Original Article

Avicenna Journal of Medical Biotechnology
Volume 13, Issue no. 1, 15-23
http://dx.doi.org/10.18502/ajmb.v13i1.4581

Integrated Network and Gene Ontology Analysis Identifies Key Genes and Pathways for Coronary Artery Diseases

Tejaswini Prakash and Nallur B Ramachandra *

Genetics and Genomics Lab, Department of Studies in Genetics and Genomics, University of Mysore, Karnataka, India

*** Corresponding author:**
Nallur B Ramachandra, Ph.D., Department of Genetics and Genomics, University of Mysore, Manasagangothri, Mysuru, Karnataka, India
Tel: +91 9916344326, +91 9880033687
E-mail:
prakash.tejaswini@gmail.com
nallurbr@gmail.com
Received: 17 Jun 2020
Accepted: 23 Sept 2020

Abstract

Background: The prevalence of Coronary Artery Disease (CAD) in developing countries is on the rise, owing to rapidly changing lifestyle. Therefore, it is imperative that the underlying genetic and molecular mechanisms be understood to develop specific treatment strategies. Comprehensive disease network and Gene Ontology (GO) studies aid in prioritizing potential candidate genes for CAD and also give insights into gene function by establishing gene and disease pathway relationships.

Methods: In the present study, CAD-associated genes were collated from different data sources and protein-protein interaction network was constructed using STRING. Highly interconnected network clusters were inferred and GO analysis was performed.

Results: Interrelation between genes and pathways were analyzed on ClueGO and 38 candidates were identified from 1475 CAD-associated genes, which were significantly enriched in CAD-related pathways such as metabolism and regulation of lipid molecules, platelet activation, macrophage derived foam cell differentiation, and blood coagulation and fibrin clot formation.

Discussion: Integrated network and ontology analysis enables biomarker prioritization for common complex diseases such as CAD. Experimental validation and future studies on the prioritized genes may reveal valuable insights into CAD development mechanism and targeted treatment strategies.

Avicenna J Med Biotech 2021; 13(1): 15-23

Keywords: Coronary artery disease, Fibrin, Gene ontology, Pathway enrichment, Protein interaction maps

Introduction

Most common diseases that are manifested in human populations are complex in nature. Identifying the genes that cause these diseases has been challenging. However, the progress in recent years in high- throughput technology has led to the development of techniques that aid in gaining novel and important insights, resulting in a better understanding of the genetic architecture of complex diseases [1]. Identification of genes and pathways involved in predisposing individuals to complex diseases is crucial to elucidate the pathophysiology of these diseases, which may eventually lead to the development of novel treatments [2]. Coronary Artery Disease (CAD) is one such disease with complex interaction network and pathways which requires much attention for more specific and effective strategies of treatment.

CAD, a complex multifactorial disease, continues to be the leading cause of morbidity and mortality all over the world [3]. India, in particular, has seen a steady raise in the number of CAD cases over the last 40 years [4]. Several risk factors are known to contribute to CAD, such as, diabetes, hypertension, dyslipidemia, lifestyle and genetics [5]. Over the years, multiple genetic studies have identified hundreds of causal and susceptibility loci associated with CAD that have improved our knowledge and understanding of its rudimentary causal factors [6-8]. However, these studies have been successful in explaining only a small proportion of disease etiology [6], suggesting the involvement of many other genes in development of CAD that are yet to be defined. As the search for underlying genetic factors continues, biological validation of all the genes associated with CAD manifestation becomes difficult and impractical. Therefore, integration of high-throughput technology and bioinformatics has gained significant clinical relevance by improving the speed and efficiency of candidate gene discovery and prioritization [9]. Leveraging these methods to understand the molecular mechanisms of CAD might improve our knowledge of biological relevance of CAD-associated genes, particularly in the context of complex networks [10].

In silico analyses using network and gene ontology-

based approaches have been used widely to discover and prioritize potential candidate genes. The huge data generated from genome wide linkage and association or gene expression studies are extensively used to prioritize candidate genes and predict drug targets by means of network based approaches [11,12]. The primary focus of network based approach is to discern the relationship between two components using Protein-Protein Interaction (PPI) patterns. A vast majority of research has gone into combining differential gene expression with network and pathway analysis to delineate the molecular mechanisms underlying CAD [10,13-17]. Tan *et al* [10] used microarray data combined with network analysis to identify key genes and pathways in advanced coronary atherosclerosis. Similarly, Kashyap *et al* [18] established known candidates *APOA1*, *CFTR*, *SRC*, *ICAM1*, *ESR1* and *HNF1A* as hub genes for CAD, Single Vessel Disease (SVD) and Triple Vessel Disease (TVD), respectively using PPI network and Gene Ontology (GO) analysis on differentially expressed genes in CAD subjects and normal controls. More recently, Miao *et al* integrated gene expression changes in CAD subjects with GO annotation and cluster analysis of PPI network, implicating *IL1B*, *ICAM1*, *JUN* and *CCL2* in fluid shear stress, AGE-RAGE signaling pathway, Tumor Necrosis Factor (TNF) and cytokine-cytokine receptor interaction [19]. There is also growing literature where studies have identified important CAD associated genes from existing literature and re-established their roles in CAD pathophysiology using PPI network-based approaches [4,13,17,20].

In view of these studies, an attempt was made to utilize phenotype association, PPI network, cluster analysis and GO annotation to identify highly interacting gene clusters and the interrelation between their biological functions to prioritize plausible candidates from CAD-associated genes.

Materials and Methods

Selection of genes

Genes associated with CAD were retrieved from CardioGenBase (http://www.CardioGenBase.com/) [21]. Collection of data from CardioGenBase was stopped on 26-03-2018. Additionally, Polysearch2 (http://polysearch.ca/) [22] text mining tool was used to extract all associated genes/proteins given the disease term "coronary artery disease". The genes obtained from all these sources were combined and replicates were removed.

Disease association and phenotype enrichment analysis

Web-based Gene Analysis Toolkit (WebGestalt) [23] was used to perform disease association and phenotype enrichment analysis on the retrieved gene list with *Homo sapiens* as the organism of interest. Hypergeometric statistical test was used with Benjamini and Hochberg procedure to determine the false discovery rate. Significance level was set to "Top 10". Phenotype enrichment was established using Mammalian Phenotype Ontology and Human Phenotype Ontology. The genes enriched in phenotype analysis were further used for PPI analysis.

Protein-protein interaction network construction and module analysis

STRING v11.0 (Search Tool for the Retrieval of Interacting Genes/Proteins) (https://string-db.org/) [24] was used to construct PPI network using the gene set enriched in phenotype association analysis. Gene symbols were collated and duplicates were removed, retaining only non-redundant, unique genes. *Homo sapiens* was selected as the organism of interest. Experiments, databases and co-expression were used as active interaction sources, with a median confidence score of >0.4. Rest of the parameters were set to default. The resultant network was visualized in Cytoscape v3.8.0.0 [25].

Module analysis was performed using Cytoscape's Molecular Complex Detection (MCODE) plug-in. MCODE identifies and clusters highly interconnected nodes in a network. PPI network was analyzed and clusters were inferred using the following parameters: node score cut-off - 0.2, node density cut-off - 0.1, K-core - 2, maximum depth - 100. Network scoring was done with a cut-off score of 2 and singly connected nodes were removed from clusters.

Gene ontology and pathway interrelation analysis of PPI network clusters

The genes identified from the top scoring clusters were subjected to function-based categorization to assess their functional association with biological processes, molecular function and cellular components with a significant p-value of <0.05. GO and pathway interrelation analysis of these genes was carried out on ClueGO, a Cytoscape plug-in, using annotations from KEGG. Right-sided hypergeometric test was employed to evaluate enrichment of genes in pathways with significant p-value (<0.05) corrected using Benjamini-Hochberg method.

Results

Selection of genes

A total of 1441 genes were retrieved from CardioGenBase database. PolySearch text-mining source yielded 100 genes with Z-score>1. The two gene sets were combined to get a total of 1475 non-redundant, unique genes which were subjected to disease and phenotype association analysis.

Disease association and phenotype enrichment analysis

Disease association analysis performed on WebGestalt categorized the 1475 genes into 10 highly significant disease phenotypes -vascular diseases, cardiovascular diseases, myocardial ischemic disease, coronary disease, coronary artery disease, arterial occlusive diseases, arteriosclerosis, inflammation, heart diseases and genetic predisposition to diseases (Table 1).

Table 1. Disease association analysis of CAD-associated genes showing top 10 significantly enriched disease categories

Disease	Gene count	p-value	Adj p-value
Vascular diseases	274	0.00E+00	0.00E+00
Cardiovascular diseases	286	0.00E+00	0.00E+00
Myocardial ischemia	217	4.67E-281	3.14E-278
Coronary artery disease	204	1.20E-257	6.05E-255
Coronary disease	203	3.43E-256	1.38E-253
Arterial occlusive diseases	191	1.79E-255	6.02E-253
Arteriosclerosis	188	8.95E-253	2.58E-250
Inflammation	245	9.64E-249	2.43E-246
Heart diseases	218	3.09E-228	6.93E-226
Genetic predisposition to disease	293	8.87E-228	1.79E-225

Phenotype enrichment analysis showed the 1475 genes being enriched in three categories of phenotypic abnormalities; the interactions were represented as Directed Acyclic Graph (DAG)-abnormality of the integument, abnormality of blood and blood forming tissues (183 genes, adjP=9.83e-12), and abnormality of the cardiovascular system (Figure 1). The parent nodes (Shown in black) were further classified into nine significantly enriched phenotypic classes (Shown in red), with a collection of 253 genes.

Protein-protein interaction network construction and module analysis

PPI network was constructed on STRING database

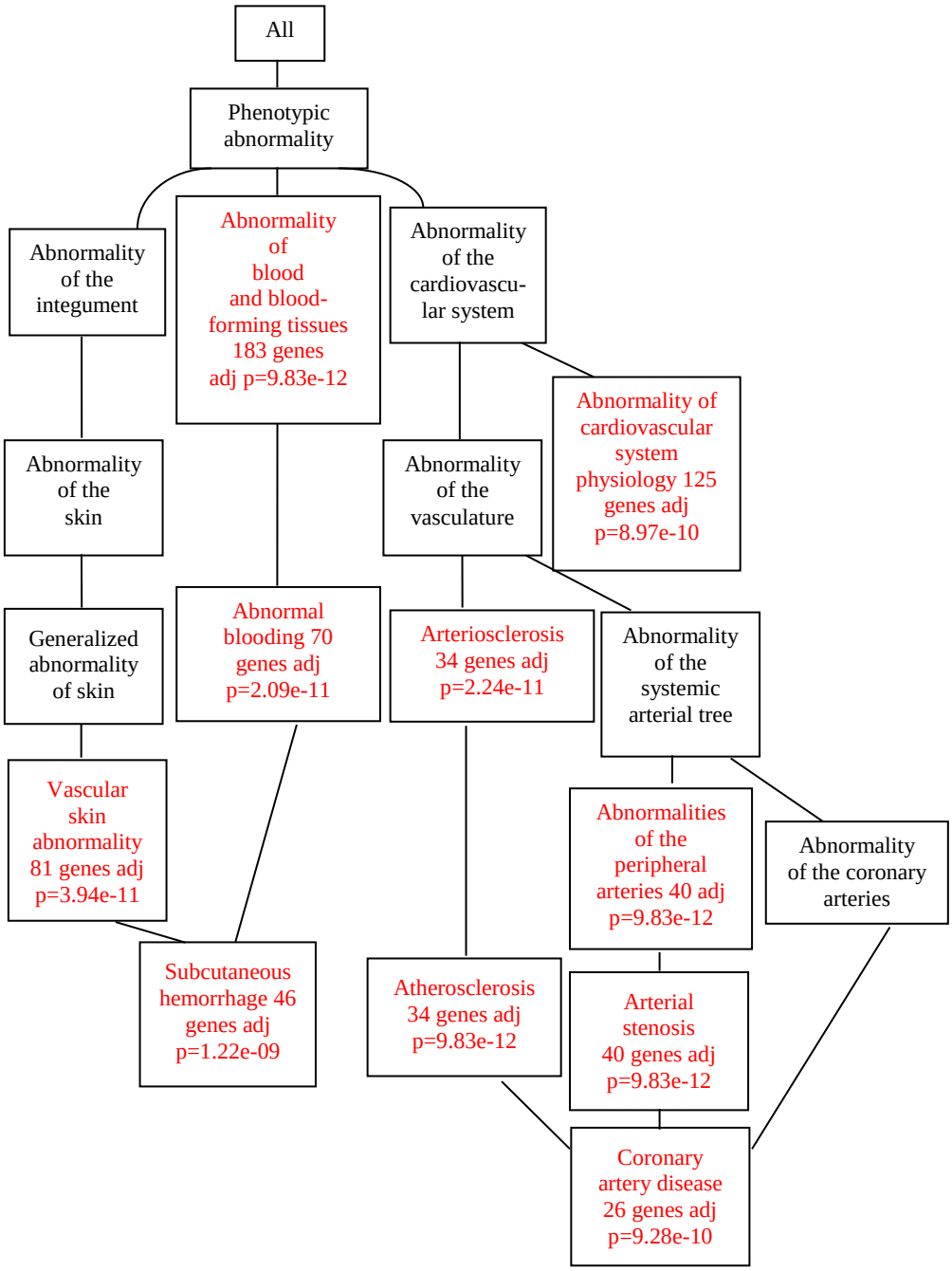

Figure 1. Directed acyclic graph showing phenotype enrichment of 1475 CAD-associated genes. Nodes in black represent non-enriched parent nodes. Nodes in red represent significantly enriched phenotypes.

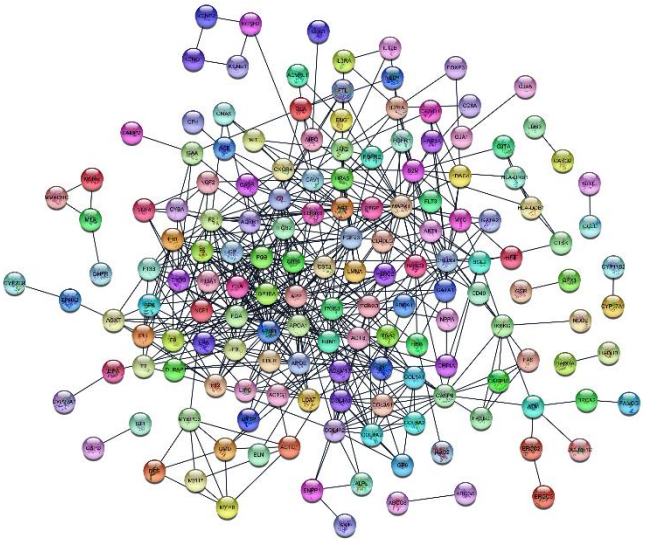

Figure 2. Protein-protein interaction network constructed from genes derived from phenotype enrichment analysis.

(Figure 2) using 253 seed genes collated from significantly enriched phenotypes. The resultant network had 158 nodes with 555 interactions. The network was analyzed for highly interconnected proteins using Cytoscape's MCODE plug-in and 13 clusters of connected nodes were obtained (Table 2), of which the top 3 high scoring clusters (Figures 3A and C) were selected for functional annotation. These clusters comprise 38 genes that can be potentially high-risk candidate genes for CAD.

Gene ontology and pathway interrelation analysis of PPI network clusters

GO analysis and pathway enrichment provided further insights into CAD related molecular mechanisms. The interrelation between pathways and 38 genes inferred from 3 high scoring network clusters was investigated using ClueGO (Figure 4). Significant enrichment was seen in GO terms such as regulation of plasma lipoprotein particle levels, regulation of lipid localization, regulation of lipid storage, plasma lipoprotein particle clearance, lipoprotein particle binding, plasma lipoprotein particle indicating that lipid metabolism and regulation plays a central role in CAD pathophysiology. Additionally, macrophage derived foam cell differentiation, platelet activation, complement and coagulation cascades, blood coagulation and fibrin clot formation, focal adhesion and AGE-RAGE signaling pathway in diabetes complications were also significantly enriched, giving insights into the pathophysiology of CAD. Detailed results of GO analysis are shown in table 3. The plausible biological function of each gene in relation to CAD is summarized in table 4.

Discussion

The major underlying cause for CAD is Atherosclerosis (AS), which is developed from a complicated etiology that involves complex interactions between several genes and proteins [26]. The treatment strategies for AS-related diseases are currently limited to the use of anticoagulant and lipid-lowering drugs. Therefore, furthering our knowledge of AS mechanisms and identifying novel targets for therapeutics holds significant clinical value. Although Genome-Wide Association Studies (GWAS) and the more recent high throughput technologies such as, Next Generation Sequencing (NGS) and genome expression profiling have established association of hundreds of probable candidate genes to CAD, identifying specific disease-causing genes and the mechanism through which they contribute to disease progression has been challenging. Also, experimental validation of vast number of genes is time consuming and laborious [4]. Systems biology offers a feasible approach for prioritization of genes and elucidation of their roles in manifestation of disease. Genes and proteins that are most relevant to disease can be mined and analyzed from large public databases using *in silico* analyses. Thus, it is deemed to be a promising technology to discover novel targets for treatment.

Table 2. Highly interconnected genes in network clusters derived from module analysis using Cytoscape's MCODE plug-in

Cluster	Cluster score	Nodes, edges	Genes
1	12	12, 66	FBN1, APOE, FGA, CST3, APP, C4A, FGF23, F5, CP, APOB, APOA1, C3
2	7.4	21, 74	F9, ITGB3, F2, HRAS, F12, COL4A2, F11, COL4A1, CASP8, LMNA, F10, GP1BA, ITGA2, MAPK1, COL1A1, COL1A2, COL3A1, COL5A2, GP9, AGXT, ITGA2B
3	5	5, 10	LPL, LCAT, LPA, LDLR, LIPC
4	5	5, 10	NCF1, NCF2, CYBA, NCF4, CYBB
5	4	4, 6	AGT, AGTR1, EDNRB, CASR
6	4	4, 6	ACTB, ACTG1, DMD, ACTC1
7	3.667	7, 11	MYC, CREBBP, CASP10, IKBKG, BCL2, FASLG, FAS
8	3.333	4, 5	FGB, F13A1, F8, F13B
9	3.333	4, 5	DES, MYBPC3, MYH7, MYH6
10	3	3, 3	MPO, FTL, GLA
11	3	3, 3	HLA-DRB1, HLA-DQB1, CIITA
12	3	3, 3	MTRR, MTR, MMACHC
13	2.8	6, 7	IL12B, FGFR1, IL2RA, IL3RA, AKT1, GATA1

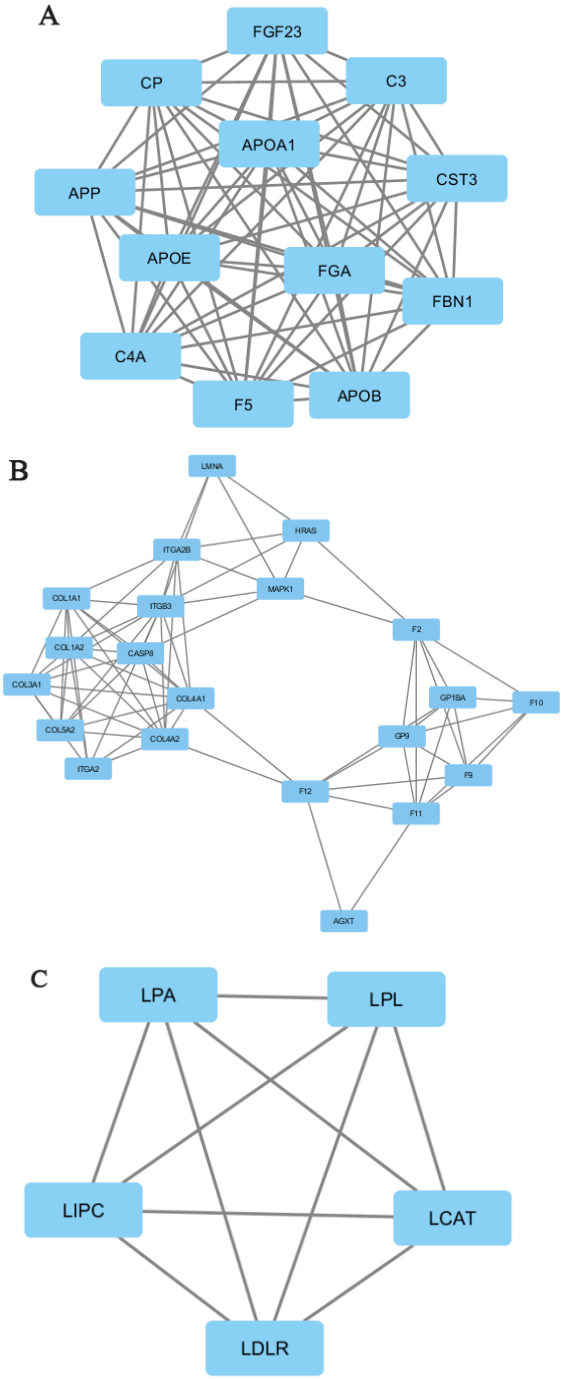

Figure 3. A-C) Top scoring network clusters from module analysis using Cytoscape's MCODE plug-in.

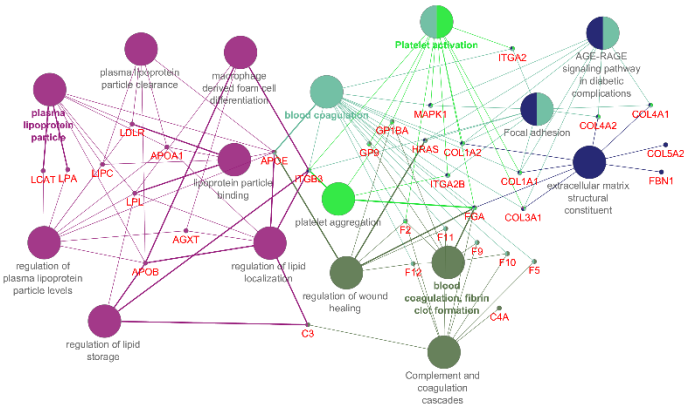

Figure 4. Pathway interrelation analysis of genes derived from top scoring network clusters.

Recent years have seen remarkable progress in biomarker identification for CAD using network-based approaches. Genes associated with CAD and related phenotypes have been used to construct PPI networks and identify selected hub genes that are possibly most relevant to disease manifestation and progression. In the present study, an integrated approach of phenotype association, PPI network and gene ontology analysis was employed to identify 38 plausible high-risk candidate genes from 1475 CAD-associated genes. GO analysis and pathway enrichment analysis of 38 candidate genes demonstrated several of them are established CAD driver genes (*AGXT*, *APOA1*, *APOB*, *APOE*, *LCAT*, *LDLR*, *LIPC*, *LPA* and *LPL*) and participate in the lipid metabolic processes. Significant enrichment was observed in macrophage derived foam cell differentiation (*AGXT*, *APOB*, *ITGB3* and *LPL*). Macrophages are critical players in the development of atherosclerosis as they contribute to plaque development, regulate and maintain vascular inflammatory response, and stimulate thrombus formation, making them desirable therapeutic targets [27]. Blood coagulation and fibrin clot formation (*F10*, *F11*, *F12*, *F2*, *F9*, *FGA*, *GP1BA*, *GP9*) and regulation of wound healing process (*APOE*, *F11*, *F12*, *F2*, *FGA*, *GP1BA*, *HRAS*) were also significantly enriched. GO annotation also showed enrichment of genes in Extracellular Matrix (ECM) structural constituents which play crucial roles in cell integrity, adhesion and communication. ECM governs the proper functioning of different organs and changes in its components have been implicated in CAD mechanisms such as atherothrombosis, plaque rupture and calcification. Vascular smooth muscle cells and macrophages functioning are also shown to be influenced by matrix constituents such as collagens [28]. Previous network analysis by Zhao *et al* [16] has also established the importance of ECM proteins in CAD development and its link with complement and coagulation cascades, and inflammation.

The study also identified less well-studied possible candidates for CAD in addition to established CAD driver genes such as, *APP* and *FBN1*. Amyloid Precursor Protein (*APP*) encodes amyloid-β (Aβ) peptides which are main drivers of Alzheimer's Disease (AD), a progressive neurodegenerative disorder. The biological function of APP with respect to Alzheimer's disease has been scrutinized extensively over the last 30 years [29]. Interestingly, in addition to being expressed in hu-

Table 3. Gene ontology and pathway enrichment analysis of 38 identified candidate genes. Enrichment is classified into pathways, biological processes, molecular functions and cellular components

GO ID	GO term	Term p-value	Term p-value corrected with benjamini-hochberg	Group p-value	Group p-value corrected with benjamini-hochberg	Associated genes found
Pathway						
KEGG:04610	Complement and coagulation cascades	4.00E-14	2.18E-13	2.31E-16	5.77E-16	C3, C4A, F10, F11, F12, F2, F5, F9, FGA
KEGG:04510	Focal adhesion	9.40E-11	2.82E-10	1.08E-20	5.40E-20	COL1A1, COL1A2, COL4A1, COL4A2, HRAS, ITGA2, ITGA2B, ITGB3, MAPK1
KEGG:04611	Platelet activation	3.15E-16	9.45E-15	1.08E-20	5.40E-20	COL1A1, COL1A2, COL3A1, F2, FGA, GP1BA, GP9, ITGA2, ITGA2B, ITGB3, MAPK1
KEGG:04933	AGE-RAGE signaling pathway in diabetic complications	6.62E-10	1.81E-09	1.08E-20	5.40E-20	COL1A1, COL1A2, COL3A1, COL4A1, COL4A2, HRAS, MAPK1
Biological process						
GO:0072378	Blood coagulation, fibrin clot formation	7.34E-16	1.47E-14	2.31E-16	5.77E-16	F10, F11, F12, F2, F9, FGA, GP1BA, GP9
GO:0061041	Regulation of wound healing	2.02E-08	3.28E-08	2.31E-16	5.77E-16	APOE, F11, F12, F2, FGA, GP1BA, HRAS
GO:0097006	Regulation of plasma lipoprotein particle levels	1.37E-13	6.31E-13	3.19E-11	3.19E-11	AGXT, APOA1, APOB, APOE, LCAT, LDLR, LIPC, LPA, LPL
GO:1905952	Regulation of lipid localization	2.20E-08	3.47E-08	3.19E-11	3.19E-11	AGXT, APOA1, APOB, APOE, C3, ITGB3, LPL
GO:0010883	Regulation of lipid storage	3.16E-06	3.58E-06	3.19E-11	3.19E-11	APOB, C3, ITGB3, LPL
GO:0010742	Macrophage derived foam cell differentiation	9.32E-07	1.19E-06	3.19E-11	3.19E-11	AGXT, APOB, ITGB3, LPL
GO:0034381	Plasma lipoprotein particle clearance	8.39E-08	1.20E-07	5.25E-13	6.56E-13	APOA1, APOB, APOE, LDLR, LIPC
Molecular function						
GO:0005201	Extracellular matrix structural constituent	1.23E-09	3.06E-09	1.08E-20	5.40E-20	COL1A1, COL1A2, COL3A1, COL4A1, COL4A2, COL5A2, FBN1, FGA
GO:0071813	Lipoprotein particle binding	4.12E-09	8.25E-09	5.25E-13	6.56E-13	APOA1, APOE, LDLR, LIPC, LPL
Cellular component						
GO:0034358	Plasma lipoprotein particle	4.25E-15	4.25E-14	5.25E-13	6.56E-13	APOA1, APOB, APOE, LCAT, LDLR, LIPC, LPA, LPL

man brain, kidney, and platelets, APP is also expressed in vascular endothelium of coronary, cerebral and peripheral blood vessels. Additionally, the endothelial cells of large conduit arteries, resistance arteries, and microvessels also show the expression and activity of enzymes responsible for proteolytic cleavage of APP [29]. In a study of apolipoprotein E knockout (ApoE-/-) mice, Tibolla *et al* [30] and, Austin and Combs [31] demonstrated accelerated aortic atherosclerotic development and endothelial dysfunction induced by APP overexpression. Austin and Combs also went on to show that in ApoE-/- mice and AD patients, there was overexpression of APP mediated monocyte adhesion to the endothelium, implying that vascular dysfunction associated with AD and atherosclerosis involves endothelial APP function. However, the molecular mechanisms and physiological outcomes underlying APP action are incompletely understood, thereby making it an interesting candidate for future study. Similarly, the effect of impaired *FBN1* action was studied using mouse models by Van der Donckt *et al* [32]. *FBN1* mutations have been established to cause Marfan syndrome, affecting the connective tissues that support the body's joints and organs. Van der Donckt *et al* [32] showed that C1039G +/− mutation in mice resulted in increased arterial stiffness owing to fragmentation of the elastic fibers of the vessel wall. It was also seen that extremely unstable plaques were developed due to the mutation leading to sporadic ruptures. Although several single nucleotide polymorphisms and mutations have been mapped to *FBN1*, the precise mechanistic action through which it exerts arterial stiffness and aids atherosclerotic progression has not been established.

Pathway enrichment analysis performed on the prioritized genes asserted some results of previous studies. Over-representation of genes was seen in AGE-

Table 4. Plausible function of 38 prioritized candidate genes for CAD

Gene symbol	Gene name	Potential function in CAD
FBN1	Fibrillin 1	Essential for proper formation of extracellular matrix (ECM), including the biogenesis and maintenance of elastic fibers
APOE	Apolipoprotein E	Efficient uptake of lipoprotein particles by hepatic cells, stimulation of cholesterol efflux from macrophage foam cells in the atherosclerotic lesion, and regulation of immune and inflammatory responses
FGA	Fibrinogen alpha chain	Blood clot formation (Coagulation)
CST3	Cystatin C	Inhibition of cathepsins that degrade extracellular matrix elastin and collagen
APP	Amyloid beta precursor protein	Formation of arterial plaque
C4A	Complement C4A	Initial activation of classical complement pathway
FGF23	Fibroblast growth factor 23	Arterial calcification and stiffness
F5	Coagulation factor V	Regulation of blood coagulation
CP	Ceruloplasmin	Decreases nitric oxide bioavailability in blood
APOB	Apolipoprotein B	LDL cholesterol and lipoprotein organization
APOA1	Apolipoprotein A1	Triglyceride-rich lipoproteins organization
C3	Complement C3	Activation of the complement system
F9	Coagulation factor IX	Participates in the intrinsic pathway of blood coagulation
ITGB3	Integrin subunit beta 3	Receptor for FBN1, functions in clot formation
F2	Coagulation factor II, thrombin	Participates in blood clotting process
HRAS	HRas proto-oncogene, GTPase	Mediates VSMC proliferation during vascular injury
F12	Coagulation factor XII	Participates in blood clotting process
COL4A2	Collagen type IV alpha 2 chain	Cellular proliferation and vascular remodeling
F11	Coagulation factor XI	Participates in blood coagulation process
COL4A1	Collagen type IV alpha 1 chain	Cellular proliferation and vascular remodeling
CASP8	Caspase 8	Uncertain
LMNA	Lamin A	Regulation of nuclear structure and cell functions
F10	Coagulation factor X	Participates in blood coagulation process
GP1BA	Glycoprotein Ib platelet subunit alpha	Plays a role in blood clotting
ITGA2	Integrin subunit alpha 2	Mediates the adhesion of platelets and other cell types to the extracellular matrix
MAPK1	Mitogen-activated protein kinase 1	Signaling cascades during foam cell formation
COL1A1	Collagen type I alpha 1 chain	Constitutes the major protein of plaque extracellular matrix
COL1A2	Collagen type I alpha 2 chain	
COL3A1	Collagen type III alpha 1 chain	
COL5A2	Collagen type V alpha 2 chain	
GP9	Glycoprotein IX platelet	Formation of fibrin clot
AGXT	Alanine--glyoxylate and serine--pyruvate aminotransferase	Uncertain
ITGA2B	Integrin subunit alpha 2b	Crucial role in the blood coagulation system, by mediating platelet aggregation
LPL	Lipoprotein lipase	Lipid clearance from the blood stream, lipid utilization and storage
LCAT	Lecithin-cholesterol acyltransferase	Esterification of cholesterol required for cholesterol transport
LPA	Lipoprotein (A)	Contributes to sustained plaque development in atherosclerosis
LDLR	Low density lipoprotein receptor	Binds LDL, the major cholesterol-carrying lipoprotein of plasma, and transports it into cells
LIPC	Lipase C	Catalyzes the hydrolysis of triglycerides and phospholipids present in circulating plasma lipoproteins

RAGE signaling pathway in diabetes complications. Diabetes Mellitus (DM) is a well-established risk factor for CAD and is known to promote vascular calcification through mechanisms of hyperglycemia, hypercalcemia and oxidative stress. Several studies have implicated the AGE/RAGE signaling pathway in accelerating vascular calcification in DM patients which increases the risk of atherosclerotic plaque rupture, resulting in the possibility of heart attack or stroke [33-35]. AGE-RAGE signaling pathway is also a key component of inflammation and immune response. Miao *et al* [19] identified 413 differentially expressed genes using two datasets of gene expression studies in CAD subjects and controls and emphasized the contribution of AGE-RAGE signaling pathways in the pathogenesis of CAD. Significant enrichment was also observed in focal adhesion, platelet activation pathways and complement-coagulation cascades. Focal adhesion mole-

cules are important mediators of interaction between blood vessels and circulating leukocytes [36]. They control adhesion dynamics and influence the interaction between platelets and endothelium. They also play a key role in vascular dysfunction and tissue injury in atherosclerosis [37]. An integrative network and pathway analysis by Chan *et al* [38] showed that the CAD and diabetes molecular pathways share focal adhesion as one of the key pathways in different ethnicities. Several key processes involved in the development and progression of atherosclerosis are shown to be influenced by the complement system through *in vitro* studies. It is also said to influence the extent of thrombus formation when the atherosclerotic plaque ruptures [39]. As demonstrated by the pathway interrelation analysis, these pathways function with a complex network of interconnected genes. Therefore, dysfunction or disruption of these important genes may result in a chain of functional pathological outcomes.

Conclusion

The current study focused on prioritizing CAD-associated genes and identified 38 potential high-risk CAD candidates using a combination of network and pathway analysis in a stepwise manner. These genes were found to be involved in key biological processes and pathways that contribute to the pathogenesis of CAD. Therefore, it can be concluded that if any of these genes in the network fails to function normally owing to variations and/or mutations, the downstream processes can be disrupted resulting in development of CAD. The strength of the study lies in ontology-based gene filtration. However, there are certain limitations to this study. Since the knowledge of CAD genetics and mechanisms is still an ongoing process, network construction could be incomplete or certain interactions could be missing as it is derived from published reports and these results need to be further validated with appropriate wet-lab experiments to delineate the biological mechanisms.

Acknowledgement

We are grateful to Department of Studies in Genetics and Genomics, University of Mysore for providing facility to conduct this work. We also thank members of the Genetics and Genomics lab, Department of Studies in Genetics and Genomics for their support and encouragement. We thank Institution of Excellence for providing fellowship to Ms. Tejaswini Prakash.

Conflict of Interest

The authors declare no competing financial interests.

References

1. Lightbody G, Haberland V, Browne F, Taggart L, Zheng H, Parkes E, et al. Review of applications of high-throughput sequencing in personalized medicine: barriers and facilitators of future progress in research and clinical application. Brief Bioinform 2019;20(5):1795-811.
2. Hofker MH, Fu J, Wijmenga C. The genome revolution and its role in understanding complex diseases. Biochim Biophys Acta 2014;1842(10):1889-95.
3. Mc Namara K, Alzubaidi H, Jackson JK. Cardiovascular disease as a leading cause of death: how are pharmacists getting involved? Integr Pharm Res Pract 2019;4(8):1-11.
4. Ghatge M, Nair J, Sharma A, Vangala RK. Integrative gene ontology and network analysis of coronary artery disease associated genes suggests potential role of ErbB pathway gene EGFR. Mol Med Rep 2018;17(3):4253-64.
5. Hajar R. Risk factors for coronary artery disease: Historical perspectives. Heart Views 2017;18(3):109-14.
6. CARDIoGRAMplusC4D Consortium; Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, et al: Large-scale association analysis identifies new risk loci for coronary artery disease. Nature Genetics 2013;45:25-33.
7. Mehta NN. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Circ Cardiovasc Genet 2011;4(3):327-9.
8. Arvind P, Jayashree S, Jambunathan S, Nair J, Kakkar VV. Understanding gene expression in coronary artery disease through global profiling, network analysis and independent validation of key candidate genes. J Genet 2015;94(4):601-10.
9. Doncheva NT, Kacprowski T, Albrecht M. Recent approaches to the prioritization of candidate disease genes. Wiley Interdiscip Rev Syst Biol Med 2012;4(5):429-42.
10. Tan X, Zhang X, Pan L, Tian X, Dong P. Identification of key pathways and genes in advanced coronary atherosclerosis using bioinformatics analysis. Biomed Res Int 2017;2017:4323496.
11. Kaimal V, Sardana D, Bardes EE, Gudivada RC, Chen J, Jegga AG. Integrative systems biology approaches to identify and prioritize disease and drug candidate genes. Methods Mol Biol 2011;700:241-59.
12. Chen BS, Yang SK, Lan CY, Chuang YJ. A systems biology approach to construct the gene regulatory network of systemic inflammation via microarray and databases mining. BMC Med Genomics 2008;1:46.
13. Nair J, Ghatge M, Kakkar VV, Shanker J. Network analysis of inflammatory genes and their transcriptional regulators in coronary artery disease. PLoS One 2014;9(4): e94328.
14. Hofree M, Shen JP, Carter H, Ideker T. Network-based stratification of tumor mutations. Nat Methods 2013;10 (11):1108-15.
15. Makinen VP, Civelek M, Meng Q, Zhang B, Zhu J, Levian C, et al. Integrative genomics reveals novel molecular pathways and gene networks for coronary artery disease. PLoS Genet 2014;10(7):e1004502.
16. Zhao Y, Chen J, Freudenberg JM, Meng Q, Rajpal DK, Yang X. Network-based identification and prioritization of key regulators of coronary artery disease loci. Arterioscler Thromb Vasc Biol 2016;36(5):928-41.

17. Sarajlic A, Janjić V, Stojković N, Radak D, Przulj N. Network topology reveals key cardiovascular disease genes. PLoS One 2013;8(8):e71537.
18. Kashyap S, Kumar S, Agarwal V, Misra DP, Phadke SR, Kapoor A. Protein protein interaction network analysis of differentially expressed genes to understand involved biological processes in coronary artery disease and its different severity. Gene Reports 2018;12:50-60.
19. Miao L, Yin RX, Huang F, Yang S, Chen WX, Wu JZ. Integrated analysis of gene expression changes associated with coronary artery disease. Lipids Health Disease 2019;18(1):92.
20. Zhang L, Li X, Tai J, Li W, Chen L. Predicting candidate genes based on combined network topological features: a case study in coronary artery disease. PLoS One 2012;7 (6):e39542.
21. Alexandar V, Nayar PG, Murugesan R, Beaulah M, Darshana P, Ahmed SSSJ. CardioGenBase: A literature based multi-omics database for major cardiovascular diseases. PLoS One 2015;10(12):e0143188.
22. Liu Y, Liang Y, Wishart DS. PolySearch 2.0: A significantly improved text-mining system for discovering associations between human diseases, genes, drugs, metabolites, toxins, and more. Nucleic Acids Res 2015;43 (W1):W535-42.
23. Wang J, Duncan D, Shi Z, Zhang B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. Nucleic Acids Res 2013;41(Web Server issue):W77-83.
24. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019;47(D1):D607-D613.
25. Shannon P, Markiel A, Ozier O, Baliga NS, Wang AT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13(11):2498-504.
26. Zhang YM, Meng LB, Yu SJ, Ma DX. Identification of potential crucial genes in monocytes for atherosclerosis using bioinformatics analysis. J Int Med Res 2020;48(4): 0300060520909277.
27. Barrett TJ. Macrophages in Atherosclerosis Regression. Arteriosclerosis, Thrombosis, Vascular Biology 2020;40 (1):20-33.
28. Chistiakov DA, Sobenin IA, Orekhov AN. Vascular extracellular matrix in atherosclerosis. Cardiol Rev 2013; 21:270-88.
29. d'Uscio LV, He T, Katusic ZS. Expression and processing of amyloid precursor protein in vascular endothelium. Physiology (Bethesda) 2017;32(1):20-32.
30. Tibolla G, Norata GD, Meda C, Arnaboldi L, Uboldi P, Piazza F, et al. Increased atherosclerosis and vascular inflammation in APP transgenic mice with apolipoprotein E deficiency. Atherosclerosis 2010;210(1):78-87.
31. Austin SA, Combs CK. Amyloid precursor protein mediates monocyte adhesion in AD tissue and apoE(-)/(-) mice. Neurobiology Aging 2010;31(11):1854-66.
32. Van der Donckt C, Van Herck JL, Schrijvers DM, Vanhoutte G, Verhoye M, Blockx I, et al. Elastin fragmentation in atherosclerotic mice leads to intraplaque neovascularization, plaque rupture, myocardial infarction, stroke, and sudden death. Eur Heart J 2015;36(17):1049-58.
33. Kay AM, Simpson CL, Stewart JA Jr. The role of AGE/ RAGE signaling in diabetes-mediated vascular calcification. J Diabetes Res 2016;2016(4):1-8.
34. Liu W, Zhang Y, Yu CM, Ji QW, Cai M, Zhao YX, et al. Current understanding of coronary artery calcification. J Geriatr Cardiol 2015;12(6):668-75.
35. Senatus LM, Schmidt AM. The AGE-RAGE axis: implications for age-associated arterial diseases. Front Genet 2017;8:187.
36. Charchar FJ, Bloomer LD, Barnes TA, Cowley MJ, Nelson CP, Wang Y, et al. Inheritance of coronary artery disease in men: An analysis of the role. Lancet 2012;379 (9819):915-22.
37. Krieglstein CF, Granger DN. Adhesion molecules and their role in vascular disease. Am J Hypertens 2001;14 (6Pt 2):44S-54S.
38. Chan KH, Huang YT, Meng Q, Wu C, Reiner A, Sobel EM, et al. Shared molecular pathways, and gene networks for cardiovascular disease, and type 2 diabetes mellitus in women across diverse ethnicities. Circ Cardiovasc Genet 2014;7(6):911-9.
39. Carter AM. Complement activation: an emerging player in the pathogenesis of cardiovascular disease. Scientifica (Cairo) 2012;2012:402783.