

Acknowledgments

The authors would like to acknowledge authors of primary study, the Topol group.

References

1. Lehmann JM, Holzmann B, Breitbart EW, Schmiegelow P, Riethmüller G, Johnson JP. Discrimination between benign and malignant cells of melanocytic lineage by two novel antigens, a glycoprotein with a molecular weight of 113,000 and a protein with a molecular weight of 76,000. Cancer research. 1987;47(3):841-5.

2. Shih IM. The role of CD146 (Mel‐CAM) in biology and pathology. The Journal of pathology. 1999;189(1):4-11.

3. Tu T, Zhang C, Yan H, Luo Y, Kong R, Wen P, et al. CD146 acts as a novel receptor for netrin-1 in promoting angiogenesis and vascular development. Cell research. 2015;25(3):275-87.

4. Widemann A, Sabatier F, Arnaud L, Bonello L, AL‐MASSARANI G, Paganelli F, et al. CD146‐based immunomagnetic enrichment followed by multiparameter flow cytometry: a new approach to counting circulating endothelial cells. Journal of thrombosis and haemostasis. 2008;6(5):869-76.

5. Espagnolle N, Guilloton F, Deschaseaux F, Gadelorge M, Sensébé L, Bourin P. CD 146 expression on mesenchymal stem cells is associated with their vascular smooth muscle commitment. Journal of cellular and molecular medicine. 2014;18(1):104-14.

6. Delorme B, Basire A, Gentile C, Sabatier F, Monsonis F, Desouches C, et al. Presence of endothelial progenitor cells, distinct from mature endothelial cells, within human CD146^+ blood cells. THROMBOSIS AND HAEMOSTASIS-STUTTGART-. 2005;94(6):1270.

7. Luo Y, Duan H, Qian Y, Feng L, Wu Z, Wang F, et al. Macrophagic CD146 promotes foam cell formation and retention during atherosclerosis. Cell research. 2017;27(3):352-72.

8. Kamiyama T, Watanabe H, Iijima M, Miyazaki A, Iwamoto S. Coexpression of CCR6 and CD146 (MCAM) is a marker of effector memory T‐helper 17 cells. The Journal of dermatology. 2012;39(10):838-42.

9. Elshal MF, Khan SS, Takahashi Y, Solomon MA, McCoy JP. CD146 (Mel-CAM), an adhesion marker of endothelial cells, is a novel marker of lymphocyte subset activation in normal peripheral blood. Blood. 2005;106(8):2923-4.

10. Duda DG, Cohen KS, di Tomaso E, Au P, Klein RJ, Scadden DT, et al. Differential CD146 expression on circulating versus tissue endothelial cells in rectal cancer patients: implications for circulating endothelial and progenitor cells as biomarkers for antiangiogenic therapy. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 2006;24(9):1449.

11. Hadjinicolaou A, Wu L, Fang B, Watson P, Hall F, Busch R. Relationship of CD 146 expression to activation of circulating T cells: exploratory studies in healthy donors and patients with connective tissue diseases. Clinical & Experimental Immunology. 2013;174(1):73-88.

12. Muse ED, Kramer ER, Wang H, Barrett P, Parviz F, Novotny MA, et al. A whole blood molecular signature for acute myocardial infarction. Scientific reports. 2017;7(1):1-9.

13. Fürstenberger G, Von Moos R, Senn H, Boneberg E. Real-time PCR of CD146 mRNA in peripheral blood enables the relative quantification of circulating endothelial cells and is an indicator of angiogenesis. British journal of cancer. 2005;93(7):793-8.

14. Guezguez B, Vigneron P, Lamerant N, Kieda C, Jaffredo T, Dunon D. Dual role of melanoma cell adhesion molecule (MCAM)/CD146 in lymphocyte endothelium interaction: MCAM/CD146 promotes rolling via microvilli induction in lymphocyte and is an endothelial adhesion receptor. The Journal of Immunology. 2007;179(10):6673-85.

15. Dagur PK, McCoy Jr JP. Endothelial-binding, proinflammatory T cells identified by MCAM (CD146) expression: Characterization and role in human autoimmune diseases. Autoimmunity reviews. 2015;14(5):415-22.

16. Leroyer AS, Blin MG, Bachelier R, Bardin N, Blot-Chabaud M, Dignat-George F. CD146 (Cluster of Differentiation 146) An Adhesion Molecule Involved in Vessel Homeostasis. Arteriosclerosis, thrombosis, and vascular biology. 2019;39(6):1026-33.

17. Gallastegi T, Soto B, Romero JM, Galán M, Escudero JR, Camacho M. MCAM/CD146 Which is Differentially Expressed in Vascular Diseases, is a Potential Biomarker in Abdominal Aortic Aneurysm. European Journal of Vascular and Endovascular Surgery. 2019;58(6):e454.

18. Wang Z, Yan X. CD146, a multi-functional molecule beyond adhesion. Cancer letters. 2013;330(2):150-62.

19. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. Network motifs: simple building blocks of complex networks. Science. 2002;298(5594):824-7.

20. Zhao Z, Li C, Zhang X, Chiclana F, Viedma EH. An incremental method to detect communities in dynamic evolving social networks. Knowledge-Based Systems. 2019;163:404-15.

411 21. Liu Y-Y, Slotine J-J, Barabási A-L. Controllability of complex networks. nature.
412 2011;473(7346):167-73.
413 22. Goh K-I, Cusick ME, Valle D, Childs B, Vidal M, Barabási A-L. The human disease
414 network. Proceedings of the National Academy of Sciences. 2007;104(21):8685-90.
415 23. Akat KM, Morozov P, Brown M, Gogakos T, Da Rosa JC, Mihailovic A, et al.
416 Comparative RNA-sequencing analysis of myocardial and circulating small RNAs in human
417 heart failure and their utility as biomarkers. Proceedings of the National Academy of
418 Sciences. 2014;111(30):11151-6.
419 24. Eicher JD, Wakabayashi Y, Vitseva O, Esa N, Yang Y, Zhu J, et al. Characterization of
420 the platelet transcriptome by RNA sequencing in patients with acute myocardial infarction.
421 Platelets. 2016;27(3):230-9.
422 25. Xiao M, Zheng WX, Jiang G, Cao J. Stability and bifurcation analysis of arbitrarily high-
423 dimensional genetic regulatory networks with hub structure and bidirectional coupling. IEEE
424 Transactions on Circuits and Systems I: Regular Papers. 2016;63(8):1243-54.
425 26. De Domenico M, Nicosia V, Arenas A, Latora V. Structural reducibility of multilayer
426 networks. Nature communications. 2015;6(1):1-9.
427 27. Cheng M, An S, Li J. Identifying key genes associated with acute myocardial infarction.
428 Medicine. 2017;96(42).
429 28. Ge WH, Lin Y, Li S, Zong X, Ge ZC. Identification of biomarkers for early diagnosis of
430 acute myocardial infarction. Journal of Cellular Biochemistry. 2018;119(1):650-8.

29. Qiu L, Liu X. Identification of key genes involved in myocardial infarction. European journal of medical research. 2019;24(1):22.

30. Guo Y, Wu C, Guo M, Liu X, Keinan A. Gene-Based Nonparametric Testing of Interactions Using Distance Correlation Coefficient in Case-Control Association Studies. Genes. 2018;9(12):608.

31. Wang Y, Chi X, Meng D, editors. The Application of Network Structure Aanalysis in the Study of Disease Mechanisms. 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2019: IEEE.

32. Opsahl T. Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. Social Networks. 2013;35(2):159-67.

33. Wang Y, Kou Y, Meng D. Network Structure Analysis Identifying Key Genes of Autism and Its Mechanism. Computational and mathematical methods in medicine. 2020;2020.

34. Galton F. Typical laws of heredity. III. Nature. 1877;15(389):512-4.

35. Jeong H, Tombor B, Albert R, Oltvai ZN, Barabási A-L. The large-scale organization of metabolic networks. Nature. 2000;407(6804):651-4.

36. Jeong H, Mason SP, Barabási A-L, Oltvai ZN. Lethality and centrality in protein networks. Nature. 2001;411(6833):41-2.

37. Barrett P, Topol EJ. TCT-639 NR4A3 as a gene expression marker of acute atherosclerotic plaque rupture in STEMI. Journal of the American College of Cardiology. 2013;62(18 Supplement 1):B194.

451 38. Gabrielsen A, Lawler PR, Yongzhong W, Steinbrüchel D, Blagoja D, Paulsson-Berne G,
452 et al. Gene expression signals involved in ischemic injury, extracellular matrix composition
453 and fibrosis defined by global mRNA profiling of the human left ventricular myocardium.
454 Journal of molecular and cellular cardiology. 2007;42(4):870-83.
455 39. Haubner BJ, Adamowicz-Brice M, Khadayate S, Tiefenthaler V, Metzler B, Aitman T, et
456 al. Complete cardiac regeneration in a mouse model of myocardial infarction. Aging (Albany
457 NY). 2012;4(12):966.
458 40. Jiang Y, Feng Y-P, Tang L-X, Yan Y-L, Bai J-W. The protective role of NR4A3 in acute
459 myocardial infarction by suppressing inflammatory responses via JAK2-STAT3/NF-κB
460 pathway. Biochemical and biophysical research communications. 2019;517(4):697-702.
461 41. Maxwell MA, Muscat GE. The NR4A subgroup: immediate early response genes with
462 pleiotropic physiological roles. Nuclear receptor signaling. 2006;4(1):nrs. 04002.
463 42. Meares GP, Ma X, Qin H, Benveniste EN. Regulation of CCL20 expression in astrocytes
464 by IL‐6 and IL‐17. Glia. 2012;60(5):771-81.
465 43. Beider K, Abraham M, Begin M, Wald H, Weiss ID, Wald O, et al. Interaction between
466 CXCR4 and CCL20 pathways regulates tumor growth. PloS one. 2009;4(4):e5125.
467 44. Hosokawa Y, Shindo S, Hosokawa I, Ozaki K, Matsuo T. IL-6 trans-signaling enhances
468 CCL20 production from IL-1β-stimulated human periodontal ligament cells. Inflammation.
469 2014;37(2):381-6.

45. Lin C-F, Su C-J, Liu J-H, Chen S-T, Huang H-L, Pan S-L. Potential effects of CXCL9 and CCL20 on cardiac fibrosis in patients with myocardial infarction and isoproterenol-treated rats. Journal of clinical medicine. 2019;8(5):659.

46. Safa A, Rashidinejad H, Khalili M, Dabiri S, Nemati M, Mohammadi M, et al. Higher circulating levels of chemokines CXCL10, CCL20 and CCL22 in patients with ischemic heart disease. Cytokine. 2016;83:147-57.

47. Herrmann M, Stanić B, Hildebrand M, Alini M, Verrier S. In vitro simulation of the early proinflammatory phase in fracture healing reveals strong immunomodulatory effects of CD146‐positive mesenchymal stromal cells. Journal of tissue engineering and regenerative medicine. 2019;13(8):1466-81.

48. Song J, Wu C, Korpos E, Zhang X, Agrawal SM, Wang Y, et al. Focal MMP-2 and MMP-9 activity at the blood-brain barrier promotes chemokine-induced leukocyte migration. Cell reports. 2015;10(7):1040-54.

49. McMorrow JP, Murphy EP. Inflammation: a role for NR4A orphan nuclear receptors? Biochemical Society Transactions. 2011;39(2):688-93.

50. Prince LR, Prosseda SD, Higgins K, Carlring J, Prestwich EC, Ogryzko NV, et al. NR4A orphan nuclear receptor family members, NR4A2 and NR4A3, regulate neutrophil number and survival. Blood, The Journal of the American Society of Hematology. 2017;130(8):1014-25.

51. Liu H, Liu P, Shi X, Yin D, Zhao J. NR4A2 protects cardiomyocytes against myocardial infarction injury by promoting autophagy. Cell death discovery. 2018;4(1):1-11.

491 52. Croker BA, Mielke LA, Wormald S, Metcalf D, Kiu H, Alexander WS, et al. Socs3
492 maintains the specificity of biological responses to cytokine signals during granulocyte and
493 macrophage differentiation. Experimental hematology. 2008;36(7):786-98.

494

495 **Figure Legends**

496 **Figure 1.** Flow chart for study design. DEGs, differential expression genes; DCGs, differential
497 connectivity genes.

498 **Figure 2.** Gene expression profiles of DEGs. 126 genes show significant differential
499 expressions between the AMI and the control groups in the discovery cohort (**A**) and validation
500 cohort (**B**), thus define as DEGs. DEGs, differential expression genes.

501 **Figure 3.** Assessment of DEGs' networks. Networks in the control and AMI groups are
502 independent and separable according to the average clustering coefficients of DEGs (**A**).
503 Number of clusters within DEGs' networks progressively decline when thresholds increase
504 from 0.1 to 0.9 (**B**). The AMI group has a lower decline slope. The gene networks of DEGs in
505 the AMI group has more complex connection compare to that in the control group (**C**).
506 Networks are present under threshold 0.5 and 0.7. Darker line represents connections under
507 threshold 0.7; lighter line represents connections under threshold 0.5. DEGs, differential
508 expression genes.

Figure 4. Identification of DCGs. Genes that clustering coefficient increased over 0.1 in the AMI group, in discovery cohort and validation cohort, are revealed as DCGs (**A**). Gene expression profile of DCGs shows stable increase in AMI group in two cohorts (**B**). The connection among DCGs in the AMI group are denser (**C**) and the average degrees of DCGs in AMI group are higher (**D**) compare to the control group in two cohorts. Networks are presented under threshold 0.5 and 0.7. Darker line represents connections under threshold 0.7; lighter line represents connections under threshold 0.5. Degree are presented as mean ± SEM. DCGs, differential connectivity genes.

Figure 5. Visualization of DCGs in DEGs' networks. The networks of DEGs in discovery cohort (**A**) and in the validation cohort (**B**) indicate that the DCGs participate in distinctive ways in the control group and in the AMI group. DCGs switch to central functional position of networks and participate in more intricate connections under AMI situation. Yellow nodes indicate the DCGs. Red gene names indicate the hub genes. DEGs, differential expression genes; DCGs, differential connectivity genes.

Figure 6. Analysis of gene connection and expression of DCGs in discovery cohort. The analysis of clustering coefficient and gene expression revealed *CCL20* and *NR4A3* as hub genes (**A**). The *CCL20* is a chemoattractant while *NR4A3* is a nuclear factor receptor (**B**). Subgraphs of *CCL20* and *NR4A3* substantiate their important roles in AMI development (**C**). Networks are presented under threshold 0.5 and 0.7. Darker line represents connections under threshold 0.7; lighter line represents connections under threshold 0.5. DCGs, differential connectivity genes; CC, clustering coefficient; GeExp, gene expression.

Figure 7. *CCL20* cluster and *NR4A3* cluster formation in early-stage AMI. *CCL20* and *NR4A3* stay in the peripheral position of DCGs' network under normal state (**A**). However, they shift to the primary position of DCGs' network dominating two functional clusters under AMI stimulation (**B**). DCGs, differential connectivity genes.

Figures

Figure 1.

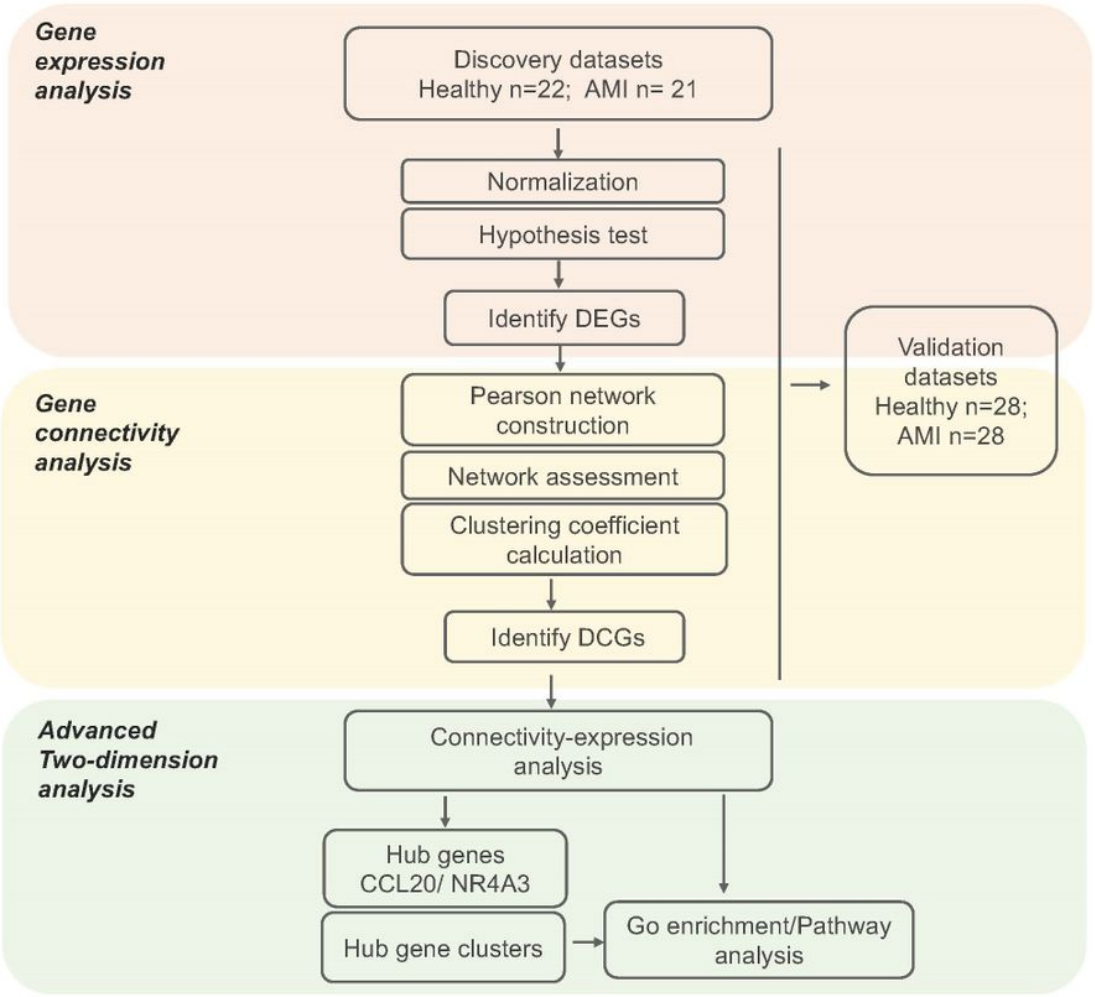

Figure 1

Flow chart for study design. DEGs, differential expression genes; DCGs, differential connectivity genes.

Figure 2.

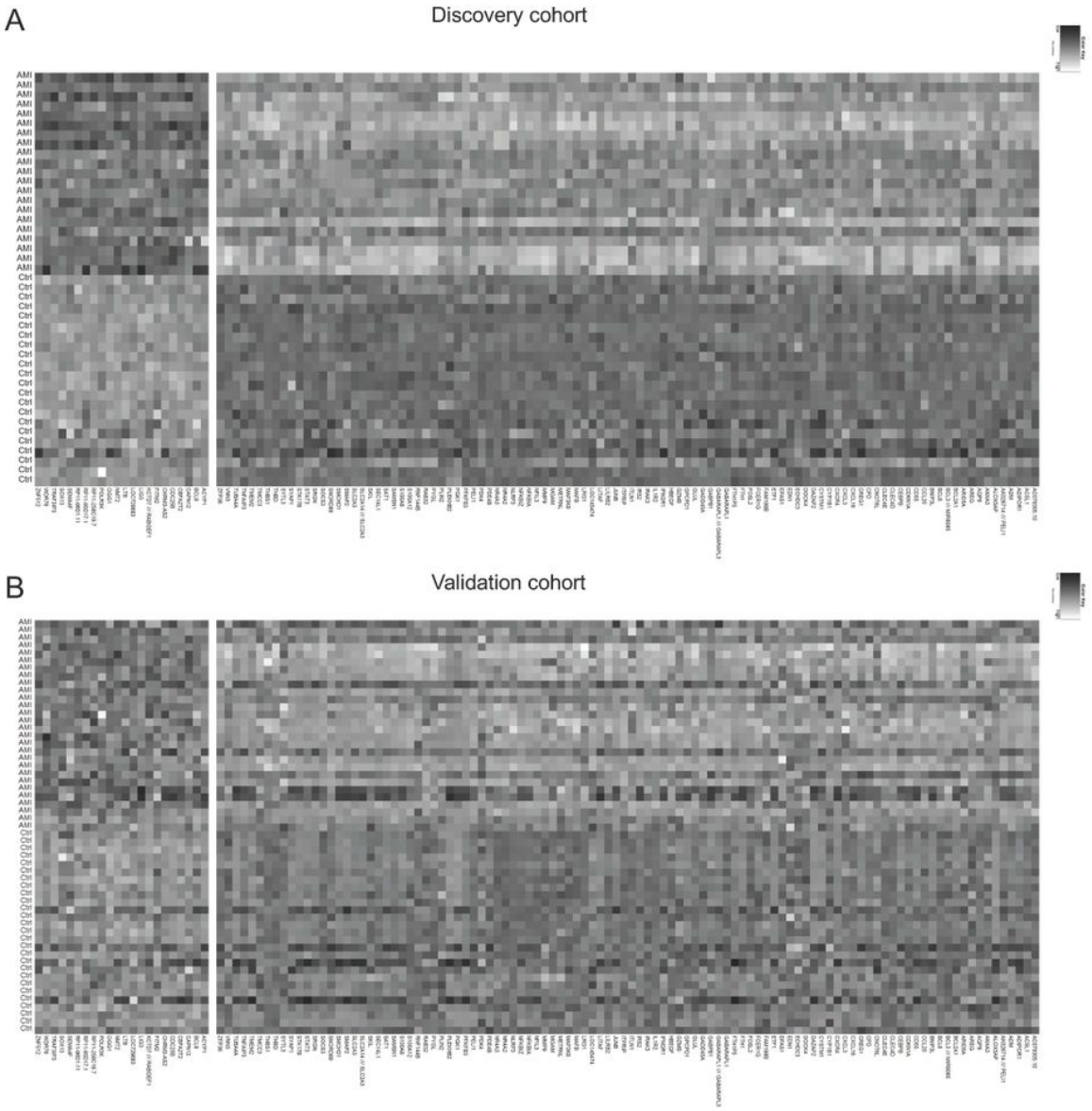

Figure 2

Gene expression profiles of DEGs. 126 genes show significant differential expressions between the AMI and the control groups in the discovery cohort (A) and validation cohort (B), thus define as DEGs. DEGs, differential expression genes.

Figure 3.

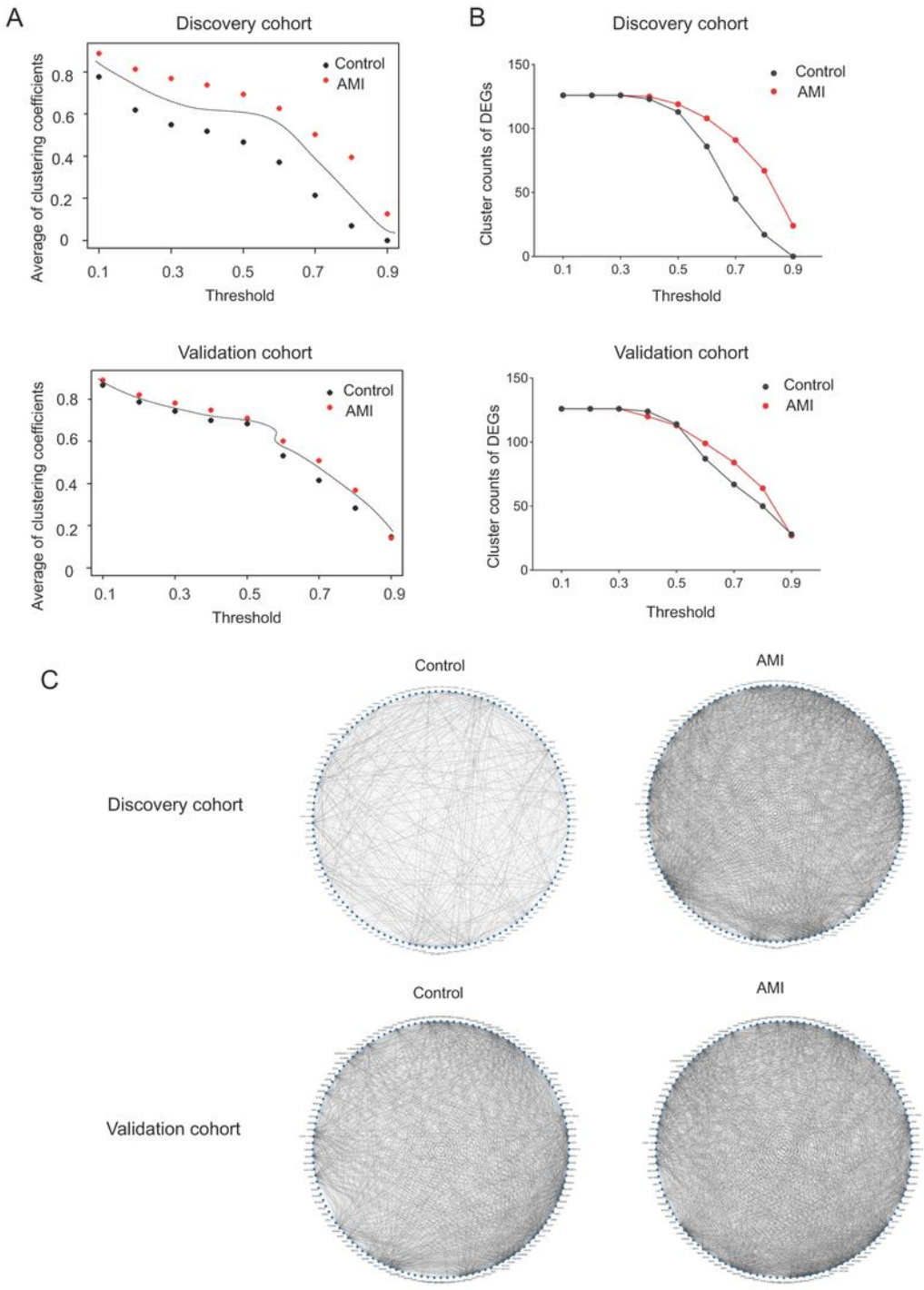

Figure 3

Assessment of DEGs' networks. Networks in the control and AMI groups are independent and separable according to the average clustering coefficients of DEGs (A). Number of clusters within DEGs' networks progressively decline when thresholds increase from 0.1 to 0.9 (B). The AMI group has a lower decline slope. The gene networks of DEGs in the AMI group has more complex connection compare to that in the control group (C). Networks are present under threshold 0.5 and 0.7. Darker line represents connections

under threshold 0.7; lighter line represents connections under threshold 0.5. DEGs, differential expression genes.

Figure 4.

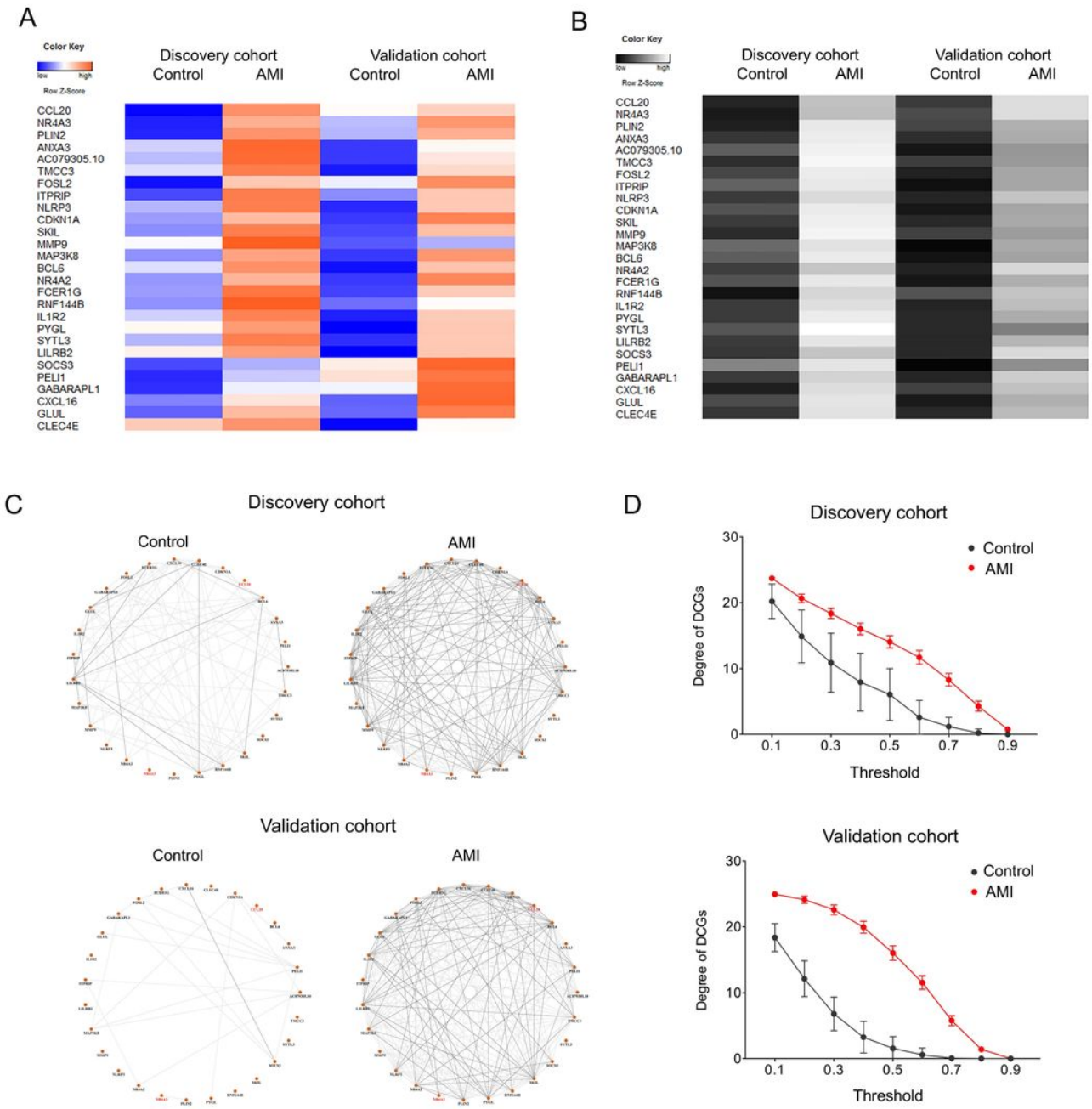

Figure 4

Identification of DCGs. Genes that clustering coefficient increased over 0.1 in the AMI group, in discovery cohort and validation cohort, are revealed as DCGs (A). Gene expression profile of DCGs shows stable increase in AMI group in two cohorts (B). The connection among DCGs in the AMI group are denser (C) and the average degrees of DCGs in AMI group are higher (D) compare to the control group in two

cohorts. Networks are presented under threshold 0.5 and 0.7. Darker line represents connections under threshold 0.7; lighter line represents connections under threshold 0.5. Degree are presented as mean ± SEM. DCGs, differential connectivity genes.

Figure 5.

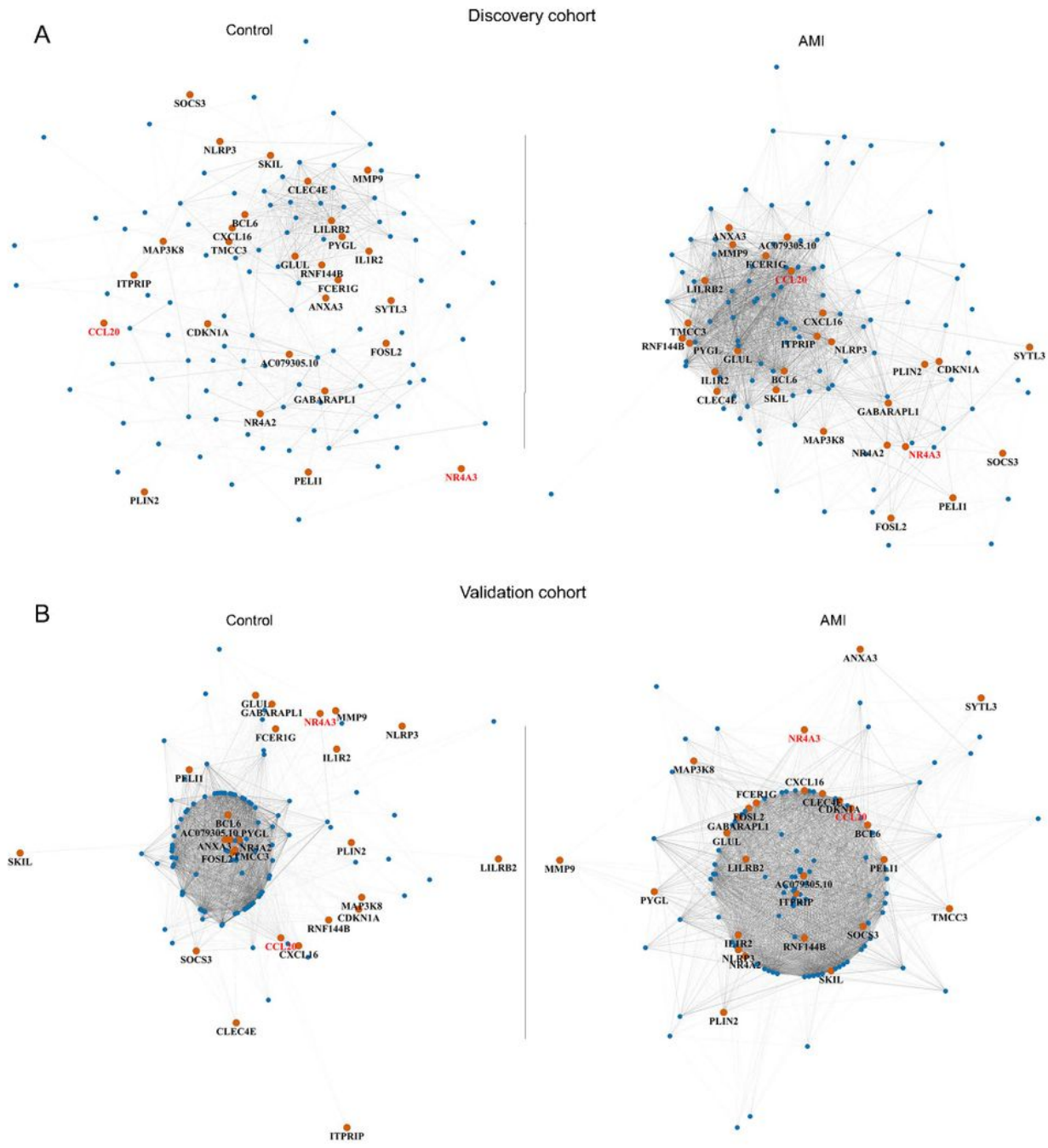

Figure 5

Visualization of DCGs in DEGs' networks. The networks of DEGs in discovery cohort (A) and in the validation cohort (B) indicate that the DCGs participate in distinctive ways in the control group and in the

AMI group. DCGs switch to central functional position of networks and participate in more intricate connections under AMI situation. Yellow nodes indicate the DCGs. Red gene names indicate the hub genes. DEGs, differential expression genes; DCGs, differential connectivity genes.

Figure 6.

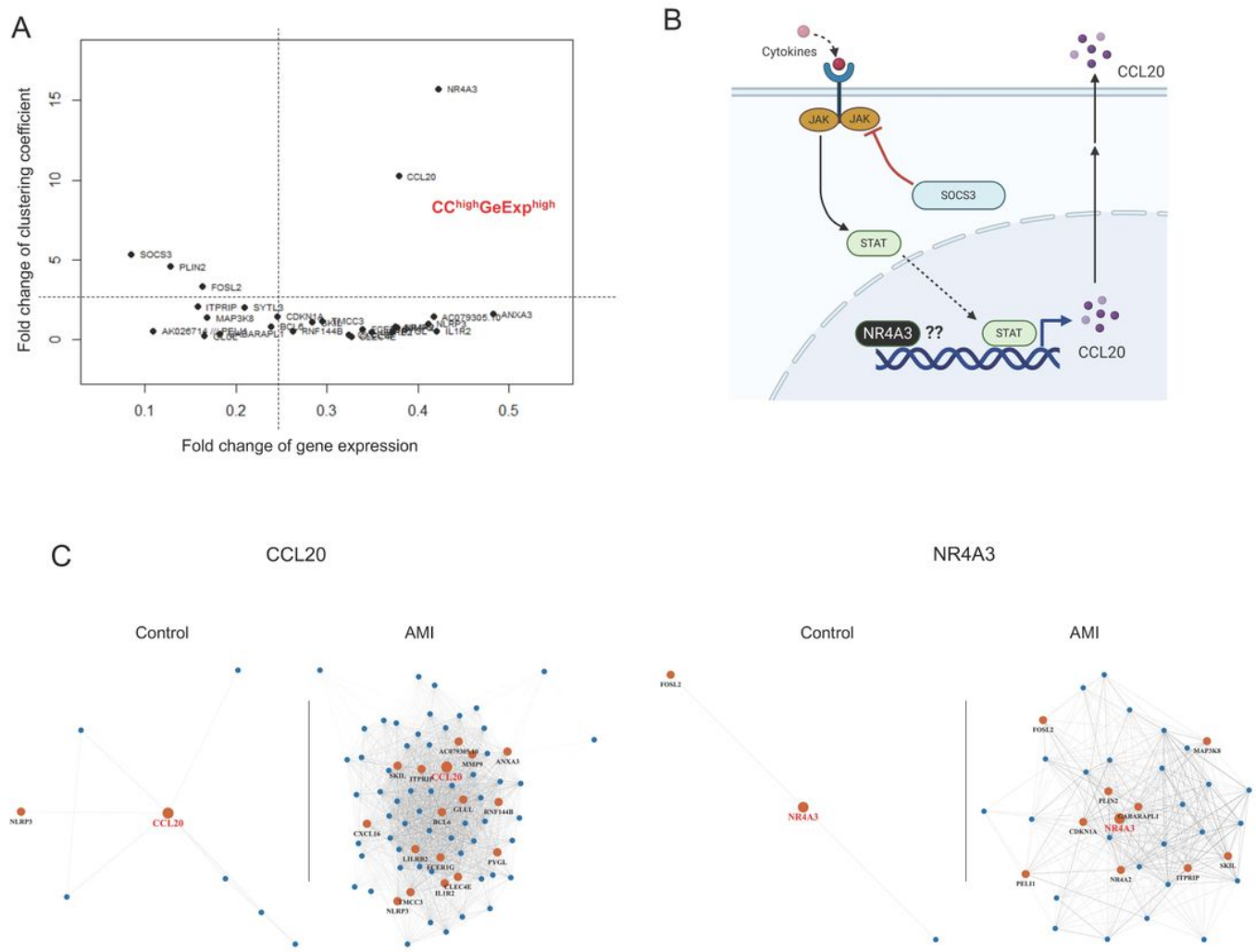

Figure 6

Analysis of gene connection and expression of DCGs in discovery cohort. The analysis of clustering coefficient and gene expression revealed CCL20 and NR4A3 as hub genes (A). The CCL20 is a chemoattractant while NR4A3 is a nuclear factor receptor (B). Subgraphs of CCL20 and NR4A3 substantiate their important roles in AMI development (C). Networks are presented under threshold 0.5 and 0.7. Darker line represents connections under threshold 0.7; lighter line represents connections under threshold 0.5. DCGs, differential connectivity genes; CC, clustering coefficient; GeExp, gene expression.

Figure 7.

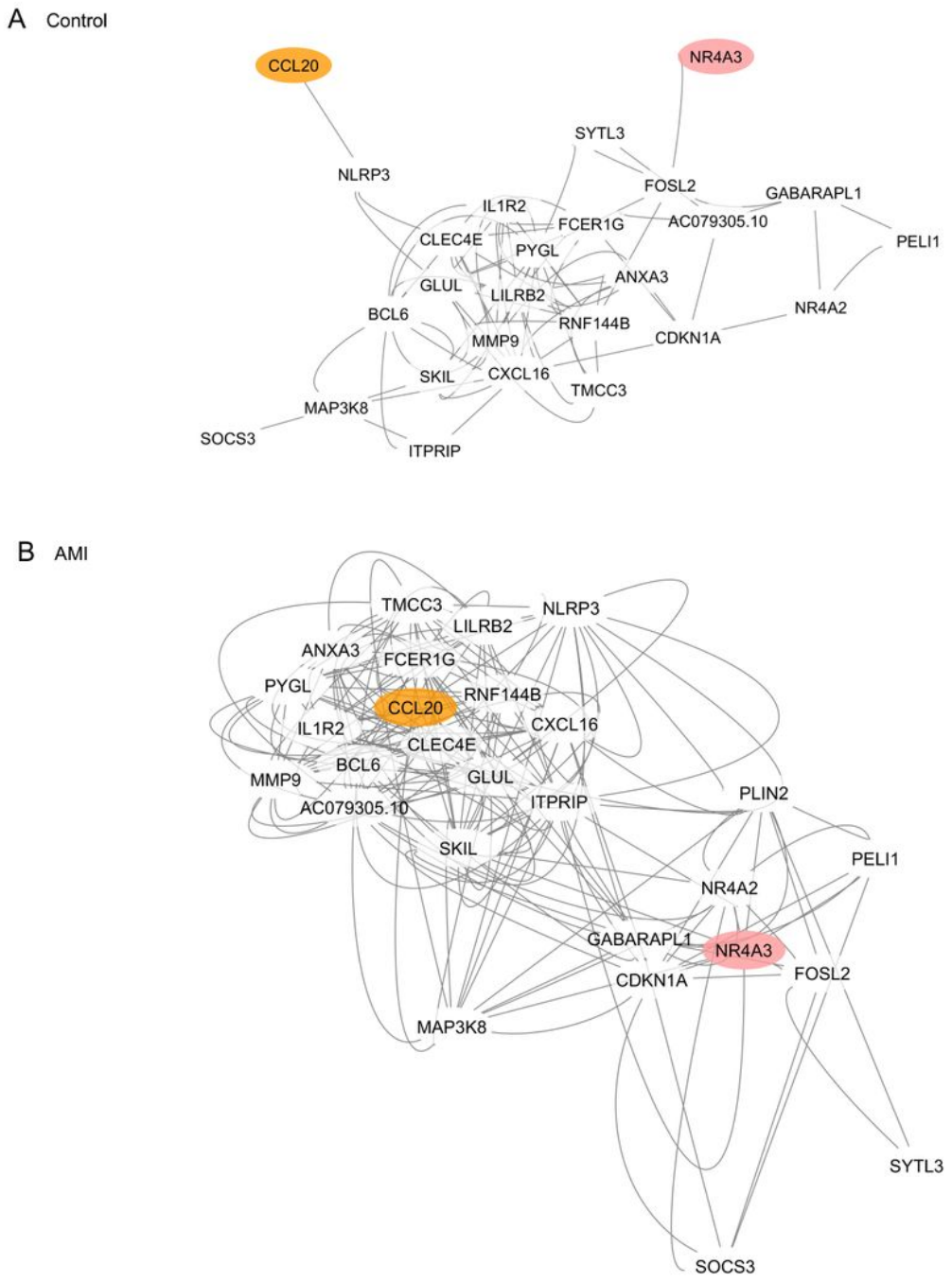

Figure 7

CCL20 cluster and NR4A3 cluster formation in early-stage AMI. CCL20 and NR4A3 stay in the peripheral position of DCGs' network under normal state (A). However, they shift to the primary position of DCGs' network dominating two functional clusters under AMI stimulation (B). DCGs, differential connectivity genes.

Supplementary Files

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

- Tables.docx
- supplementaryfigures.pdf
- supplementarytables.pdf