Abstract
Cardiovascular diseases are affected by multiple factors like genetic as well as environmental hence they reveal factorial nature. The evidences that genetic factors are susceptible for developing cardiovascular diseases come from twin studies and familial aggregation. Different ethnic populations reveal differences in the prevalence of coronary artery disease (CAD) pointing towards the genetic susceptibility. With progression in molecular techniques different developments have been made to comprehend the disease physiology. Molecular markers have also assisted to recognize genes that may provide evidences to evaluate the role of genetic factors in causation of susceptibility towards CAD. Numerous studies suggest the contribution of specific "candidate genes", which correlate with various roles/pathways that are involved in the coronary heart disease. Different studies have revealed that there are large numbers of genes which are involved towards the predisposition of CAD. However, these reports are not consistent. One of the reasons could be weak contribution of genetic susceptibility of these genes. Genome wide associations show different chromosomal locations which dock, earlier unknown, genes which may attribute to CAD. In the present review different ApoAI-CIII-AIV gene clusters have been discussed.

Key words: ApoAI-CIII-AIV gene cluster; Haplotype analysis; Single nucleotide polymorphism; Candidate gene study; Genome wide association studies

Core tip: Cardiovascular disease analysis requires holistic approach using genomic, epigenomic and exposomic techniques to improve the quality of life of patients and contribution towards personalised medicine.

INTRODUCTION
Coronary artery disease (CAD), is mostly fatal if remain untreated result into atherosclerosis in the epicardial coronary arteries[1]. Atherosclerotic plaques progressively narrow the coronary artery lumen and impair antegrade myocardial blood flow. This reduction in coronary artery flow may lead to a myocardial infarction.

Cardiovascular disease is a multifarious disorder showing large diversity of phenotypes. The accurate, and analogous phenotypic evidences are crucial for detailed understanding of the affiliation between disease and genes, as well as understanding the role of various extrinsic factors on different component of various genotypes.
This complexity also contributes to difficulties in diagnosis and prognosis of the disease. Diagnostic difficulties also hamper the optimal and personalised treatment for patients. In recent years the role of genetic variability on the development of CAD has been extensively been studied which is impacting upon our understanding of phenotypic outcomes and clinical complications. New developments in genomics, epigenomics and exposomics (environmental risk factors across the life span) would result into the improved understanding of the different phenotypes observed in CAD and would help in the better regimen of treatment. In the last century, there has been rapid increases in the global prevalence of CAD, which has become the important cause of cardiovascular mortality all over the world, i.e., > 4.5 million deaths in the developing countries. By 2020, it is predictable that CAD will be the major source of disease burden universally. The prevalence of CAD varies in different ethnic groups which may show higher/lower genetic and environmental susceptibilities. India has also witnessed consistent increases in the prevalence of CAD over the past few decades and could become the number one killer if appropriate interventions are not planned and implemented. In Table 1 the incidence of CAD is shown in different parts of India.

It has been reported that CAD is increasing in a linear fashion as it has increased from 4% in 1960 to 11% in 2001, almost every 25th individual in 1960 was having CAD, while in 2001 every 9th individual was having CAD. The CAD is declining internationally among Indians settled abroad, whereas, these rates are growing in the Indian subcontinent. Presently, 10%–12% of metropolitan Indians have CAD compared to 3% of the United States population. Many studies document that Asian Indians are at 3-4 times greater risk of CAD than white Americans/Europeans, 6 times higher than Chinese, and almost 20 times higher than Japanese. CAD prevalence has increased from 3.5% in the 1960s to 9.5% in the 1990s in urban populations of India. Current studies recognized occurrence of CAD to be 13.9% in the urban south Indians, 9.6% in urban north Indians. In 1990s, 33% of cardiovascular deaths have been reported from India and it is likely that deaths from non-communicable diseases such as CAD will increase two times higher i.e., 4.5 million in 1998 to 8 million in 2020 in India.

As CAD has a multifactorial nature and the occurrence of the familial clustering in CAD led investigators to start searching for susceptibility genes. In Figure 1 different risk factors involved in the causation of CAD are summarized.

It is vital to keep in mind that certain genes may show population specific effects. There are hundreds of genes known to have functional allelic variations that may be important for determining an individual's vulnerability to CAD. There is much argument on results of published epidemiological studies until now. The differences may be due to differences in the techniques used or the population used to calculate the incidence, prevalence and other risks. It has been proposed that if multiple markers are used for assigning the risk, the results would be more conclusive clinically. Most important reason of concern in developing countries like India is the incomplete detection, treatment and control of CAD risk factors. The benefits of addressing the root cause of CAD, such as inflammation, smoking and cholesterol, together with preventive methodology will be useful in improving quality of life and saving lives. This in turn may be translated into preventive approaches to help reduce the risk of CAD using genetic and epigenetic approaches. Although CAD mortality in the Indians is highest than other populations, the reason for increased risk, which has been recorded in both the Asian immigrants and among Indians in urban India; are not yet clear hence more systematic and comprehensive studies are required to understand the spectrum of genetic and epigenetic influences on CAD.

### GENETIC BASIS OF CAD

Atherosclerosis involves multiple factors, hence understanding the genetic and environmental basis of this complex disease requires holistic approaches. A range of candidate genes (e.g., APOE, APOB, LPL, iNOS, ACE, COX2, CD14, P-Selectin, E-Selectin, MTHFR, PON1, TNFz) have been investigated in relation to initiation, development and progression of CAD. A large number of studies using of candidate genes and genome-wide association analyses have shown some promising signals, but only a few have been confirmed to some extent which may be playing a role in CAD.

There are very few examples where single genes have played a role in causing atherosclerosis. Mostly, CAD is caused by the environmental factors however the risk increases when some risk associated genes are also present. Research on identical twins consistently shows significant genetic effect in the development of CAD or its risk factors (Table 2). Heritability for CHD vary from 40% to 60%, suggesting a strong role of genes in the development of the disease. A detailed analysis of the many known CAD susceptibility genes and studies is be-

### Table 1 Prevalence of coronary artery disease in different Indian surveys

<table>
<thead>
<tr>
<th>City</th>
<th>Prevalence</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban population</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chandigarh</td>
<td>(6.60%)</td>
<td>Sarvatham et al[49]</td>
</tr>
<tr>
<td>Rohtak</td>
<td>(3.80%)</td>
<td>Gupta et al[50]</td>
</tr>
<tr>
<td>Jaipur</td>
<td>(7.60%)</td>
<td>Gupta et al[51]</td>
</tr>
<tr>
<td>Delhi</td>
<td>(9.70%)</td>
<td>Chada et al[52]</td>
</tr>
<tr>
<td>Rural population</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jaipur</td>
<td>(3.50%)</td>
<td>Gupta et al[50]</td>
</tr>
<tr>
<td>Ludhiana</td>
<td>(8.08%)</td>
<td>Wander et al[54]</td>
</tr>
<tr>
<td>South Indians</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamil Nadu</td>
<td>(14.30%)</td>
<td>Ramachandran et al[53]</td>
</tr>
<tr>
<td>Tamil Nadu</td>
<td>(11.00%)</td>
<td>Mohan et al[54]</td>
</tr>
<tr>
<td>Migrant Indians</td>
<td></td>
<td></td>
</tr>
<tr>
<td>London, United Kingdom</td>
<td>(17.00%)</td>
<td>Bahl et al[55]</td>
</tr>
<tr>
<td>Illinois, United States</td>
<td>(10.00%)</td>
<td>Enas et al[56]</td>
</tr>
</tbody>
</table>
Beyond the scope of this overview. This overview will focus on selected candidate genes in the ApoAI-CIII-AIV gene region.

**SINGLE GENE DISORDERS AND CAD**

**Familial hypercholesterolemia**

Familial hypercholesterolemia (FH) is a classic genetic disease in which increased cholesterol, tendon xanthomas, and early heart disease segregates together. Joseph Goldstein and Michael Brown showed that FH results from mutations in the low-density lipoprotein (LDL) receptor, which leads to impaired binding, internalization and degradation of LDL. Dose dependent relationship was observed, homozygotes patients had higher levels of cholesterol (> 600 mg/dL), whereas heterozygotes had levels of approximately 400 mg/dL. This variable penetrance is modified by genes and other risk factors such as diet, smoking, and physical activity level[23]. Heterozygote frequency for this disease relatively high, approximately 1 in 500[24] in most populations, however DNA screening and effective treatments are available now[25,26].

**Familial defective apolipoprotein B causing hypercholesterolemia**

This comparatively common hypercholesterolemia (approximately 1 in 800), results from mutations in the major protein of LDL called Apolipoprotein B (ApoB). The mutations in ApoB prevent LDL binding to the LDL receptor. The majority of patients of this disorder carry a dominant mutation (codon 3500) and have lower cholesterol levels compared to FH patients. Other single-gene CHD/CAD traits are rare and of lower clinical/population significance[25].

**CANDIDATE GENES AND CAD**

During last 30 years, there have been many advancements in molecular genetic technology, development of sophisticated statistical tools and analyses which have contributed to improvements in human genetic research. One of the early developments was positional cloning technique, which allowed genetic mapping of many Mendelian diseases and traits. However for complex diseases, which involve many genes and environmental influences, this technique did not provide any major insights into genetic basis. Majority of our understanding of the genetic basis of CAD/CHD has been gained from studies of “candidate genes,” and more recently genome wide association (GWA) studies. These population based studies have provided further insights into genetic susceptibilities/contributions to complex diseases. Some examples of these are given below.

**APOLIPOPROTEIN E AND APOAI-CIII-AIV GENE CLUSTER**

Apolipoprotein E (ApoE) is one of the extensively studied genetic locus as it plays a pivotal role in lipid metabolism and mediates the uptake of chylomicron and very low-density lipoprotein (VLDL) remnants. Utermann and colleagues[27] identified genetic polymorphism at ApoE locus and its association with cholesterol levels and type III hyperlipidemia. The polymorphism and its CAD associations have been replicated in many global populations. E3 allele is the most common (approximately 60%) followed by E4 allele (approximately 30%) and E2

---

**Table 2 Genetic and environmental risk factors for coronary heart disease**

<table>
<thead>
<tr>
<th>Risk factors</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevated LDL and VLDL cholesterol</td>
<td>40%-60%</td>
</tr>
<tr>
<td>Low HDL cholesterol</td>
<td>45%-75%</td>
</tr>
<tr>
<td>Elevated triglycerides</td>
<td>40%-80%</td>
</tr>
<tr>
<td>Increased body mass index</td>
<td>25%-60%</td>
</tr>
<tr>
<td>Elevated systolic blood pressure</td>
<td>80%-70%</td>
</tr>
<tr>
<td>Elevated diastolic blood pressure</td>
<td>50%-65%</td>
</tr>
<tr>
<td>Elevated lipoprotein(a) levels</td>
<td>90%</td>
</tr>
<tr>
<td>Elevated homocysteine levels</td>
<td>45%</td>
</tr>
<tr>
<td>Type 2 diabetes mellitus</td>
<td>40%-80%</td>
</tr>
<tr>
<td>Elevated fibrinogen</td>
<td>20%-50%</td>
</tr>
<tr>
<td>Elevated C-reactive protein</td>
<td>40%</td>
</tr>
<tr>
<td>Elevated homocysteine levels</td>
<td>45%</td>
</tr>
</tbody>
</table>

Risk factors with a significant genetic component (heritability)

Gender
Family history
Environmental risk factors
-smoking
-diet
-exercise
-infection
-foetal environment
-air pollution (particulates)

Risk factors for coronary heart disease can be subdivided into those that are determined significantly by genetic differences and those that are largely environmental (Based on Lusis et al[58] 2004). VLDL: Very low density lipoprotein; LDL: Low density lipoprotein; HDL: High density lipoprotein.

---

Figure 1 Cardiovascular risk factors.

Figure 1 Cardiovascular risk factors.

Table 2 Genetic and environmental risk factors for coronary heart disease

Risk factors for coronary heart disease can be subdivided into those that are determined significantly by genetic differences and those that are largely environmental (Based on Lusis et al[58] 2004). VLDL: Very low density lipoprotein; LDL: Low density lipoprotein; HDL: High density lipoprotein.
(approximately 10) in world populations. E4 allele carriers have increased plasma cholesterol levels compared to E3 allele carriers while E2 carriers have decreased plasma cholesterol levels. The allelic variation at ApoE locus explains approximately 5% of the variation in cholesterol levels. Type III hyperlipidemia, a relatively rare phenotype, are homozygous for the E2 allele, but not all E2 homozygous individuals have this disorder. Therefore, genotype-phenotype relationships may require contribution of other genetic or environmental factors.

In addition to ApoE, there is now strong evidence that mutations in hepatic lipase influence the levels of high-density lipoprotein (HDL), and the ApoAI-CIII-AIV-AV locus contributes to plasma triglyceride levels. Many studies have shown that Lp(a) levels are strongly influenced by Apo(a) gene. In addition, both hepatic lipase and the ApoAI-CIII-AIV-AV cluster influence LDL particle size, which significantly contributes to CHD risk. However, taken together, these genetic differences only explain a small amount of variation in plasma lipids and CHD/CAD phenotypes.

Dyslipidemia, a metabolic disorder, caused due to the defects in the synthesis, processing and catabolism of lipoprotein particles. Increased total cholesterol (TC), triglyceride (TG), LDL cholesterol (LDL-C), and apolipoprotein (Apo)B together with lower levels of ApoA1 and HDL cholesterol (HDL-C) have been found to increase coronary artery disease (CAD) risk. Epidemiological and clinical studies have documented that above genetic factors/polymorphisms play a significant role in dyslipidemia susceptibilities along with environmental factors. Twin and family studies suggest there are considerable genetic contributions in the inter-individual variation in plasma lipid phenotypes with the heritability estimates ranging from 40%-60%. It has been suggested that understanding variation at these loci along with other newer genetic loci will provide a better understanding of the disease processes and contribution to personalized medicine.

ApoA1, is the main protein component of HDL-C, it functions in the activation of lecithin: cholesterol acyltransferase, and facilitates the reverse cholesterol transport from peripheral tissues. ApoC3, is a 79-amino-acid protein formed mainly in the liver, is one of the major component of chylomicrons and VLDL and a minor component of HDL. ApoC3 prevents protein lipase and plays a key role in the catabolism of TG-rich lipoproteins. ApoA5 is detectable in very low-density lipoprotein, HDL, and chylomicrons and its concentrations are low compared to other apolipoproteins. Human ApoA1/C3/A5 genes resides in the ApoA1/C3/A4/A5 gene cluster on chromosome 11q23-q24. The ApoA1/C3/A4/A5 gene cluster has emerged as a significant risk factor for hypertriglyceridemia and atherosclerosis. A number of studies have shown significant associations between single nucleotide polymorphisms (SNPs) in the ApoA1/C3/A4/A5 gene cluster and raised plasma or serum lipid levels in humans, while others have reported negative or inconsistent results. In addition there are many other SNPs involved in the inflammation and cell signalling with CAD and/or MI, some of these are summarized in Table 3.

One of the limitations of case control studies is that many false positive or false negative associations may emerge between different genetic markers and complex diseases like CAD. The reason for such results are: (1) controls are not properly selected; (2) sample size of both controls and cases because of which accurate power of the study is not generated and replication of results is not possible; and (3) position of single-nucleotide polymorphisms (SNP’s) in terms of their effect on transcription of gene or protein expression. In general, results of small sample size studies (200-300 patients and control subjects) should be interpreted with caution and should be replicated with larger sample sizes. It is important to confirm that genotype distributions are not skewed, especially in the control group. Large deviations from the Hardy-Weinberg equilibrium, may suggest that the control group is not necessarily the representative of healthy and randomly sampled individuals. This departure may also highlight issues with genotype scoring. Recent genome-wide sequencing research has revealed extensive level of variation and heterogeneity between individuals and populations, which should be considered when choosing SNPs and interpreting SNP data. Some of the early SNP association studies failed to include the effect of the polymorphism on gene expression or protein function and genotype-phenotype correlations. This information could reveal if an SNP is the actual cause or solely a marker which may be in linkage disequilibrium another causal variant. These analyses could provide significant clues for understanding the pathophysiologic mechanisms behind clinical outcomes. It is important to correct/control for the age, gender, ethnicity, and other confounders in heart disease genetic association studies. There should be a holistic approach to understand the role of genes, environment and life style factors in CAD susceptibilities and progression.

Recently, genetic analyses have expanded to whole genome sequence analysis and genome-wide association studies (GWAS) as these analyses eliminates biases in the selection of the candidate genes. A number of GWAS studies have identified new loci in previously unsuspected genomic regions. These analyses have shown, novel biological pathways involved in the disease states and development of novel therapies. Many recent studies have shown only limited evidences may exist where the genetic variants may be associated with MI or only with CAD. A care has to be taken in interpreting the GWAS data as large number of variant alleles may be found but one should consider only elegant systems genetics approach to Plaisier et al. used similar approach and found that FADS3 is a causal gene for familial combined hyperlipidemia and elevated triglycerides in Mexicans. The authors used network gene co-expression analysis and SNP data to assign a function to the genetic variants of healthy and randomly sampled individuals. This departure may also highlight issues with genotype scoring. Recent genome-wide sequencing research has revealed extensive level of variation and heterogeneity between individuals and populations, which should be considered when choosing SNPs and interpreting SNP data. Some of the early SNP association studies failed to include the effect of the polymorphism on gene expression or protein function and genotype-phenotype correlations. This information could reveal if an SNP is the actual cause or solely a marker which may be in linkage disequilibrium another causal variant. These analyses could provide significant clues for understanding the pathophysiologic mechanisms behind clinical outcomes. It is important to correct/control for the age, gender, ethnicity, and other confounders in heart disease genetic association studies. There should be a holistic approach to understand the role of genes, environment and life style factors in CAD susceptibilities and progression.

Recently, genetic analyses have expanded to whole genome sequence analysis and genome-wide association studies (GWAS) as these analyses eliminates biases in the selection of the candidate genes. A number of GWAS studies have identified new loci in previously unsuspected genomic regions. These analyses have shown, novel biological pathways involved in the disease states and development of novel therapies. Many recent studies have shown only limited evidences may exist where the genetic variants may be associated with MI or only with CAD. A care has to be taken in interpreting the GWAS data as large number of variant alleles may be found but one should consider only elegant systems genetics approach to Plaisier et al. used similar approach and found that FADS3 is a causal gene for familial combined hyperlipidemia (FCHL) and elevated triglycerides in Mexicans. The authors used network gene co-expression analysis and SNP data to assign a function to the genetic variants.
This overview has highlighted some of the important factors involved in inflammation and cell signalling with coronary artery disease and/or myocardial infarction.

Table 3  Example of association studies of factors involved in inflammation and cell signalling with coronary artery disease and/or myocardial infarction

<table>
<thead>
<tr>
<th>Gene</th>
<th>Polymorphism</th>
<th>Ref.</th>
<th>Suggested results</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRP</td>
<td>1059 G/C</td>
<td>Zee et al[46]</td>
<td>No significant association with non-fatal MI, stroke or cardiovascular death</td>
</tr>
<tr>
<td>ICAM-1</td>
<td>Lys-469-glu, Ser-128-Arg, Leu-554-phe, G98T</td>
<td>Jiang et al[47]</td>
<td>Association with MI and CAD</td>
</tr>
<tr>
<td>E-selectin</td>
<td>Ser-128-Arg, G98T</td>
<td>Herrmann et al[48]</td>
<td>Associated with angiographic proof of severe CAD in patients &lt; 50 yr</td>
</tr>
<tr>
<td>P-selectin</td>
<td>Ser-128-Arg, Pro715</td>
<td>Ye et al[49]</td>
<td>T allele more common in younger patients with angiographic CAD</td>
</tr>
<tr>
<td>TNF-α and β</td>
<td>C-212G, A-196G, Thr715Pro</td>
<td>Barbaux et al[50]</td>
<td>Polymorphisms associated with P-selectin levels but not with MI</td>
</tr>
<tr>
<td>TNF-β</td>
<td>-863C/A, -308G/A (TNF-α), 252G/A</td>
<td>Koch et al[51]</td>
<td>No association of TNF or IL-10 polymorphisms with MI or CAD</td>
</tr>
<tr>
<td>IL-1 cluster</td>
<td>IL-1α (-889), IL-1β (-511),</td>
<td>Keso et al[52]</td>
<td>No association to old MI by autopsy or CAD</td>
</tr>
<tr>
<td>IL-1-RA</td>
<td>IL-1α intron 2 VNTR</td>
<td>Francis et al[53]</td>
<td>No association to angiographic CAD</td>
</tr>
<tr>
<td>IL-1-RA</td>
<td>IL-1β intron 2 VNTR</td>
<td>Vohnout et al[54]</td>
<td>IL-1RA VNTR allele 2 associated with single-vessel CAD</td>
</tr>
<tr>
<td>IL-1α, IL-1β, IL-RA</td>
<td>IL-1α, IL-1β intron 2 VNTR</td>
<td>Zee et al[55]</td>
<td>No association to angiographic CAD with either polymorphisms</td>
</tr>
<tr>
<td>IL-6</td>
<td>IL-6 G (-174) G/C promoter polymorphism</td>
<td>Nauck et al[56]</td>
<td>No association with risk for future MI</td>
</tr>
<tr>
<td>IL-10</td>
<td>3 IL-10 promoter polymorphisms</td>
<td>Koch et al[57]</td>
<td>IL-1β (-511)/C/C and IL-1-RA (intron 2)/2- or 3- repeat allele both associated with CAD, association with MI only in patient who are seropositive for Chlamydia pneumonia</td>
</tr>
<tr>
<td>TGF-β1</td>
<td>29 T/C, -509T</td>
<td>Donger et al[58]</td>
<td>No association with the risk for CAD or MI</td>
</tr>
<tr>
<td>Stromelysin (MMP-3)</td>
<td>5A-117/6A promoter polymorphism</td>
<td>Syrris et al[59]</td>
<td>-174 C associated with MI (OR = 1.34)-174 C more frequent in patients with two or fewer stressed vessels than in patients with three vessel lesions</td>
</tr>
<tr>
<td>PECAM-1 (CD31)</td>
<td>Val 125Leu, Asn563Ser and Gly670Arg</td>
<td>Sasaoaka et al[60]</td>
<td>Association with early-onset CAD</td>
</tr>
</tbody>
</table>

This overview has highlighted some of the important factors involved in inflammation and cell signalling with coronary artery disease and/or myocardial infarction.

rs3737787 (1q21-q23) in USF1 gene, which was previously identified to be associated with FCHL, It is envisaged that new methods like Network medicine[61] will play an important role in these analyses and the advancement of our understanding of pathophysiological mechanisms of diseases like CAD and MI.

CONCLUSION

This overview has highlighted some of the important
challenges regarding the use of genetic approaches to investigate complex diseases. The recent research using genomic, epigenomics and exposomic approaches is providing a range of patient centric tools which will help better classification of phenotypes and personalised medicine for CAD patients. The mechanisms underlying the association of these loci to CAD/MI remain largely unknown and the effects are relatively small. Hence the future challenges are (1) discovering new genetic variants through large-scale meta-analyses, using pathway-based approaches, and high throughput sequencing; (2) illustrating the mechanisms for the identified loci to CAD; and (3) translating the findings from CAD-GWASs and epigenetic analyses to novel and optimized therapeutic strategies.

REFERENCES


30 Lusis AJ, Ivanidc B, Castellani LW. Lipoprotein and lipid


38 Kilis¸-Pstrusin’ska K. Genetic factors in the development and progression of chronic kidney disease. Postepy Hig Med Dosw 2010; 64: 50–57


Agrawal S et al. Genetic contribution of apolipoproteins in CHD


Wenze K, Baumann G, Felix SB. The homoyzogous combination of Leu125Val and Ser663Asn polymorphisms in the PECAM1 (CD31) gene is associated with early severe coro...


P- Reviewer: Iacoviello M, Ong HT  
S- Editor: Wen LL  
L- Editor: A  
E- Editor: Wu HL