European Review for Medical and Pharmacological Sciences 2015; 19: 1092-1100

Common genes in coronary artery disease from Europe, Asia and North America regardless of race and lifestyle

R. GONG[1], M.-H. CHEN[2], L.-S. PENG[1], S.-L. WEI[1]

[1]Department of Gerontology, and [2]Department of Cardiothoracic Surgery; The Third Municipal Hospital of Shijiazhuang City, Hebei, China

Abstract. – OBJECTIVE: Coronary artery disease (CAD) remains one of the major causes of death worldwide. Despite considerable advances in the prevention and treatment of CAD, its complications, morbidity and mortality still remain very high, and vary widely across different ethnic groups.

MATERIALS AND METHODS: To detect genes involved in the development of CAD, we collected gene expression studies in the blood samples of CAD patients from different continents by searching the Gene Expression Omnibus database (GEO), performed a comparative analysis of gene expression between CAD patients and normal controls (NC) in each continent and identified the common set of differentially expressed genes (DEGs) between CAD patients and NC across different continents. PPI networks of the common set of DEGs were established by Cytoscape software to understand their biological role in CAD.

RESULTS: A total of 575, 868 and 476 genes were identified to be significantly differentially expressed between CAD patients and NC in Asia, Europe and North America. 24 genes were found common in three different continents, and 6 genes were previously linked to CAD or atherosclerosis. In the PPIs network the significant hub proteins contained IRF4 (Degree = 23), PLAUR (Degree = 17) and HIST1H2AE (Degree = 15).

CONCLUSIONS: Not only did we detect gene expression differences in the blood samples between CAD and NC in Asia, Europe and North America population, but analysis of the three population groups revealed a common set of 24 genes regardless of differences related to race, ethnicity, lifestyle, and environmental factor which may provide key factors to understand the pathogenesis of CAD and lead to development of diagnostic markers and/or effective therapeutic strategies.

Key Words:
Coronary artery disease, Different continents, Gene expression, Common genes.

Introduction

Coronary artery disease (CAD), as a complex chronic disease affected by many genetic and environmental factors[1,2] as well as lifestyle patterns, remains one of the major causes of death. Despite considerable advances in the prevention and treatment of CAD, its complications, morbidity and mortality still remain very high[3], and vary widely across different ethnic groups. The South Asian (SA), including Indian, Pakistan, Bangladesh, or Sri Lankan ancestry, is believed to have a higher risk of CAD than other ethnic groups[4,5] possibly due to some specific infections such as M. tuberculosis and CMV[6]. Nevertheless, the association of self reported race/ethnicity with CAD is complicated and poorly understood due to the heterogeneity within racial/ethnic groups.

Recently large-scale gene expression-profiling studies, which could display genetic contribution and patterns of altered gene expression related to CAD, have provided valuable information regarding the molecular environment of CAD across different continents[7-10]. Identifying the common mechanisms underlying differential manifestations of CAD could potentially identify novel targets for prevention and treatment of CAD.

In this study, we suppose that CAD is mediated by a common set of genes regardless of differences related to race, ethnicity, lifestyle, and/or environmental factors. We analyze and compare gene expression signatures of CAD populations of different ancestral continents to identify a common set of genes, which may provide key factors to understand pathogenesis of CAD and lead to development of diagnostic markers and/or effective therapeutic strategies.

Corresponding Author: Rui Gong, MD; e-mail: rui_gong619@163.com

Materials and Methods

Identification of Eligible Gene Expression Profiles in CAD

Gene expression studies in the blood samples of CAD patients across different continents were identified by searching the Gene Expression Omnibus database (GEO, http://www.ncbi.nlm.nih.gov/geo)[11], and the following key words and their combinations were used: "coronary artery disease, gene expression, microarray". We only adopted the original experimental articles that analyzed the gene expression profiling between CAD and normal control (NC) blood. Non-human studies, review articles and integrated analysis of expression profiles were not considered.

Data Preprocessing

The heterogeneity among varying microarray data arising from different microarray platforms, gene nomenclature and clinical samples makes it difficult to compare the gene expression data directly. In terms of current issues, a global normalization approach should be included to minimize the inconsistency. Consequently, we preprocessed the raw microarray data of each study by Quantile normalization and log2 transformation to obtain intensity values.

Statistical Analysis

The MATrix LABoratory (MATLAB) software was used to identify differently expressed genes (DEGs) between CAD and NC blood. A gene specific *t*-test was carried out, and then *p*-value of individual microarray study was calculated. For microarray studies from the same continent, Fisher's method was used to combine *p*-value of individual microarray studies. The Genes with p-value < 0.05 were selected as the significantly DEGs.

Functional Annotation of DEGs from Different Continents

Gene Ontology (GO) was performed to interpret biological roles of DEGs indifferent continents. GO would display functional annotation and classification of analyzed gene sets. Furthermore the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was also performed to detect the potential pathway of DEGs. KEGG pathway database is a recognized and comprehensive database, which could help to represent knowledge on the molecular interaction and reaction network for metabolism, cellular processes, human diseases and others[12]. The GENECODIS[13], a web-accessible program, was used for those functional analysis.

Identification of the Common Set of Genes Across Different Continents

We compared DEGs from different continents, and identified the common set of genes across different continents. In order to understand the biological role of the common set of genes in CAD, we performed the protein-protein interactions (PPIs) analysis. PPI analysis could display protein functions at the molecular level and the rules of cellular activities including growth, development, metabolism, differentiation and apoptosis[14]. We adopted Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) (http://www.string-db.org/), which is a database of known and predicted protein interactions to construct the PPI network, then visualized the distribution characteristics of the common set of genes in the PPI network with Cytoscape software[15].

Results

Detecting DEGs in CAD from Different Continents

In this study, we collected a total of 4 CAD expression profiles derived from the peripheral blood according to the inclusion criteria, among which 1 CAD expression profile was from Europe (GSE12288), 1 from Asia (GSE42148) and 2 from North America (GSE20680, GSE20681). For the purpose of global normalization, we preprocessed the raw microarray data by Quantile normalization and log2 transformation to obtain intensity values for each probe in the gene expression profiling. MATLAB software was adopted to identify DEGs between CAD and NC samples. Finally, 575, 868 and 476 genes were regarded as significantly differentially expressed in samples of CAD compared with NC under the threshold of p-value < 0.05 in Asia, Europe and North America (Table SI, Table SII, Table SIII).

Functional Annotation of DEGs Indifferent Continents

To delineate the predominant biological theme of DEGs in CAD from different continents, we conducted GO categories and KEGG pathway enrichment analysis by web-based software

GENECODIS (Figures 1, 2, and 3). Despite distinct gene expression patterns, some common functions was involved in CAD across different continents, such as signal transduction, blood coagulation, cell adhesion, immune response, phospholipid biosynthetic process, lipid metabolic process and infections.

Identification of the Common Set of Genes Across Different Continents

24 genes were found differentially expressed in CAD for the cases from Europe, Asia and North America populations together (Figure 4, Table I). Six genes were previously linked to CAD or atherosclerosis. PPI networks of the common set of genes across different continents were established by Cytoscape software to understand their biological role in CAD. In the PPIs network the nodes with high degree are defined as hub proteins, and degrees are defined to measure how many neighbors a node directly connect to. The significant hub proteins containing IRF4 (Degree = 23), PLAUR (Degree = 17) and HIST1H2AE (Degree = 15) (Figure 5).

Discussion

We identified 868, 476 and 575 DEGs in CAD patients from Europe, North America and Asia. These genes encode proteins involved in signal transduction, blood coagulation, cell adhesion, immune response phospholipid biosynthetic process, lipid metabolic process and in-

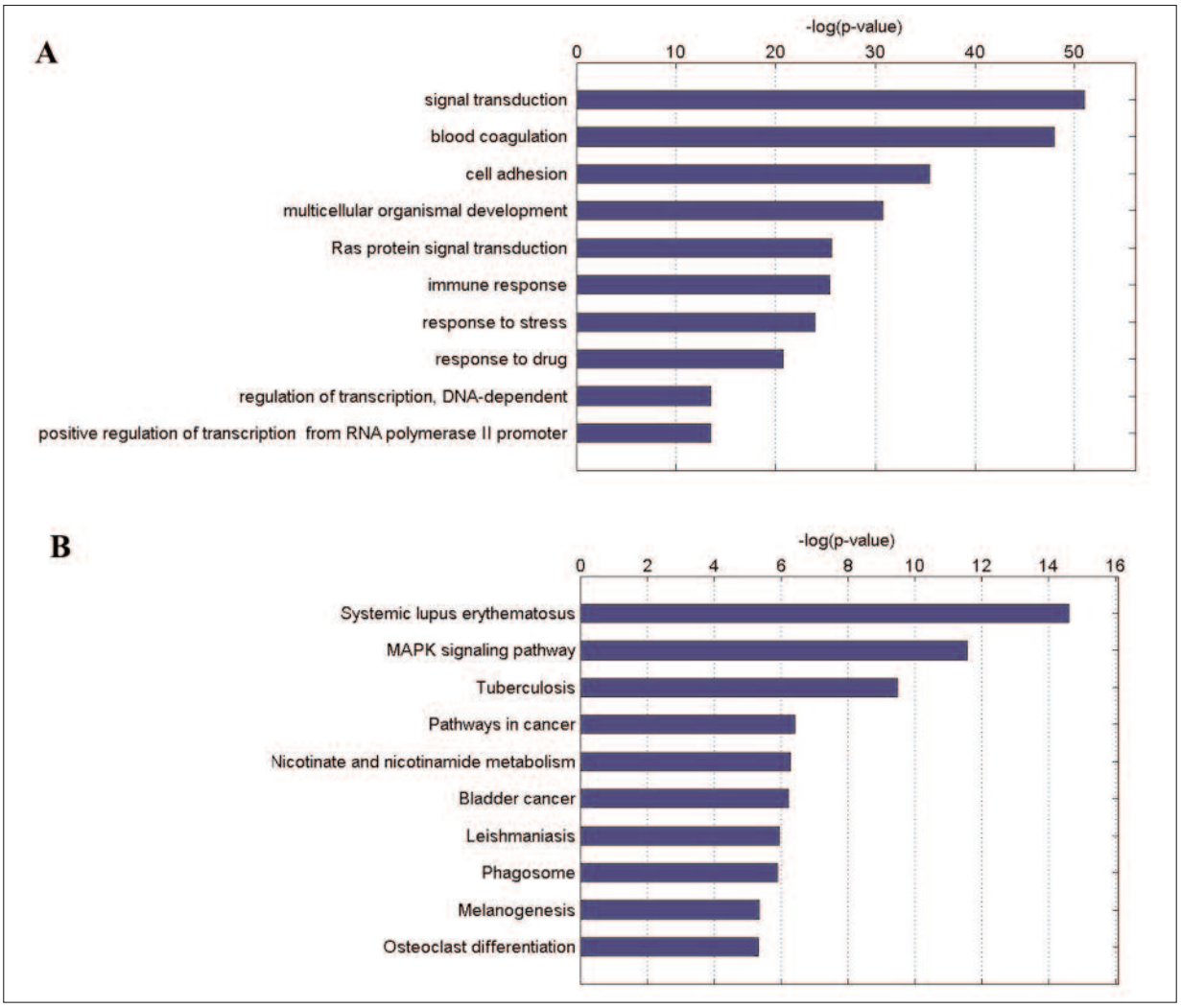

Figure 1. The significantly enriched functional annotation of DEGs detected in CAD from Asia population. The top 10 enriched GO categories for biological process. ***B,*** The top 10 enriched KEGG pathway.

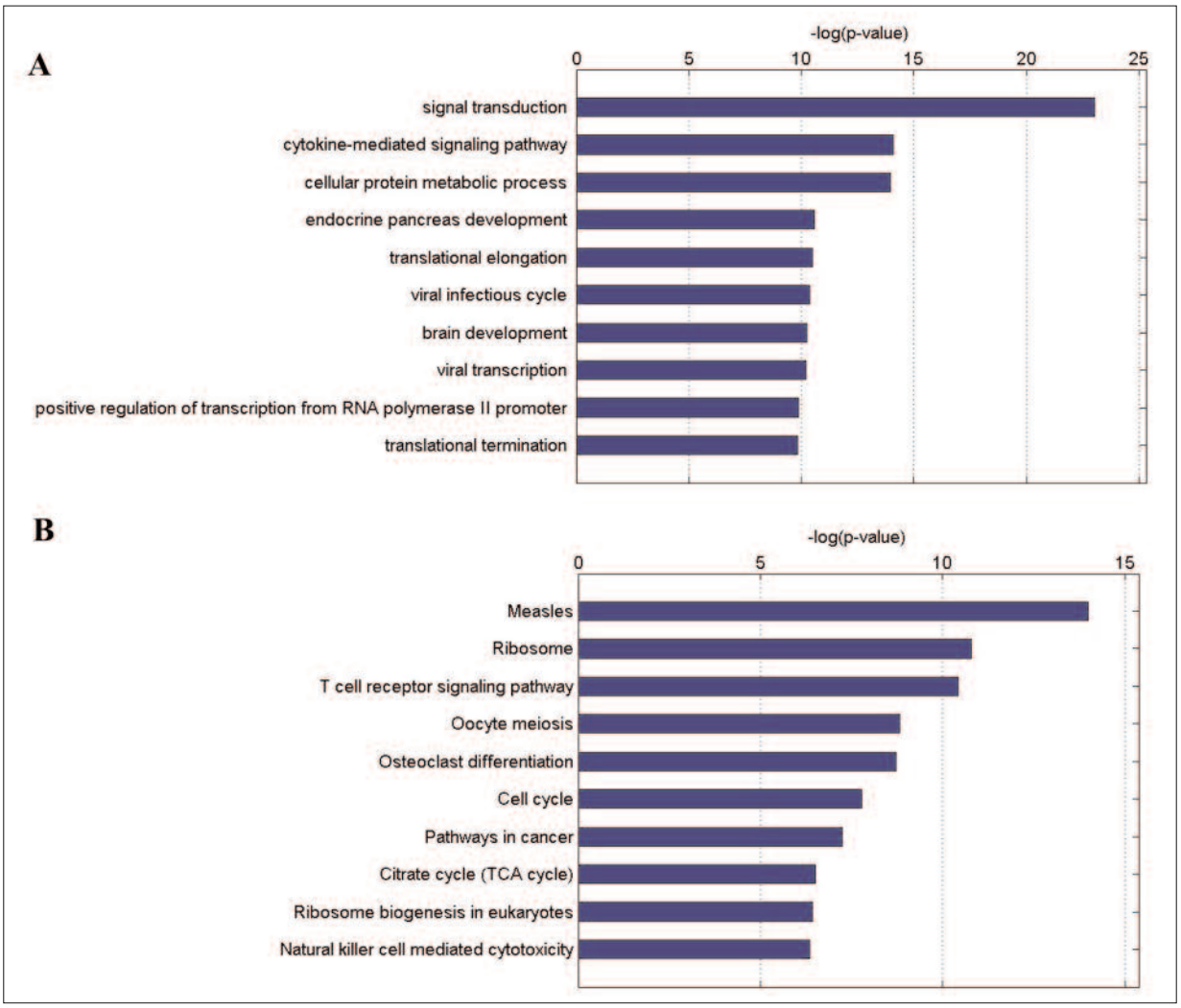

Figure 2. The significantly enriched functional annotation of DEGs detected in CAD from Europe population. ***A,*** The top 10 enriched GO categories for biological process. ***B,*** The top 10 enriched KEGG pathway.

fections, which were found to play an essential role in the development of CAD. Systemic and local inflammation including activation and migration of immune cells into the vessel wall has been shown to play an essential pathologic role in atherosclerotic coronary artery disease[16,17]. It has been confirmed that adhesion of leukocytes to activated endothelial cells and their migration into the arterial wall would initiate, propagate, and destabilize coronary plaques[18]. A recent integrative genomics study detected CAD-associated gene networks of lipid metabolism, coagulation, immunity[19]. A large number of studies have reported on associations of human CAD with certain persistent bacterial and viral infections[20,21].

Although the sets of DEGs identified in CAD patients from Europe, North America and Asia, are largely non-overlapping because of health disparities, 24 genes were found common in the three continents, and 6 genes were previously linked to CAD or atherosclerosis. PLAUR, is a crucial component in plasminogen activator (PA) system which may predispose to thrombosis and modulate the vascular response to injury. PLAUR was found to be up-regulated significantly by experiments of cholesterol-fed rabbits in early atherosclerotic lesions[22]. As with tissue-type PA, atherosclerotic arteries tended to have increased intimal PLAUR expression, and the ratio of intima-to-media protein expression increased the most for PLAUR[23].

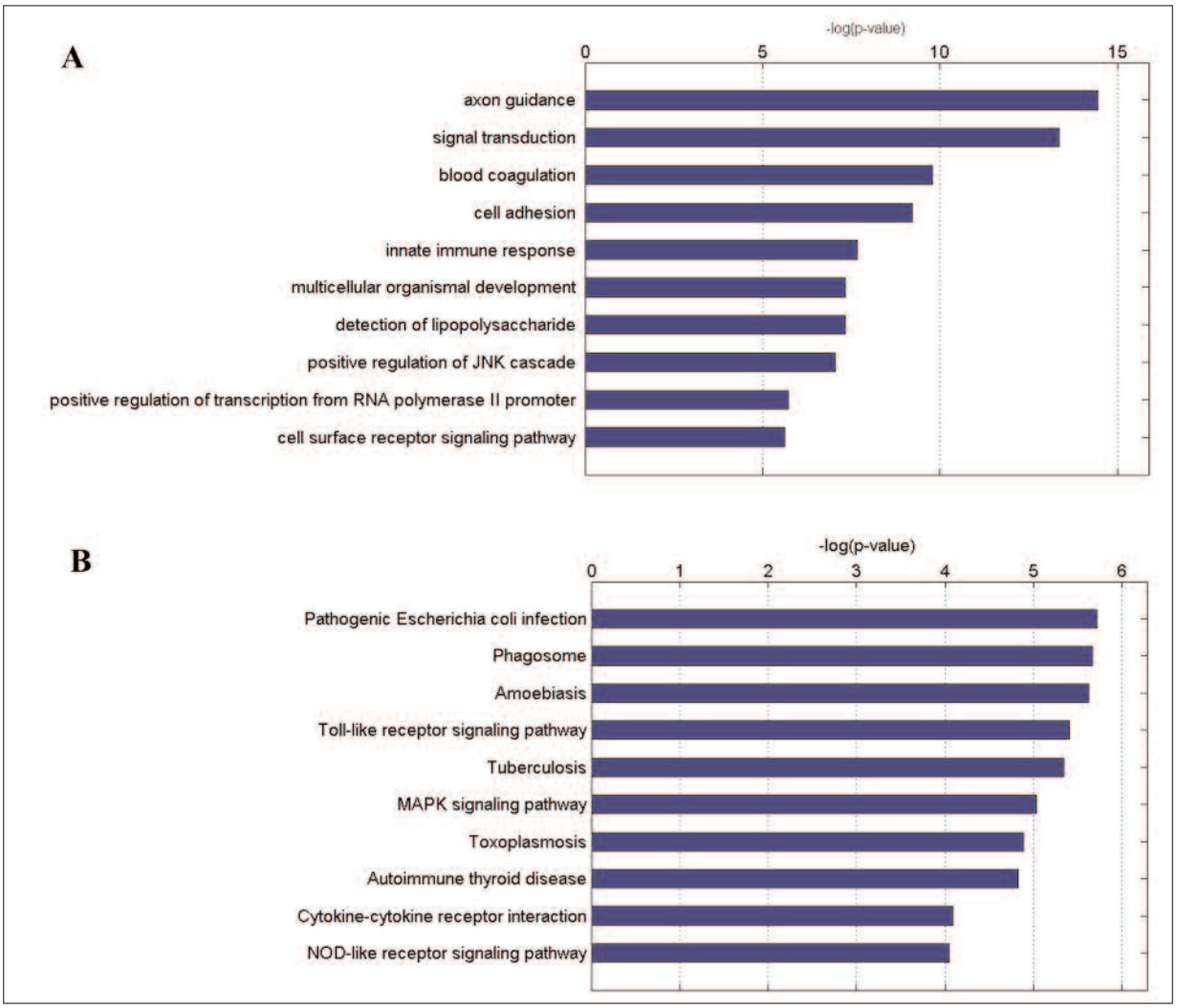

Figure 3. The significantly enriched functional annotation of DEGs detected in CAD from North America population. ***A,*** The top 10 enriched GO categories for biological process. ***B,*** The top 10 enriched KEGG pathway.

TNFRSF10C with an extracellular TRAIL-binding domain and a transmembrane domain is a membrane-bound decoy receptor of TRAIL, which may play an important role in atherosclerosis by regulating IGF1R expression in vascular smooth muscle cell (VSMC) in an NF-κB-dependent manner. TNFRSF10C and DR4 promoted VSMC proliferation at low concentrations of TRAIL without inducing human VSMC apoptosis due to lack of death domains or functional death domains[24,25].

ID3, emerged as an important upstream regulator of atheroprotective pathways in immune and vessel wall cells, was found to be directly with human coronary artery pathology by identification of functionally significant polymorphism ofID3 in patients with cardiovascular disease[26,27]. The protein encoded by FGA is the alpha component of fibrinogen, which is implicated in coagulation reactions and atherosclerosis development[28,29]. Some prospective studies revealed that fibrinogen levels in plasma have been associated with CAD and myocardial infarction risk[30,31]. FGB rs1800787 and rs1800789 SNPs seem to impose protection to CAD onset reducing the risk by about 50% in homozygotes for the minor alleles[32].

IRF4, one member of the interferon regulatory factor (IRF) family, which are initially described as downstream regulators of interferon signaling and function in the regulation of innate and adaptive immune response[33], was iden-

Table I. The 24 genes that showed up-regulation or down-regulation in the CAD examined in patients from Asia, Europe and North America.

ID	Symbol	Official full name
129642	MBOAT2	Membrane bound O-acyltransferase domain containing 2
1824	DSC2	Desmocollin 2
3012	HIST1H2AE	Histone cluster 1, H2ae
4122	MAN2A2	Mannosidase, alpha, class 2A, member 2
4608	MYBPH	Myosin binding protein H
5329	PLAUR	Plasminogen activator, urokinase receptor
7940	LST1	Leukocyte specific transcript 1
79930	DOK3	Docking protein 3
8358	HIST1H3B	Histone cluster 1, H3b
8794	TNFRSF10C	Tumor necrosis factor receptor superfamily, member 10c, decoy without an intracellular domain
8798	DYRK4	Dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 4
55082	ARGLU1	Arginine and glutamate rich 1
2165	F13B	Coagulation factor XIII, B polypeptide
3399	ID3	Inhibitor of DNA binding 3, dominant negative helix-loop-helix protein
3983	ABLIM1	Actin binding LIM protein 1
26960	NBEA	Neurobeachin
55314	TMEM144	Transmembrane protein 144
2243	FGA	Fibrinogen alpha chain
3662	IRF4	Interferon regulatory factor 4
441024	MTHFD2L	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2-like
9215	LARGE	Like-glycosyltransferase
57515	SERINC1	Serine incorporator 1
6236	RRAD	Ras-related associated with diabetes
9947	MAGEC1	Melanoma antigen family C, 1

tified as the most significant hub proteins in the PPI network. IRF4 is also closely related to inflammation[34]. IRF4 was highly expressed in immune cells, human and mouse cardiomyocytes as a transcription factor. Interestingly IRF4 was downregulated in human dilated cardiomyopathy hearts and mouse hypertrophic hearts, implying the important role of IRF4 in the process of cardiac hypertrophy and failure. Out of genes previously reported to be associated with CAD in the common set of 24 genes, some genes such as F13B, SERINC1, MBOAT2, HIST1H2AE and HIST1H3B may be implicated in CAD by involving blood coagulation, phospholipid biosynthetic process and other related functions.

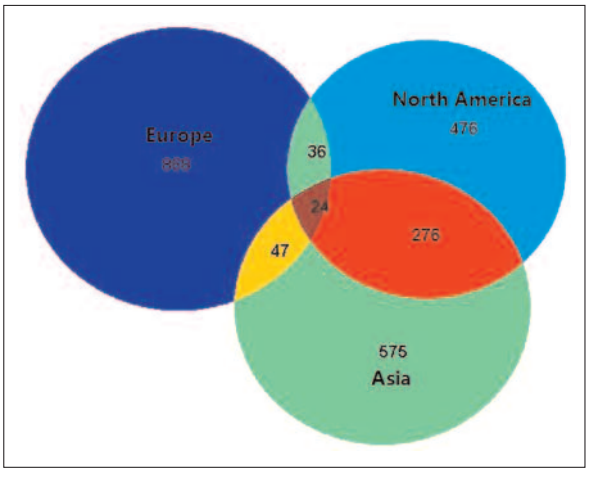

Figure 4. Comparison of DEGs in CAD samples from Europe, Asia and North America.

Conclusions

The data presented here have identified genes differentially expressed between CAD and NC samples in Asia, Europe and North America population. Although the sets of DEGs identified in CAD from different continents are largely non-overlapping because of health disparities, a common set of 24 genes was identified. Consistency among the results obtained on Asia, Europe and North America samples indicates a common biology of CAD, which seems to be mediated by similar pathways regardless of existing differences related to ethnicity, lifestyle, or environmental factors.

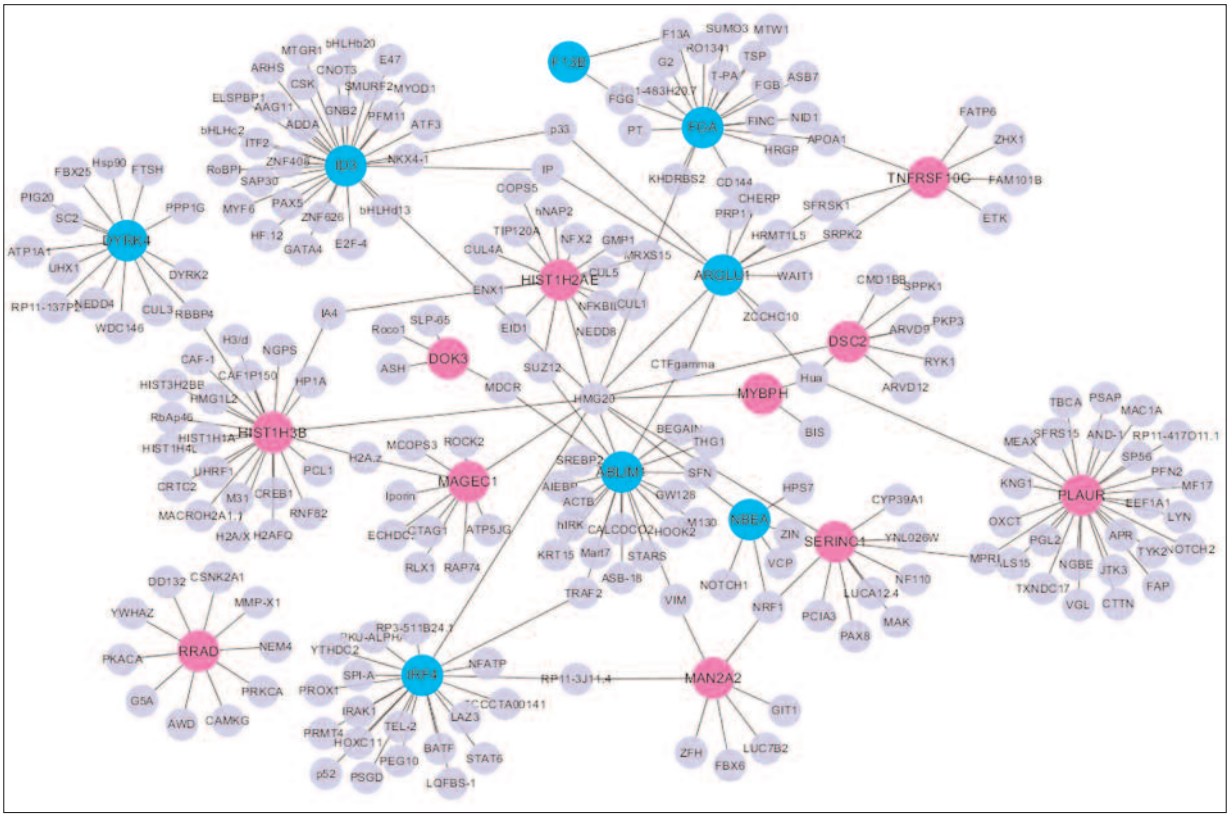

Figure 5. PPI networks of the common set of genes across different continents. Nodes represent proteins, edges represent interactions between two proteins. Red- and blue-color nodes represent products of up- and down-regulated DEGs, respectively. Purple nodes denote products of genes predicted to interact with the DEGs.

Acknowledgements

The authors would like to thank Professor Wei Cui for his support throughout this work.

Conflict of Interest

The Authors declare that there are no conflicts of interest.

References

1) Cao Y, Lu L, Liu M, Li XC, Sun RR, Zheng Y, Zhang PY. Impact of epigenetics in the management of cardiovascular disease: a review. Eur Rev Med Pharmacol Sci 2014; 18: 3097-3104.
2) Balta S, Demirkol S, Kurt O, Unlu M, Iyisoy A, Sarlak H. Inflammatory status in patients with coronary artery disease. Eur Rev Med Pharmacol Sci 2014; 18: 939-940.
3) Liu CY, Chen CQ. Intra- and extracranial atherosclerotic stenosis in China: epidemiology, diagnosis, treatment and risk factors. Eur Rev Med Pharmacol Sci 2014; 18: 3368-3379.
4) Barnett AH, Dixon AN, Bellary S, Hanif MW, O'Hare JP, Raymond NT, Kumar S. Type 2 diabetes and cardiovascular risk in the UK south Asian community. Diabetologia 2006; 49: 2234-2246.
5) Uppaluri CR. Heart disease and its related risk factors in Asian Indians. Ethn Dis 2002; 12: 45-53.
6) Stefler D, Bhopal R, Fischbacher CM. Might infection explain the higher risk of coronary heart disease in South Asians? Systematic review comparing prevalence rates with white populations in developed countries. Public Health 2012; 126: 397-409.
7) Huang CC, Lloyd-Jones DM, Guo X, Rajamannan NM, Lin S, Du P, Huang Q, Hou L, Liu K. Gene expression variation between African Americans and whites is associated with coronary artery calcification: the multiethnic study of atherosclerosis. Physiol Genomics 2011; 43: 836-843.
8) Sinnaeve PR, Donahue MP, Grass P, Seo D, Vonderscher J, Chibout SD, Kraus WE, Sketch M, Jr, Nelson C, Ginsburg GS, Goldschmidt-Clermont PJ, Granger CB. Gene expression patterns in peripheral blood correlate with the extent of coronary artery disease. PLoS One 2009; 4: e7037.

9) Elashoff MR, Wingrove JA, Beineke P, Daniels SE, Tingley WG, Rosenberg S, Voros S, Kraus WE, Ginsburg GS, Schwartz RS, Ellis SG, Tahirkheli N, Waksman R, McPherson J, Lansky AJ, Topol EJ. Development of a blood-based gene expression algorithm for assessment of obstructive coronary artery disease in non-diabetic patients. BMC Med Genomics 2011; 4: 26.

10) Beineke P, Fitch K, Tao H, Elashoff MR, Rosenberg S, Kraus WE, Wingrove JA, PREDICT Investigators. A whole blood gene expression-based signature for smoking status. BMC Med Genomics 2012; 5: 58.

11) Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res 2013; 41: D991-D995.

12) Altermann E, Klaenhammer TR. PathwayVoyager: pathway mapping using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. BMC Genomics 2005; 6: 60

13) Tabas-Madrid D, Nogales-Cadenas R, Pascual-Montano A. GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. Nucleic Acids Res 2012; 40: W478-483.

14) Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B, Li Y, Hao Y, Ooi C, Godwin B, Vitols E. A protein interaction map of Drosophila melanogaster. science 2003; 302: 1727-1736.

15) Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.

16) Hansson GK, Libby P, Schonbeck U, Yan ZQ. Innate and adaptive immunity in the pathogenesis of atherosclerosis. Circ Res 2002; 91: 281-291.

17) Libby P, Ridker PM, Maseri A. Inflammation and atherosclerosis. Circulation 2002; 105: 1135-1143.

18) Blankenberg S, Barbaux S, Tiret L. Adhesion molecules and atherosclerosis. Atherosclerosis 2003; 170: 191-203.

19) Makinen VP, Civelek M, Meng Q, Zhang B, Zhu J, Levian C, Huan T, Segre AV, Ghosh S, Vivar J, Nikpay M, Stewart AF, Nelson CP, Willenborg C, Erdmann J, Blakenberg S, O'Donnell CJ, Marz W, Laaksonen R, Epstein SE, Kathiresan S, Shah SH, Hazen SL, Reilly MP, Lusis AJ, Samani NJ, Schunkert H, Quertermous T, McPherson R, Yang X, Assimes TL. Integrative genomics reveals novel molecular pathways and gene networks for coronary artery disease. PLoS Genetics 2014; 10: e1004502.

20) Danesh J, Collins R, Peto R. Chronic infections and coronary heart disease: is there a link? Lancet 1997; 350: 430-436.

21) Gupta S. Chronic infection in the aetiology of atherosclerosis--focus on Chlamydia pneumoniae. Atherosclerosis 1999; 143: 1-6.

22) Noda-Heiny H, Daugherty A, Sobel BE. Augmented urokinase receptor expression in atheroma. Arterioscler Thromb Vasc Biol 1995; 15: 37-43.

23) Raghunath PN, Tomaszewski JE, Brady ST, Caron RJ, Okada SS, Barnathan ES. Plasminogen activator system in human coronary atherosclerosis. Arterioscler Thromb Vasc Biol 1995; 15: 1432-1443.

24) Kavurma MM, Schoppet M, Bobryshev YV, Khachigian LM, Bennett MR. TRAIL stimulates proliferation of vascular smooth muscle cells via activation of NF-kappaB and induction of insulin-like growth factor-1 receptor. J Biol Chem 2008; 283: 7754-7762.

25) Zhang J, Kabra NH, Cado D, Kang C, Winoto A. FADD-deficient T cells exhibit a disaccord in regulation of the cell cycle machinery. J Biol Chem 2001; 276: 29815-29818.

26) Doran AC, Lehtinen AB, Meller N, Lipinski MJ, Slayton RP, Oldham SN, Skaflen MD, Yeboah J, Rich SS, Bowden DW, McNamara CA. Id3 is a novel atheroprotective factor containing a functionally significant single-nucleotide polymorphism associated with intima-media thickness in humans. Circ Res 2010; 106: 1303-1311.

27) Manichaikul A, Rich SS, Perry H, Yeboah J, Law M, Davis M, Parker M, Ragosta M, Connelly JJ, McNamara CA, Taylor AM. A functionally significant polymorphism in ID3 is associated with human coronary pathology. PLoS One 2014; 9: e90222.

28) Koenig W. Fibrin(ogen) in cardiovascular disease: an update. Thromb Haemost 2003; 89: 601-609.

29) Harley SL, Sturge J, Powell JT. Regulation by fibrinogen and its products of intercellular adhesion molecule-1 expression in human saphenous vein endothelial cells. Arterioscler Thromb Vasc Biol 2000; 20: 652-658.

30) Danesh J, Collins R, Appleby P, Peto R. Association of fibrinogen, C-reactive protein, albumin, or leukocyte count with coronary heart disease: meta-analyses of prospective studies. JAMA 1998; 279: 1477-1482.

31) Fibrinogen Studies C, Danesh J, Lewington S, Thompson SG, Lowe GD, Collins R, Kostis JB, Wilson AC, Folsom AR, Wu K, Benderly M, Goldbourt U, Willeit J, Kiechl S, Yarnell JW, Sweetnam PM, Elwood PC, Cushman M, Psaty BM, Tracy RP, Tybjaerg-Hansen A, Haverkate F, Maat MP, Fowkes FG, Lee AJ, Smith FB, Salomaa V, Harald K, Rasi R, Vahtera E, Jousilahti P, Pekkanen J, D'Agostino R, Kannel WB, Wilson PW, Tofler G, Arocha-Pinango CL, Rodriguez-Larralde A, Nagy E, Mijares M, Espinosa R, Rodriquez-Roa E, Ryder E, Diez-Ewald MP, Campos G, Fernandez V, Torres E, Marchioli R, Valagussa F, Rosengren A, Wilhelmsen L, Lappas G, Eriksson H, Cremer P, Nagel D, Curb JD, Rodriguez B, Yano K, Salonen JT, Nyyssonen K, Tuomainen TP, Hedblad B, Lind P, Loewel H, Koenig W, Meade TW, Cooper JA, De Stavola B, Knottenbelt C, Miller GJ, Cooper JA, Bauer KA, Rosenberg RD, Sato S, Kitamura A, Naito Y, Palosuo T, Ducimetiere P, Amouyel P, Arveiler D, Evans AE, Ferrieres J, Juhan-Vague I,

Bingham A, Schulte H, Assmann G, Cantin B, Lamarche B, Despres JP, Dagenais GR, Tunstall-Pedoe H, Woodward M, Ben-Shlomo Y, Davey Smith G, Palmieri V, Yeh JL, Rudnicka A, Ridker P, Rodeghiero F, Tosetto A, Shepherd J, Ford I, Robertson M, Brunner E, Shipley M, Feskens EJ, Kromhout D, Dickinson A, Ireland B, Juzwishin K, Kaptoge S, Lewington S, Memon A, Sarwar N, Walker M, Wheeler J, White I, Wood A. Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. JAMA 2005; 294: 1799-1809.

32) Theodoraki EV, Nikopensius T, Suhorutsenko J, Peppes V, Fili P, Kolovou G, Papamikos V, Richter D, Zakopoulos N, Krjutskov K, Metspalu A, Dedoussis GV. Fibrinogen beta variants confer protection against coronary artery disease in a Greek case-control study. BMC Med Genetics 2010; 11: 28.

33) Eguchi J, Wang X, Yu S, Kershaw EE, Chiu PC, Dushay J, Estall JL, Klein U, Maratos-Flier E, Rosen ED. Transcriptional control of adipose lipid handling by IRF4. Cell Metab 2011; 13: 249-259.

34) Mudter J, Amoussina L, Schenk M, Yu J, Brustle A, Weigmann B, Atreya R, Wirtz S, Becker C, Hoffman A, Atreya I, Biesterfeld S, Galle PR, Lehr HA, Rose-John S, Mueller C, Lohoff M, Neurath MF. The transcription factor IFN regulatory factor-4 controls experimental colitis in mice via T cell-derived IL-6. J Clin Invest 2008; 118: 2415-2426.