Hypercholesterolemia and Atherosclerosis: A Genetic Approach

Brian Vail

Florida Atlantic University

Abstract

Hypercholesterolemia is characterized by high levels of cholesterol and leads to atherosclerosis. Targeting the genetic associations and associated biological pathways can help researchers study cardiovascular and related diseases. In this article, genes associated with hypercholesterolemia and atherosclerosis are identified in pre- and post-GWAS studies and Gene Enrichment Analysis is performed to identify associated biological pathways. It was found that biological processes involved with cholesterol, sterols, and lipoproteins play significant roles in hypercholesterolemia and atherosclerosis, results which can help direct future research in a variety of diseases, such as hypertension, obesity, diabetes, kidney disease. As well, this article can act as a framework for extending Gene Enrichment Analysis to other diseases such as respiratory diseases, schizophrenia, depression, Alzheimer's disease, diabetes, cancer, etc.

Background

High levels of cholesterol have been associated with increased risk of atherosclerosis and heart disease for decades. Atherosclerosis is a disease where deposits of plaques and fatty acids can build up within the artery walls, thereby impeding blood and oxygen flow and cause further conditions such as heart attacks, strokes, and peripheral vascular diseases. The study of atherosclerosis due to hypercholesterolemia is a wide field, tackled by many different researchers such as physicians, biologists, geneticists and, most recently, even computer scientists. Aspects studied include treatment, prevention and reducing risk factors that impact a range of health-related issues, most notably cardiovascular disease. For example, reducing risk factors, such as cholesterol, through preventative care have shown up to a 30% to 40% reduction in morbidity

and mortality (Shepherd et al., 1995; Downs et al., 1998). Others argue that with time, we could prevent up to 90% of coronary heart diseases (McGill, McMahan & Gidding, 2008).

With recent technological advances, the study of genetic related diseases is on the rise. Utilizing tools such as the GWAS (Genome-Wide Association Study) Catalog and the GO (Gene Ontology) database, this article seeks to pinpoint genes related to atherosclerosis and hypercholesterolemia, and determine related biological pathways that are likely to have critical roles in the disease. The results will help researchers of all kinds find new avenues of study related to atherosclerosis due to hypercholesterolemia, that may extend to different perspectives such as prevention, treatment, reducing risk factors, hypertension, obesity, or coronary heart disease.

Methods

For the literature review section of this paper, research articles were found using the online resources Google Scholar, The National Center for Biotechnology Information (NCBI), and PubMed. Key terms used for search results included “Genetics”, “Atherosclerosis”, “Hypercholesterolemia”, “Prevention”, “Treatment”, “Risk Factors”, and others. The GWAS Catalog was used to expand the list of cholesterol related genes. The search function in the GWAS catalog website (https://www.ebi.ac.uk/gwas/) was utilized to narrow down and identify associated genes. Key terms used for search results included “Cholesterol”, “Atherosclerosis”, “Hypercholesterolemia”, “Obesity”, “Physical Activity”, “Diet”, and others. To prune results, genes were selected from articles that were relevant to the topics of atherosclerosis and hypercholesterolemia and genes resulted in the highest significance levels (minimum p value <

1E-07) from those studies. Identified genes were further identified on the HUGO Gene Nomenclature Committee (HGNC) website and the Online Mendelian Inheritance in Man (OMIM) website, confirming correct gene identification within their database. GO enrichment analysis was performed on the combined list of genes previously identified and the current expanded genes using Gene Ontology (www.geneontology.org). This provided a list of related GO Terms and their statistical significance in terms of their p values. The top ten most statistically significant biological processes were selected. As well, biological processes of interest were also noted so long as their significance was < 1E-07 and the process represented a low-level node in AmiGO 2's inferred tree view. To summarize and visualize the GO term results, GO terms and corresponding p values were entered into ReviGO to generate a scatterplot, interactive graph and treemap. The related GO terms were further explored using the GO database in Web Protégé.

Wikipedia and WebMD were utilized for further understanding of related biological and medical concepts.

Results

In a review article by Roberts (2014), several pre-GWAS identified genes that were associated with high levels of low-density lipoprotein cholesterol (LDL-C) were also found to lead to increased risk of atherosclerotic cardiovascular disease. This included genes such as PCSK9, LDLR, and APOE. Furthermore, he identified genes associated with low levels of high-density lipoprotein (HDL) which leads to increased risk of high cholesterol and atherosclerotic cardiovascular disease. Then genetic mechanisms and their contributions to the disease varied.

For example, the PCSK9 gene creates the enzyme PCSK9 that increases degradation of LDL-C receptors, leading to excess LDL-C and hypercholesterolemia (Seidah, 2003). Mutations of the LDLR receptor gene may lead in hypercholesterolemia as the receptors are a major mechanism for the removal of LDL-C. The ANSK1A gene produces high-density lipoproteins (HDL), which remove cholesterol from the blood. Mutations of this gene lead to increased cholesterol levels and increased risk of atherosclerosis (Roberts, 2014).

Expanding the gene list through the GWAS Catalog resulted in several more genes associated with LDL-C, HDL-C, and triglycerides. Specifically, several trait aspects related to cholesterol were identified, including LDL-C to HDL-C ratio, cholesterol metabolism, triglyceride to HDL-C ratio, and LDL peak particle diameter measurement. This was seen in genes such as CETP, APOC1, and THOC5. Also, in light of recent research into obesity, the potential modulators of the disease risk including diet and physical activity were also included in the GWAS catalog search and provided associated genes such as ARHGEF38, INTS12, and CELSR2.

A summary of the finding can be found in Table 1.

Gene	Variant	Gene	Variant	Gene	Variant
ABCG5 - ABCG8	rs6544713	**CELSR2**	rs599839 rs7528419 rs9989419	**LPL**	rs4244457 rs17482753
ABO	rs579459	**CETP**	rs200751500 rs72786786	**MAIP1**	rs17445774
AC092979.1	rs3819340	**CNNM2**	rs12413409	**MIR31HG**	rs7849420
ANKS1A	rs12205331	**CYP17A1**	rs12413409	**NECTIN2**	rs41290120
APOA1	rs964184	**DSCAML1**	rs145556679	**NT5C2**	rs12413409
APOA4	rs964184	**FNBP1**	rs2007126 rs10760649	**PCSK9**	rs11206510 rs141502002
APOA5	rs964184	**FURIN**	rs17514846	**PSRC1**	rs599839
APOB	rs12713559 rs5742904 rs515135	**GCKR**	rs1260326	**SETD7**	rs706334
APOC1	rs4420638	**GUCY1A1**	rs7692387	**SH2B3**	rs3184504

APOC1P1	rs4420638	**HAPLN4**	rs150641967	**SORT1**	rs599839
APOC3		**HERPUD1**	rs9989419	**TDRD15**	rs562338
APOE	rs2075650	**INTS12**	rs112037309 rs768563000	**THOC5**	rs8135828
ARGFX	rs13096657	**LDLR**	rs151207122 rs121908030 rs201573863 rs137853964	**TRIB1**	rs10808546
ARHGEF38	rs112037309	**LEP**	rs10487505 rs7778167	**ZPR1**	rs964184
BEND3	rs3749872	**LPA**	rs3798220		

Table 1: Identified gene variations associated with hypercholesterolemia and atherosclerosis.

GO term enrichment analysis resulted in four major overarching categories of biological processes associated with risk of hypercholesterolemia and atherosclerosis as related to the identified genes in Table 1: regulation of plasma lipoprotein particle levels, triglyceride metabolism, sterol transport, and macromolecular complex remodeling.

Sterols are a subgroup of steroids, with cholesterol being one of the more widely researched in conjunction with hypercholesterolemia and atherosclerosis. Therefore, it is no surprise the ten most significant biological processes all included pathways related to cholesterol, sterols, lipoproteins, and lipids. Cholesterol and sterol homeostasis is an important biological function. They are needed for many processes including formation of cell membranes, and acting as agents in cellular communication and development. Sterol and cholesterol metabolism and catabolism were previously determined to play a role in hypercholesterolemia and atherosclerosis. These current results relating to homeostasis agrees with the previous findings. However, regarding regulation of sterol levels, in hypercholesterolemia there are excessive levels which lead to atherosclerosis risk. The results of the GO term enrichment also specified lipoproteins as being related to atherosclerosis and hypercholesterolemia. This includes the following biological processes: regulation of plasma lipoprotein particle levels and plasma lipoprotein particle

remodeling. This is in accordance with the previous GO term enrichment analysis that also implicated lipoproteins as being associated with hypercholesterolemia and atherosclerosis.

Looking to the identified low-level nodes for related biological pathways, it can be seen that the most prominent processes deal with cholesterol and lipoproteins. This includes chylomicron assembly. Chylomicrons are lipoprotein particles which are responsible for transporting lipids from the intestines to other parts of the body. An interesting result is that of the regulation of Cdc42 protein signal transduction. Cdc42 is a type of protein responsible for cell division control where molecular signals are mediated.

A summary of the identified biological processes can be found in Table 2 and Table 3.

Biological Process	GO ID	GO Description	P value
regulation of cholesterol transport	GO:0032374	Any process that modulates the frequency, rate or extent of the directed movement of cholesterol into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore.	1.11E-19
regulation of sterol transport	GO:0032371	Any process that modulates the frequency, rate or extent of the directed movement of sterols into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore.	1.11E-19
regulation of plasma lipoprotein particle levels	GO:0097006	Any process involved in the maintenance of internal levels of plasma lipoprotein particles within an organism.	1.17E-19
plasma lipoprotein particle remodeling	GO:0034369	The acquisition, loss or modification of a protein or lipid within a plasma lipoprotein particle, including the hydrolysis of triglyceride by hepatic lipase, with the subsequent loss of free fatty acid, and the esterification of cholesterol by phosphatidylcholine-sterol O-acyltransferase (lecithin cholesterol acyltransferase; LCAT).	2.50E-19
protein-lipid complex remodeling	GO:0034368	The acquisition, loss or modification of a protein or lipid within a protein-lipid complex.	2.50E-19

cholesterol homeostasis	GO:0042632	Any process involved in the maintenance of an internal steady state of cholesterol within an organism or cell.	3.19E-19
protein-containing complex remodeling	GO:0034367	The acquisition, loss, or modification of macromolecules within a complex, resulting in the alteration of an existing complex.	3.33E-19
sterol homeostasis	GO:0055092	Any process involved in the maintenance of an internal steady state of sterol within an organism or cell.	3.66E-19
cholesterol transport	GO:0030301	The directed movement of cholesterol, cholest-5-en-3-beta-ol, into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore.	1.28E-18
lipid homeostasis	GO:0055088	Any process involved in the maintenance of an internal steady state of lipid within an organism or cell.	1.40E-18

Table 2: Ten most significant biological processes associated with hypercholesterolemia and atherosclerosis.

Biological Process	GO ID	GO Description	P value
positive regulation of cholesterol esterification	GO:0010873	Any process that increases the frequency, rate or extent of cholesterol esterification. Cholesterol esterification is the lipid modification process in which a sterol ester is formed by the combination of a carboxylic acid (often a fatty acid) and cholesterol. In the blood this process is associated with the conversion of free cholesterol into cholesteryl ester, which is then sequestered into the core of a lipoprotein particle.	5.54E-11
negative regulation of lipoprotein lipase activity	GO:0051005	Any process that stops or reduces the activity of the enzyme lipoprotein lipase.	6.62E-07
negative regulation of very-low-density lipoprotein particle clearance	GO:0010916	Any process that decreases the rate, frequency or extent of very-low-density lipoprotein particle clearance. Very-low-density lipoprotein particle clearance is the process in which a very-low-density lipoprotein particle is removed from the blood via receptor-mediated endocytosis and its constituent parts degraded.	6.09E-05
chylomicron assembly	GO:0034378	The non-covalent aggregation and arrangement of proteins and lipids in the intestine to form a chylomicron.	8.29E-11
regulation of Cdc42 protein	GO:0032489	Any process that modulates the frequency, rate or extent of Cdc42 protein signal transduction.	9.45E-07

signal transduction			
low-density lipoprotein particle remodeling	GO:0034374	The acquisition, loss or modification of a protein or lipid within a low-density lipoprotein particle, including the hydrolysis of triglyceride by hepatic lipase, with the subsequent loss of free fatty acid, and the transfer of cholesterol esters from LDL to a triglyceride-rich lipoprotein particle by cholesteryl ester transfer protein (CETP), with the simultaneous transfer of triglyceride to LDL.	4.53E-08

Table 3: Low level biological process nodes of interest.

Discussion

According to the American Heart Association and the Center for Disease Control, cardiac disease is the leading cause of death in the United States with over 600 thousand deaths in 2016 (Heron, 2018; Benjamin et al., 2019). While the focus of this article is hypercholesterolemia and atherosclerosis, these are critical aspects to overall cardiac disease and therefore the scope of this article can be widely applied to a variety of research concerning heart health. For example, identified biological pathways may be either a starting point or allowed for deeper research and understanding of related cardiac disease health issues such as hypertension, obesity, diabetes, kidney disease, as well as health behaviors such as physical inactivity, nutrition, and smoking/tobacco use.

The results of the GO term enrichment were centered around cholesterol, which is not surprising as cholesterol has long been associated with the risk of atherosclerosis. As well, it's been found that ratios of sterol precursors are positively related to cholesterol synthesis and negatively related to cholesterol absorption (Miettinen, Tilvis, Kesaniemi, 1989) and other studies have found that the ratios of campesterol, sitosterol and cholestenol are related to the efficiency of cholesterol absorption (Miettinen, Tilvis, & Kesaniemi, 1990). And while researchers do focus on cholesterol in terms of cardiovascular disease research, the results of enrichment analysis can

help to focus specific biological pathways, such as regulation of sterol and cholesterol transport, that can have impact on several related diseases.

The results of the GO term enrichment also specified plasma lipoproteins, which share a degree of overlap with sterol homeostasis. Lipoproteins are the proteins that are combined with fats and other lipids and help to transport them through the bloodstream. As discussed previously, LDL-C is a major risk factor for hypercholesterolemia and atherosclerosis, while cholesterol with higher density lipoproteins are considered healthier. In this way, it is clear to see the regulation of lipoproteins plays an important role as a risk variant for hypercholesterolemia and atherosclerosis.

Analyzing the results of the GO term enrichment analysis, it is clear to see that the primary focus of these genes and identified biological pathways is the levels of LDL-C in those at risk of developing hypercholesterolemia and atherosclerosis. The mechanisms identified (i.e. sterol metabolism, regulation of plasma lipoprotein particle levels, etc.) are almost entirely aimed at processing cholesterol in both metabolic and catabolic pathways, and regulating cholesterol levels. This analysis allows for a deeper understanding of the links between the identified at-risk genes and the expression of hypercholesterolemia and atherosclerosis, and may allow for further, more directed, research of numerous cardiac related diseases.

The most surprising results comes from the low-level node regulation of Cdc42 protein signal transduction. In both literature review and GWAS studies, the Cdc42 gene had not come up, however the enrichment analysis shows a significant association to hypercholesterolemia and

atherosclerosis. The role of regulating communication in cell division as related to the disease of focus in this article warrants further attention and future work, with the consideration of subject matter experts (i.e. biologists or medical professionals).

All identified genes stem from studies with European-ancestry populations with two exceptions. The genes APOC1 and MAIP1 were first identified as being associated with cholesterol levels in a study that focused on an East Asian population (Kim et al., 2017). And the genes CELSR2, CETP, and NECTIN2 were associated with a studying including Hispanic or Latin American, African American or Afro-Caribbean, South East Asian and European populations (Kilpelainen et al., 2019). In future research, the genes will remain in question until their effect may be taken into account in a GO enrichment analysis.

One note of caution is that some of the studies used to identify genes were stand-alone studies, not yet replicated or re-produced. However, studies such as Sandhu et al. (2008) included almost 5000 participants, Kim et al. (2017) included over 2000 participants, and Kilpelainen et al. (2019) included over 120,000 participants. It can be argued that studies with this large an n do not require replication and their results are sound. However, it is important to note as the results of this article are predicated on these studies.

Conclusions

The identification of genes and related biological processes that lead to high risk of hypercholesterolemia and atherosclerosis is an open door for researchers to find new methods of treatment against not only coronary heart disease, but other diseases such as hypertension,

obesity, diabetes, kidney disease. Genes related to LDL-C, HDL-C, and triglycerides, previously found to be associated with hypercholesterolemia and atherosclerosis, were the basis for gene enrichment analysis. It was found cholesterol, sterols, and lipoproteins play significant roles in hypercholesterolemia and atherosclerosis, results which can help direct future research in a variety of diseases. As well, this article may serve as a framework where genes related to other diseases (i.e. respiratory disease, schizophrenia, depression, Alzheimer's disease, diabetes, cancer, etc.) may be identified and analyzed using gene enrichment databases to narrow down specific biological pathways involved/associated with the studied disease.

References

Benjamin, E.J., Muntner, P., Alonso, A., Bittencourt, M.S., Callaway, C.W., . . . Virani, S.S. (2019). Heart Disease and Stroke Statistics—2019 Update. *Circulation, 139*.

Downs, J.R., Clearfield, M., Weis, S., Whitney, E., Shapiro, D.R., . . . Gotto, A.M. Jr. (1998). Primary prevention of acute coronary events with lovastatin in men and women with average cholesterol levels. *JAMA, 279(20), 1615-1622*.

Heron, M. (2018). Deaths: Leading Causes for 2016. *National Vital Statistics Reports, 67(6)*.

Kilpelainen, T.O., Bentley, A.R., Noordam, R., Sung, Y.J., Schwander, K., . . . Loos, R.J.F. (2019) Multi-ancestry study of blood lipid levels identifies four loci interacting with physical activity. Nature Communications, 10(1):376.

Kim, T., Park, A.Y., Baek, Y., & Cha S. (2017). Genome-Wide Association Study Reveals Four Loci for Lipid Ratios in the Korean Population and the Constitutional Subgroup. *PLoS One, 12(1)*.

McGill, H.C. Jr., McMahan, C.A., & Gidding, S.S. (2008). Preventing Heart Disease in the 21st Century: implications of the Pathobiological Determinants of Atherosclerosis in Youth (PDAY) study. *Circulation, 117(9), 1216-1227*.

Miettinen, T.A., Tilvis, R.S., Kesaniemi, Y.A. (1989) Serum Cholestanol and Plant Sterol Levels in Relation to Cholesterol Metabolism in Middle-aged Men. *Metabolism, 38, 136-140*.

Miettinen, T.A., Tilvis, R.S., Kesaniemi, Y.A. (1990). Serum Plant Sterols and Cholesterol Precursors Reflect Cholesterol Absorption and Synthesis in Volunteers of a Randomly Selected Male Population. *American Journal of Epidemiology, 131, 20-31*.

Roberts, R. (2014). Genetic of Coronary Artery Disease: An Update. *Methodist Debakey Cardiovascular Journal, 10(1), 7-12*.

Sandhu, M.S., Waterworth, D.M., Debenham, S.L., Wheeler, E., Papadakis, K., . . . Mooser, V. (2008). LDL-Cholesterol Concentrations: A Genome-Wide Association Study. *Lancet, 371(9611), 483-491*.

Seidah, N.G., Benjannet, S., Wickham, L., Marcinkiewicz, J., Jasmin, S. B., . . . Chretien, M. (2003). The Secretory Proprotein Convertase Neural Apoptosis-Regulated Convertase 1 (NARC-1): Liver Regeneration and Neuronal Differentiation. *Proceeding of the National Academy of Sciences of the United States of America, 100(3), 928-933*.

Shepherd, J., Cobbe, S.M., Ford, I., Isles, C.G., Lorimer, A.R., . . . Packard, C.J. (1995). Prevention of Coronary Heart Disease with Pravastatin in Men with Hypercholesterolemia. *New England Journal of Medicine, 333(20), 1301-1308*.