



Bioinformatics Analysis to Find Novel Biomarkers for Coronary Heart Disease

Akram Gholipour¹, Farshad Shakerian^{2,3}, Ali Zahedmehr³, Shiva Irani¹,
*Mahshid Malakootian², *Seyed Javad Mowla⁴

1. Department of Biology, Science and Research Branch, Islamic Azad University, Tehran, Iran
2. Cardiogenetic Research Center, Rajaie Cardiovascular Medical and Research Center, Iran University of Medical Sciences, Tehran, Iran
3. Cardiovascular Intervention Research Center, Rajaie Cardiovascular Medical and Research Center, Iran University of Medical Sciences, Tehran, Iran
4. Department of Molecular Genetics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran

*Corresponding Authors: Email: sjmowla@yahoo.com; Malakootian@rhc.ac.ir

(Received 15 Jan 2021; accepted 19 Apr 2021)

Abstract

Background: Coronary heart disease (CHD), a major cause of death worldwide, is defined as a narrowing or blockage of the coronary arteries that supply oxygen and blood to the heart. We aimed to find potential biomarkers for coronary artery disease, by comparing the expression profile of blood exosomes of both normal and CHD samples.

Methods: Datasets of 6 CHD and 6 normal samples of blood exosomes were downloaded, and differentially expressed RNAs, with adjusted $P < 0.01$ and $\log_2\text{FoldChange} \geq 1$ were achieved. Moreover, gene ontology (GO) and pathway analysis were accomplished by PANTHER database for datasets.

Results: Our data analysis found 119 differentially expressed genes between two datasets. By comparing transcriptome profiles, we candidate the highest downregulated gene, *ACSBG1*, and the highest upregulated one, *DEF44*, as specific biomarkers for CHD. Furthermore, GO and pathway analysis depicted that aforementioned differentially expressed genes are mostly involved in different molecular metabolic process, inflammation, immune system process and response to stimulus pathways which all cause cardiovascular diseases.

Conclusion: We have provided new potential biomarkers for CHD, though experimental validation is still needed to confirm the suitability of the candidate genes for early detection of CHD.

Keywords: Blood exosomes; Coronary heart disease; Differentially expressed genes

Introduction

Coronary heart disease (CHD), also known as coronary artery disease (CAD), is the most common type of heart disease causing more than 365,000 deaths in 2017 (1). CAD is characterized

by the accumulation of plaque, combination of fat, cholesterol, and calcium deposits in the coronary arteries, and may cause heart attack and/or other clinical complications (2). The traditional



risk factors for CAD are high LDL cholesterol, low HDL cholesterol, high blood pressure, positive family history, diabetes, and smoking (3, 4). Although there has been some progress in diagnosis of CAD over recent years, finding specific and sensitive biomarkers is still essential for this common heart disease. Studies on non-invasive biomarkers have mainly focused on early detection of diseases, including CHD (5-7). Therefore, increasing demand for rapid screening as well as intervention in subsequent analyses suggests the importance of searching for such biomarkers (8). Serum or plasma (blood) samples are easily accessible and provide critical information to understand the biological characteristics of diseases (9). Exosomes is another important element in blood, which can act as a drug delivery vehicle and can be exploited for therapy in different clinical perspectives (10).

In recent years, transcriptome profiling has been broadly utilized to understand cellular states by exploring the expression patterns of genes in different samples and suggests appreciated information on the biological processes that leads to the given disease (11).

In the current study, bioinformatics analyses were applied to find and introduce specific biomarkers for CAD, by examining the differential expression profile of exosomes obtained from CAD and normal samples, and to recommend the important pathways involved in CHD.

Materials and Methods

Datasets and samples

All datasets were downloaded from the NCBI Gene Expression Omnibus (GEO). CHD and normal samples were obtained from GSE99985 and GSE100206, respectively. These data are a complete transcriptome profiling that reveals abundant RNAs in blood exosomes. 1-4 ml of plasma and serums had been used to extract exosomal RNAs (12).

Bioinformatics analysis to identify differentially expressed genes

Adapter sequences were removed by Trimmomatic version 0.36, and sequence reads were aligned to the human genome (hg38) with HISAT2 2.1.0 to analyze the downloaded datasets. Then, the GENCODEV30 (hg38) database was applied for analyzing the expression pattern of the genes. The read counts of each transcript were considered by HTSeq-0.9.1. In this study, 6 CHD versus 6 normal samples were analyzed to identify differentially expressed (DE) genes. Statistically, significant differentially expressed genes were identified using edgeR software in the R program with $\log_2\text{FoldChange} \geq 1$ and adjusted $P < 0.01$. Volcano plot was used to illustrate differentially expressed genes.

Functional enrichment analysis

We used PANTHER database to perform GO (Gene ontology) and pathway analysis to acquire important biological pathways among CHD patients and Normal samples.

Results

Differentially expressed genes in normal versus CHD samples

Our analysis by edgeR package revealed that 119 genes had differential expression between normal samples versus CHD ones. Among them, 93 genes were upregulated and 26 genes were downregulated (Fig.1, supplementary 1).

The Acyl-CoA Synthetase, Bubblegum Family member 1 (*ACSBG1*), ENSG00000103740.9, and Defensin alpha 4 (*DEFA4*), ENSG00000164821.4, with the $\log_2\text{FC} = -10.816645$ and $\log_2\text{FC} = 8.959236$, respectively, were the most downregulated and upregulated genes in patients' samples. Therefore, we introduced them as probable biomarkers with the most differential expression in patients' samples versus normal ones.

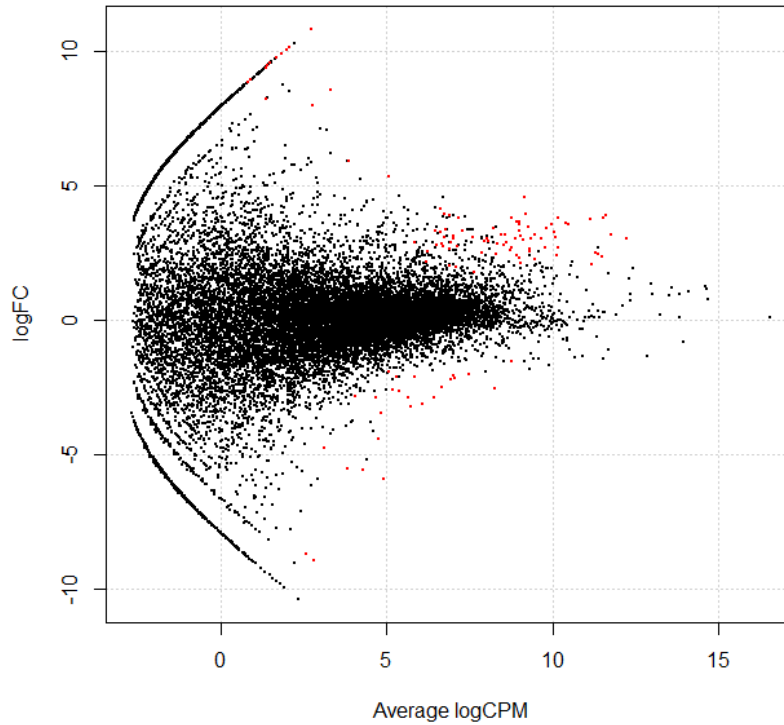


Fig. 1: Volcano plot shows differentially expressed genes. Red dots indicate differentially expressed genes in which the up-regulated genes with positive logFC were located in upper part of the graph(+0) and the down-regulated genes with negative logFC were mapped to the lower part of the graph(-0). Nonsignificant genes were illustrated in black dots

Gene ontology enrichment analysis

Gene ontology analyses were applied to explore the function of DE genes. In this part, biological processes (BP), cellular component (CC) and molecular function (MF) of genes with differential expression patterns were examined. In general, examining GO analysis, biological regulation, processes related to cell proliferation and biogenesis, metabolic and catalytic processes, as well as processes involved in immune responses had a great effect on the creation of this disease. The top significant terms enriched in GO analysis of the DE genes are shown in (Fig. 2A-C).

In addition, the presence of the *ACSBG1* and *DEFA4*, as genes with the highest differential expression, was examined in BP more precisely. *ACSBG1* gene has important role in cellular metabolic and immune responses which are major parts of cellular process and immune system process in significant terms enriched of BP, respectively. Moreover, *DEFA4* gene has vital role in immune response and response to stress that is part of the immune system process and response to stimulus as significant terms enriched in BP, as well (Fig. 3A-C).

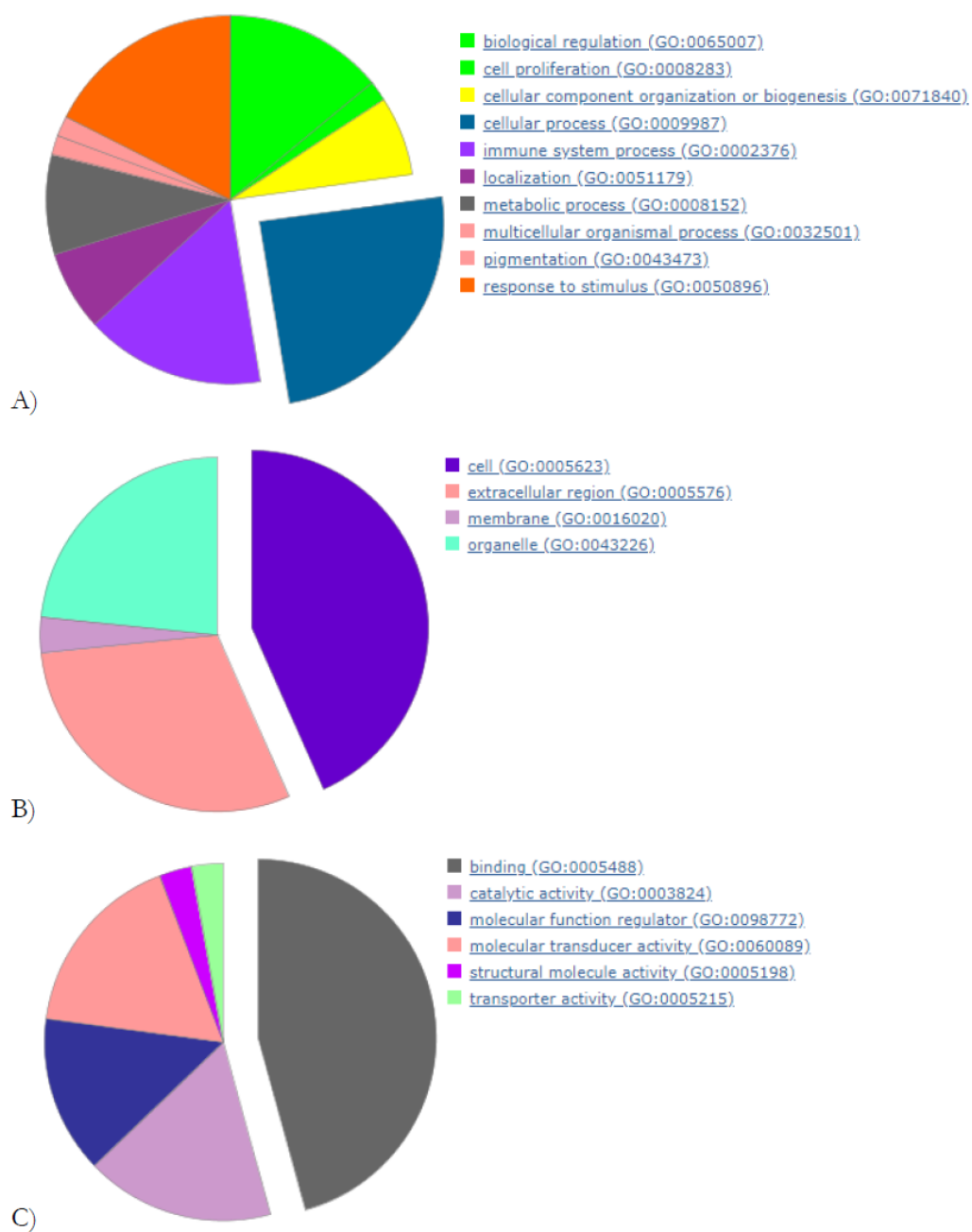


Fig. 2: Gene ontology (GO) enrichment analysis of differentially expressed genes. A) Biological process (BP) shows that the majority of the differentially expressed genes involve in biological regulation. B) Cellular component (CC) demonstrates that the genes with differential expression have roles in different cellular processes. C) Molecular function (MF) shows that those genes implicates in binding and catalytic activities

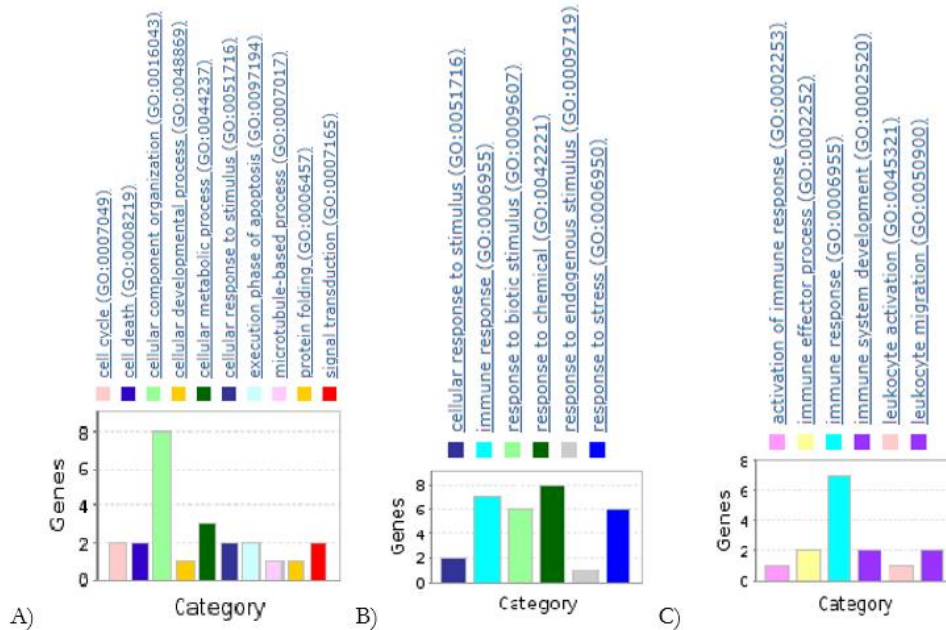


Fig. 3: Three major parts of biological process are cellular process, response to stimulus, and immune system processes which each has more detailed categories. A) Cellular process demonstrates that *ACSBG1* gene has an important role in cellular metabolic. B) Response to stimulus reveals that *DEFA4* gene has a significant role in response to stress. C) Immune system process shows that *ACSBG1* and *DEFA4* genes have important roles in immune response

Pathway enrichment analysis

Pathway enrichment analysis determined that inflammation mediated by chemokine and cytokine signaling pathways is one of the important pathways in the development of cardiovascular diseases. Examination of the genes in this pathway also revealed that platelet factor 4 (*PF4*) and regulation of G-protein signaling 18 (*RGSI8*), which is downregulated in CHD patients, have differential expression in this part of pathway enrichment analysis.

Moreover, blood coagulation and TGF-beta signaling pathways were the other influenced pathway in CHD disease. The top ten significant terms pathways enriched DE genes are shown in Fig. 4.

GO and pathway analysis results revealed that cellular and immune system processes, metabolic process, response to stimulus, blood coagulation and inflammation are the most important pathways demonstrated in blood exosomes of CHD disease.

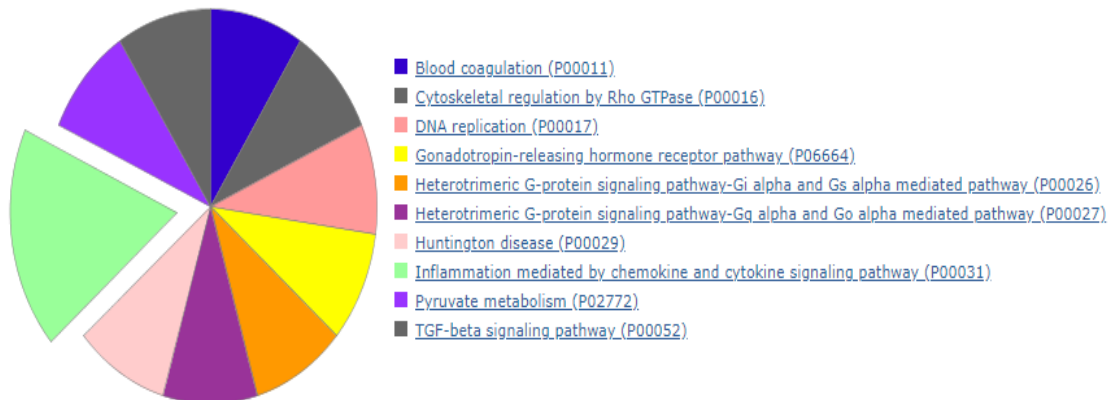


Fig. 4: The top significant terms pathways enriched differential expressed genes. The largest percentage of the DE genes involve in the inflammation mediated by chemokine and cytokine signaling pathway which is shown in green

Discussion

CHD is a prevalent disease estimated that a newly diagnosed American will suffer from the disease every 42 seconds. CHD causes about one-third of all deaths in people older than 35 yr worldwide (2).

There are numerous cardiovascular disease biomarkers such as the cardiac natriuretic peptides and the cardiac troponins which are currently powerful markers utilized in clinic (13, 14). Transcriptom-based biomarkers are recent advances in cardiovascular biomarker field. Weinberg and colleagues identified *ST2* gene by microarray analysis which is upregulated in cardiac myocytes and regulates inflammation and immunity (15, 16). Gene expression microarray analysis showed that serum and plasma level of GDF-15, a distant member of the TGF- β cytokine superfamily, is a strong predictor for acute coronary syndromes and heart failure (17). Therefore, gene expression analyses influenced the area of biomarker identification in the cardiovascular field (18). Recently, the study of cardiac markers has been extensively developed. However, there are some limitations in cardiac markers and current assays have failed to provide a good specificity for the disease, and also some of the markers are not suitable for early detection (19). Therefore, the importance of the present study is that it examines the genes in CHD that could be identified as possible biomarkers that allow us to move from efficiency to clinical effectiveness.

In addition, the introduction of accurate biomarkers capable of decreasing invasiveness and the cost of diagnostic testing of CHD will be valuable (20). We applied gene expression analysis based on RNA-sequencing analysis to find some potential transcriptomics-based biomarkers for CAD. The bioinformatic analysis introduced *ACSBG1* and *DEFA4* genes as probable diagnostic biomarkers for CHD patients. In line with this finding, *Acyl-CoA Synthetase* and *Bubblegum Family member 1* genes downregulated in patients

suffering from CHD play important roles in metabolic pathways and immune responses.

ACSBG1 is an enzyme encoded by the *ACSBG1* gene in humans. This gene influences the fatty acid beta-oxidation (peroxisome) and therefore regulates lipid metabolism. Fatty acids are incorporated into membranes and signaling molecules and have some roles in energy storage and metabolism (21, 22). *ACSBG1* gene was downregulated in skin of psoriatic patients. Psoriasis is characterized by chronic inflammatory, immunological, and vascular abnormalities. The genes involved in lipid metabolism were important mediators in this disease (23). Since some of the risk factors of CAD are related to lipid metabolism that involves the biosynthesis and degradation of cholesterol, triglycerides, and lipoproteins, the transcriptome analysis is an efficient way to survey important interactions between genes and metabolic pathways (24-26).

The most upregulated candidate exosomal biomarker was *Defensin alpha 4* which has roles in response to stress and immune responses. *DEFA4* has corticostatic activity too and may modulate lymphocyte activity. *DEFA4* also known as human neutrophil peptides is one of the abundant proteins in the granules of neutrophils and is also expressed in subset populations of lymphocytes and monocytes (27, 28). Gene expression of *DEFA1/3* and *DEFA4* were analyzed in healthy salivary gland tissue in comparison with salivary gland tumors. The gene expression of *DEFA-1/3* and *-4* was significantly increased in tumors that underline importance of immunological reactions during tumorigenesis (29).

CHD can cause cardiac injury and immediately activate neutrophil recruitment in the myocardium (30). On the other hand, coronary heart disease is associated with inflammation mechanisms that have played essential roles in formation of plaque, thrombosis and oxidative stress (31-33). This gene may have prominent role in the CHD and could be a good biomarker for this disease. Although it needs more experimental analysis to confirm this.

A couple of studies demonstrated that exosomes (Circulating extracellular vesicles) are generated by cells and released into body fluids and carried cell-specific materials like proteins, lipids, and genetic materials taken up selectively by target cells. Exosomes can regulate cell function, and as a result, examining them can be a non-invasive way to detect biomarkers in a variety of diseases, and they may reflect disease processes (34, 35). Plasma extracellular vesicles are under investigation as vehicles for proteins that play a role in cardiovascular disease. Some studies reported that the levels of exosomes and endothelial dysfunction are associated with cardiometabolic risk and heart failure (36, 37). The levels of Serpin F2, Serpin G1, Cystatin C, and CD14, as extracellular vesicle proteins, were associated with vascular events and presence of heart failure in patients (38).

Conclusion

The differentially expressed genes, *ACSBG1* and *DEFA4*, are introduced here as potential biomarkers of CHD disease. These two genes play important roles in several pathways involved in the pathogenesis of coronary heart disease such as metabolic pathways, response to stress and immune system process. The existence of these genes in exosomes of CHD patients suggests a probable role for them as suitable biomarkers of the disease. However, to further confirm the bioinformatics analysis data, we need to verify the results experimentally in big population size of normal and CHD patients.

Journalism Ethics considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Acknowledgements

The authors thank Dr. Ali Sharifi-Zarchi for his kind help during analysis of the data. This work was supported by a research grant from National Institute for Medical Research Development (NIMAD, 943104) and the Research Deputyship of Rajaie Cardiovascular, Medical, and Research Center (94111).

Conflict of interest

The authors declare no conflict of interest.

References

1. J Benjamin E, Muntner P, Alonso A, et al (2019). Heart Disease and Stroke Statistics-2019 Update: A Report from the American Heart Association. *Circulation*, 139 (10): e56-e 528.
2. Sanchis-Gomar F, Perez-Quilis C, Leischik R, et al (2016). Epidemiology of coronary heart disease and acute coronary syndrome. *Ann Transl Med*, 4 (13): 256.
3. Charlson FJ, Moran AE, Freedman G, et al (2013). The contribution of major depression to the global burden of ischemic heart disease: a comparative risk assessment. *BMC Med*, 11: 250.
4. Ambrose John A, Singh M (2015). Pathophysiology of coronary artery disease leading to acute coronary syndromes. *F1000Prime Rep*, 7: 08.
5. Parsanathan R, K.Jain S (2020). Novel Invasive and Noninvasive Cardiac-Specific Biomarkers in Obesity and Cardiovascular Diseases. *Metab syndr Relat Disord*, 18 (1): 10-30.
6. Mirzadeh Azad F, Arabian M, Maleki M, et al (2020). Small Molecules with Big Impacts on Cardiovascular Diseases. *Biochem Genet*, 58 (3): 359–383.
7. Malakootian M, Mirzadeh Azad F, Naeli P, et al (2017). Novel Spliced Variants of OCT4, OCT4C and OCT4C1, With Distinct Expression Patterns and Functions in Pluripotent and Tumor Cell Lines. *Eur J Cell Biol*, 96 (4): 347-355.
8. Md Arshad MK, Mohamad Fathil MF, C B Gopinath S, et al (2016). Cardiac Biomarkers: In-

- vasive to Non-invasive Assessments. *Curr Med Chem*, 23 (37): 4270-4284.
9. Mayeux R (2004). Biomarkers: Potential Uses and Limitations. *NeuroRx*, 1 (2): 182-188.
 10. Samanta S, Rajasingh Sh, Drosos N, et al (2018). Exosomes: New Molecular Targets of Diseases. *Acta Pharmacol Sin*, 39 (4): 501-513.
 11. Rani B, VK Sharma (2017). Transcriptome profiling: methods and applications- A review. *Agric Res*, 38 (4): 271-281.
 12. Shengli L, Yuchen L, Bing Ch, et al (2018). exoRBase: a database of circRNA, lncRNA and mRNA in human blood exosomes. *Nucleic Acids Res*, 46: 106-112.
 13. Dhingra R, Ramachandran SV (2017). Biomarkers in Cardiovascular Disease: Statistical Assessment and Section on Key Novel Heart Failure Biomarkers. *Trends Cardiovasc Med*, 27 (2): 123-133.
 14. Mark Richards A (2018). Future biomarkers in cardiology: my favourites. *Eur Heart J Suppl*, 20: 37-44.
 15. Weinberg E, Shimpo M, W De Keulenaer G, et al (2002). Expression and Regulation of ST2, an interleukin-1 Receptor Family Member, in Cardiomyocytes and Myocardial Infarction. *Circulation*, 106 (23): 2961-2966.
 16. Weinberg, E.O (2009). ST2 Protein in Heart Disease: From Discovery to Mechanisms and Prognostic Value. *Biomark Med*, 3 (5): 495-511.
 17. Wollert, K.C (2007). Growth-differentiation factor-15 in Cardiovascular Disease: From Bench to Bedside, and Back. *Basic Res Cardiol*, 102 (5):412-5.
 18. Antoon Siemelink M, Zeller T (2014). Biomarkers of Coronary Artery Disease: The Promise of the Transcriptome. *Curr Cardiol Rep*, 16 (8): 513.
 19. Plebani, M (2001). Biochemical Markers of Cardiac Damage: From Efficiency to Effectiveness. *Clin Chim Acta*, 311 (1): 3-7.
 20. Albus Ch, Barkhausen J, Fleck E, et al (2017). The Diagnosis of Chronic Coronary Heart Disease. *Dtsch Arztebl Int*, 114 (42): 712-719.
 21. Steinberg S J, Morgenthaler J, Heinzer A K, et al (2000). Very Long-Chain acyl-CoA Synthetases. Human "Bubblegum" Represents a New Family of Proteins Capable of Activating Very Long-Chain Fatty Acids. *J Biol Chem*, 275 (45): 35162-9.
 22. Watkins P A, Maiguel D, Jia Zh, et al (2007). Evidence for 26 Distinct Acyl-Coenzyme A Synthetase Genes in the Human Genome. *J Lipid Res*, 48 (12): 2736-50.
 23. Gudjonsson J E, Ding J, Li X, et al (2009). Global Gene Expression Analysis Reveals Evidence for Decreased Lipid Biosynthesis and Increased Innate Immunity in Uninvolved Psoriatic Skin. *J Invest Dermatol*, 129 (12): 2795-804.
 24. Kannel W B, Vasan, R.S (2009). Triglycerides as Vascular Risk Factors: New Epidemiologic Insights. *Curr Opin Cardiol*, 24 (4): 345-50.
 25. Smolders B, Lemmens R, Thijs V (2007). Lipoprotein (A) and Stroke: A Meta-Analysis of Observational Studies. *Stroke*, 38 (6): 1959-66.
 26. Lu H, Chen Y, Li L (2018). Metabolic Pathway Genes Associated with Susceptibility Genes to Coronary Artery Disease. *Int J Genomics*, 2018:9025841.
 27. Faurischou M, Borregaard N (2003). Neutrophil Granules and Secretory Vesicles in Inflammation. *Microbes Infect*, 5 (14): 1317-27.
 28. Agerberth B, Charo J, Werr J, et al (2000). The Human Antimicrobial and Chemotactic Peptides LL-37 and Alpha-Defensins Are Expressed by Specific Lymphocyte and Monocyte Populations. *Blood*, 96 (9): 3086-93.
 29. Winter J, Pantelis A, Kraus D, et al (2011). Human α -Defensin (DEFA) Gene Expression Helps to Characterise Benign and Malignant Salivary Gland Tumours. *BMC Cancer*, 12: 465.
 30. Entman M L, Youker K, Shoji T, et al (1992). Neutrophil Induced Oxidative Injury of Cardiac Myocytes. A Compartmented System Requiring CD11b/CD18-ICAM-1 Adherence. *J Clin Invest*, 90 (4): 1335-45.
 31. Libby P (2006). Inflammation and Cardiovascular Disease Mechanisms. *Am J Clin Nutr*, 83 (2): 456S-460S.
 32. Li JJ (2011). Inflammation in Coronary Artery Diseases. *Chin Med J (Engl)*, 124 (21): 3568-3575.
 33. Steven S, Frenis K, Oelze M, et al (2019). Vascular Inflammation and Oxidative Stress: Major Triggers for Cardiovascular Disease. *Oxid Med Cell Longev*, 2019:7092151.
 34. Zhang Y, Liu Y, Liu H, et al (2019). Exosomes: Biogenesis, Biologic Function and Clinical Potential. *Cell & Bioscience*, 9: 19.

35. Colombo M, Raposo G, Thery C (2014). Biogenesis, Secretion, and Intercellular Interactions of Exosomes and Other Extracellular Vesicles. *Annu Rev Cell Dev Biol*, 30: 255-89.
36. Nozaki T, Sugiyama S, Koga H, et al (2009). Significance of a Multiple Biomarkers Strategy Including Endothelial Dysfunction to Improve Risk Stratification for Cardiovascular Events in Patients at High Risk for Coronary Heart Disease. *J Am Coll Cardiol*, 54 (7): 601-8.
37. Amabile N, Cheng S, Renard J M, et al (2014). Association of Circulating Endothelial Microparticles With Cardiometabolic Risk Factors in the Framingham Heart Study. *Eur Heart J*, 35 (42): 2972-9.
38. Zhang Y, Vernooij F, Ibrahim I, et al (2016). Extracellular Vesicle Proteins Associated With Systemic Vascular Events Correlate With Heart Failure: An Observational Study in a Dyspnoea Cohort. *PLoS One*, 11 (1): e0148073.