The evidence for heritability of myocardial infarction (MI) is striking, with a positive family history being one of the most important risk factors for this complex trait (1). The term “complex trait” implies both gene-environment and gene-gene interaction, along with the complexity of inheritance distinct from “simple” Mendelian traits such as those with an autosomal-dominant, autosomal recessive, X-linked, or mitochondrial basis (2). Whereas a complex trait is common and probabilistic in phenotypic manifestation, a Mendelian trait is considered to be rare and deterministic (2). Willett (3) has pointed out that more than 80% of coronary heart disease may be accountable by lifestyle issues, such as weight, diet, exercise, and control of risk factors such as blood pressure and smoking. Yet, the influence of genetics is also quite important, and recent work in this field has led to determining key genes associated with increased risk of susceptibility for acute MI. This paper will review the rapid and substantive progress that has been made in identifying genes that influence susceptibility for MI.

GENETICS OF ATHEROSCLEROSIS AND MI ARE DIFFERENT

Atherosclerotic coronary artery disease (CAD) is nearly pervasive, and one can conservatively project well over 50 million American adults have some atheromatous disease. This is based on autopsy specimens of young and old adults, along with assessment by intravascular ultrasound in individuals without symptoms of angina or ischemic heart disease (4,5). Yet despite this very large universe of patients with atherosclerotic involvement of their coronary arteries, only a small fraction will ever develop an MI—in the U.S. this is estimated to be approximately one million people per year (6). In virtually all of our genetic studies to date, we have noted that the heritability of MI is much more impressive than CAD (1,7). The likely explanation for this observation is that only a limited number of individuals have particular susceptibility to have active arterial inflammation culminating in plaque rupture, erosion, or fissuring to induce an acute coronary syndrome. Recently, Fischer et al. (8) studied 401 families with MI by comparing angiograms among siblings. Remarkable concordance among 882 siblings was noted for the location of the lesions in the coronary tree, with statistical significance for ostial, left main stem, and proximal locations. Not only having underlying coronary disease, but also in the same location in the coronary tree, further emphasizes the importance of the genetic basis of this disease in families with MI. In order to understand the actual genes responsible for MI susceptibility, direct assessment of the genome of individuals is requisite.

THE HUMAN GENOME

The human genome is much simpler than originally projected when the draft sequence was announced in June 2000 (9). There are only approximately 26,000 genes, giving rise to about 100,000 proteins (more proteins than genes because of alternative splicing). Remarkably, individual heterogeneity of the 3.1 billion base pairs (bp) from person to person is based upon only 0.1%, or 3 million bps at play of the total 10 million single nucleotide polymorphisms (SNP). In Figure 1, a simplified graphic depiction of an SNP, haplotype, and microsatellite is provided. An SNP is simply a bp in which a coding letter of the four nucleotides (A, G, T, C) can be substituted. A haplotype is a collection of SNPs, often extending over thousands of bases. A microsatellite refers to a tandem repeat sequence much longer than that illustrated, and useful not only as a marker in the genome but also, at times, responsible for harboring disease-causing genes. These terms are important to understand the progress that is being made in high throughput genotyping, genomewide scans, and whole genome SNP mapping.

In Figure 2, the jumps in the field of complex trait determination are summarized in a ladder-type diagram. Starting at the bottom with case-control association studies involving a single SNP, we have moved to studying hundreds or thousands of SNPs simultaneously as an outgrowth of high throughput genotyping. Beyond these methods, genome-wide scanning using microsatellites is a classic approach to be described more thoroughly below. Very dense SNP mapping, involving millions of SNPs, may ultimately take its place. If it can be provided
without prohibitive expense, complete re-sequencing of
the individual genome could indeed lie ahead in the
future. As one goes up the ladder, there is with each
technique progressively less bias, more expense, and more
information. Advances in technology of performing
genotyping are reflected by the cost moving from $1.00
per bp analysis to a projection of a hundredth of a cent in
coming years (2).

**PROGRESS IN SNPS**

The most extensively studied polymorphism in cardio-
vascular medicine is apolipoprotein (apo) E (10). In this
gene, amino acid substitutions at residues 112 and 158
are responsible for three variants: apoE2 (cysteine, cyste-
ine), apoE3 (cysteine, arginine), and apoE4 (arginine,
arginine), respectively. A meta-analysis of 48 studies
involving over 15,000 cases and 32,000 controls has
shown the odds ratio for CAD is at least 1.4 for apoE4
carriers compared with E3 controls. This same polymor-
phism, present in 20% of the population, has been
associated with a heightened risk of Alzheimer’s disease
and reduced lifespan (11). Its replication for association
with CAD across tens of population cohorts is exemplary,
but the functional genomics of precisely how the apoE4
polymorphism sets up susceptibility for MI and CAD is
incompletely understood. Most patients with this poly-
morphism have high low-density lipoprotein cholesterol
levels, and exquisite sensitivity to statins for clinical
outcome benefits has been noted (12).

The work in apoE4 is largely derived from single-allele
case-association studies. In recent years, high throughput
SNP studies have become more commonplace. The first
high throughput study in MI demonstrated the impor-
tance of the thrombospondin (TSP) family of matricellu-
ar proteins for association with premature MI in over
400 affected families (7). Three different TSP SNPs—in
different TSP genes 1, 2, and 4—prevailed in association
with premature MI among over 100 candidate genes that
were assessed (7,13). Of note, the TSP-4 SNP is a
common SNP with the minor allele present in over 30%
of individuals and conferring a near two-fold risk of MI.
This leads to a substitution of proline for alanine in the
TSP-4 protein, and with that a marked increase in
calcium binding for this key type II repeat binding site of
the TSP-4 protein. This “gain-of-function” mutation is
correlated with extensive proinflammatory effects includ-
ing a striking reduction in endothelial proliferation or
inability to repair from endothelial injury (Fig. 3), and
enhanced neutrophil function (14). The association of
this TSP-4 SNP has been independently replicated in
other case-control studies (15–17). A number of other
SNPs have been similarly validated for MI for CAD in at
least three independent studies, with functional genomic
work, as provided in Table 1 (7,13,18–21).

Ozaki et al. (19) in Japan have taken the high-
throughput SNP study field to a new threshold. By
analyzing more than 65,000 SNPs in 13,738 genes, these
investigators found 2 SNPs in lymphotoxin alpha (LTA)
gene exon 3 with a highly significant susceptibility for MI
(p = 0.00000033, odds ratio 1.8). Furthermore, the LTA
SNPs were shown to be proinflammatory by increasing
expression of vascular cell adhesion molecule and
E-selectin. This important and groundbreaking effort
was recently extended by the discovery of a galectin-2
SNP, from the galactose-binding lectin family, represent-
ing a critical ligand to LTA (22). Galectin-2, in binding
to LTA, also regulates its secretion. Both macrophages
and smooth muscle cells in atheroma express LTA and
galactin-2. In over 4,300 patients assessed, an intron-1
SNP of galectin-2 was found to be highly significant for
association with MI (p = 0.00000026) with the TT allele
conferring protection and the CC allele susceptibility.
This discovery nicely illustrates the extension of a pri-
mary finding in MI genomics. First, identifying the LTA
anchoring effect, then moving on to explore its principal
ligand, galectin-2, thereby confirming its pivotal impor-
tance in the genomics of MI. Thus, the acceleration of
knowledge base in the field is well exemplified through
the previously established biologic pathways.
GENOME-WIDE SCANNING AND LINKAGE

The classic approach to determining linkage of a complex trait has involved the technique of genome-wide scanning with microsatellites. This involves selecting approximately 400 microsatellite markers across the genome that are evenly spaced every 10 centimorgan (approximately 1 million bps) (Fig. 4). The goal is to find a locus or "linkage peak" in which there is extensive allele sharing (and with some analyses lack of allele sharing of unaffected individuals) for the cohort of sibling pairs with a particular phenotype. This can be conceived as a simple "frequency distribution" of an allele and that in a general population there would not be any signal or linkage peak present. The critical steps in identifying a significant locus involve the following: 1) a large population of families (typically >300 with >1,000 affected individuals) with a well characterized and "restrictive" phenotype; 2) a logarithm of the odds ratio (LOD) score more than 3.5 corresponding to a p < 10^-6 (23). The finding of a significant linkage peak translates to having a gene or
genes that are in linkage disequilibrium (close in location) with a microsatellite marker. Once the linkage peak is identified, the genes in the region can be assessed to determine which one(s) is responsible.

This technique has already been successfully applied to a number of complex traits (Table 2). In each case a significant locus was ascertained, and then techniques such as fine mapping, positional cloning, or candidate gene testing were used to identify the specific gene. It is interesting to note that while we have seen over 1,600 Mendelian human traits determined, only about 15 complex traits have been thus far identified for their genomic basis (24). Accordingly, we are in the early, nascent phase of complex trait genomics.

A number of genome-wide scans have been performed in CAD and MI (1,25–31). As summarized in Table 3, there have been seven such studies in cumulatively more than 2,000 families. Each study has found different loci, but the populations differed substantially on the basis of sample size, ethnicity, age, and analysis program used to determine linkage. Furthermore, most used CAD as the phenotype rather than MI. Thus far, of the seven studies, only one has determined the specific gene accounting for a significant linkage peak.

At Cleveland Clinic, in collaboration with a number of centers in the U.S., Wang et al. (1) have reported a highly significant linkage peak on chromosome 1p34-36. In 428 multiplex families, involving 1,613 individuals, with age of MI at 44 years, several significant loci were identified by genome-wide scanning (Fig. 5). Of these, the chromosome 1 locus was highly significant with a LOD score of over 11 and genome-wide significance. This LOD score would correspond to a <1/100 million chance of finding such a linkage peak in a general, unselected population. There are over 300 genes in this broad locus spanning 32 cM, with significance across four microsatellite markers. In the months ahead, it is hoped the identity of the specific gene that accounts for this locus will be finalized. Of note, we did not find any linkage to the trait of CAD but found the important evidence of linkage disequilibrium with the premature MI phenotype.

The DeCode Genetics group (30) in Iceland have been successful in mapping a locus for both MI and stroke to chromosome 13q12-13 and identifying 5-lipoxygenase activating protein (FLAP) as the gene that accounts for this linkage peak. This finding has been confirmed in British, Scottish, and, most recently, American cohorts. The FLAP “gain-of-function” haplotype consists of a four-SNP marker that is present in approximately 10% of.
Figure 5. (A) Haseman-Elston sib-pair regression analysis for scanning loci segregating with myocardial infarction. The vertical axis of each plot is \(-\log_{10}(P)\) or \(P\), where the \(P\) is the significance level from each of two analyses. The solid line denotes the multipoint linkage profile for all pedigrees; the broken dotted/dashed line denotes the multipoint linkage profile for Caucasian-only pedigrees. The horizontal solid line in each subfigure indicates \(P = 2.2 \times 10^{-5} (-\log_{10}(P) = 4.66\) or logarithm of the odds ratio score = 3.6). The \(x\) axis denotes marker map positions. Note that the dashed line (\(p\) values for Caucasians only) is often not visible because it frequently overlaps with the solid line (\(p\) values for all study families). Continued on next page.
individuals with an MI. It carries a near two-fold risk for both MI and stroke. This gene is critical to the transformation of 5-lipoxygenase (5-LO) to leukotriene B4, one of the most potent neutrophil cytokines responsible for inflammation. The 5-LO pathway has been strongly implicated for atherosclerosis in transgenic mice, and man (32,33), and the presence of FLAP and 5-LO is indexed to the complexity of the coronary atherosclerotic lesion, with a positive correlation of plaque rupture and heightened 5-LO/FLAP (34). Recently, the DeCode investigators have performed a randomized trial of a FLAP blocker, conventionally used in patients with asthma, in over 200 patients with the FLAP gain-of-function haplotype. The FLAP inhibitor, compared with placebo, led to a significant reduction of neutrophil leukotriene B4 production and decreased serum C-reactive protein (35). This represents the first gene-specific therapeutic program in the field of MI—with a personalized or individualized approach of administering a drug to counter the genetic susceptibility. Its true validation will rely on demonstration of clinical outcome improvement in patients with the abnormal, at-risk FLAP haplotype, but this may be a harbinger of a new era of smarter therapeutics that are genomically based in the future.

Wang et al. (36) have used a rich pedigree genome-wide scanning approach to identify the transcription factor myocyte-enhancing factor 2A (MEF2A) as a cause of autosomal-dominant inheritance of MI and CAD. In a very large family consisting of 240 living members based in Iowa, spanning three generations, the nuclear family of 21 members were initially assessed with genome-wide scanning. A specific haplotype on chromosome 15 near its telomere co-segregated in the family (Fig. 6), meaning that it was found in all affected, and not found in those unaffected. In this region of over 60 genes, MEF2A was considered a candidate because of its known effect in the development of heart muscle, albeit without knowledge of coronary arterial impact. In assessing the proband's sequence of MEF2A, a 21 bp deletion was found in exon 11, the stop codon or last coding sequence of this gene. The same 21 bp deletion was identified in all members of the family affected, and not present in family members unaffected of 200 controls from our large genetic repository of individuals who have undergone coronary angiography. The functional genomics of this deletion mutation, leading to a seven amino-acid "in frame" truncation, were carried out to show the inability for this transcription factor to localize in the nucleus, the presence of MEF2A in the endothelial layer of the artery, and the shutdown of transcription using a standard assay technique. Further work is necessary to account more precisely for the specific pathogenesis of MI with an MEF2A deletion.

The work on MEF2A was recently extended by the discovery that mutations in this gene may account for MIs beyond the original family with an autosomal-dominant inheritance pattern. In over 400 MI cases and controls, MEF2A mutations were found in 1.9% of cases and no controls (37). Rather than involving exon 11, the mutations were point ones leading to a change of amino acids from exon 7. These point mutations (Fig. 7) also were responsible for significant limitation of transcription activity of MEF2A. This extension of MEF2A significance from a single family to the broader population demonstrates the value of such a rich pedigree approach in identifying novel MI genes.

Recently, Weng et al. (38) published findings challenging the MEF2A association with MI. These investigators found individuals with the same exon 11 21 bp deletion who did not have CAD (38). Furthermore, a missense MEF2A mutation (S360L), found only in a patient with CAD, was deemed benign using computational analysis (38). An accompanying editorial by Altshuler and Hirschorn (39) claimed that the "genetic evidence available to date does not demonstrate that these mutations play a causal role in CAD and/or MI in humans." Our group responded to the concern about lack of replication, pointing out that the S360L missense mutation, in the critical transcriptional activation domain, may well be a disease-causing variant (40). The lack of coronary disease among the four individuals with the 21 bp deletion may well have been accounted for by inadequate phenotyping. These patients did not undergo coronary angiography, and one proband had a transient ischemic attack suggesting cerebrovascular disease. The possibility of incomplete penetrance was raised (38,40).
Since these reports, a new study from a Spanish team of investigators has confirmed marked susceptibility for CAD with a missense MEF2A mutation in a large cohort of over 1,500 cases and controls (41). Clearly, more investigation will be required to determine the precise relationship of MEF2A variants and CAD. Notwithstanding these concerns, the proof for MEF2A as a disease-causing gene stems from its co-segregation in a large family, functional data demonstrating the deleterious effect on the function of the MEF2A protein, identification of multiple missense mutations in other patients and families, and replication by independent investigators.

**GENOME-WIDE SNP ASSOCIATION**

An exciting advance in the field has been made by Hinds et al. (42) from Perlegen Science with the genotyping of 1,586,383 SNPs in 71 individuals of diverse ethnic and racial origin. This is the most extensive and informative genotyping that has ever been accomplished beyond the initial complete sequence of a couple of individuals. It has already provided critical insight into human genomic variation, with the demonstration of approximately 30% “private” SNPs relegated to a particular ethnicity or race (Asian, African, or European-American) and 70% “cosmopolitan” referring to sharing of the SNPs across the diverse popula-
Along with this breakthrough, the International Haplotype Map (known as HapMap) (43) is proceeding to unravel nearly three million SNPs in 270 individuals or approximately one SNP per 1,000 bases across the genome (44). This will get us substantially closer to a fully informative set of approximately 500,000 haplotype-tag SNPs. Such SNPs are the index or window to a haplotype block, providing considerable information beyond an SNP that is not identifying a much more extensive pattern or block in the genome. These major advances will undoubtedly facilitate the objectives of determining MI genes in the future. Rather than finding a locus with microsatellites, the whole genome SNP association will permit the actual identity of SNPs and haplotypes that are associated with MI. Eventually, the ability to sequence the entire genome of individuals may even further accelerate the identity of key genes that determine susceptibility for MI. The implications of this rapid and revolutionary jump in high throughput genotyping comes at a time when complex trait validation has gained considerable momentum and the initial genes that are linked to MI have begun to be unraveled. As opposed to long QT syndrome and hypertrophic cardiomyopathy, in which more than 75% of the genes for these Mendelian traits have been characterized, we are clearly at a much less formative point with <10% of MI genes having been thus far determined.

**Figure 6 Continued.** (C) Δ21bp results in a deletion of seven amino acids of MEF2A (ΔQ440P441P442Q443P444Q445P446 or Δ7aa). (D) Functional characterization of WT and Δ7aa MEF2A proteins by transcriptional activation assays. The effect of Δ7aa on transcription activation activity of MEF2A was analyzed in the presence or absence of the zinc-finger transcription factor GATA-1 using the ANF promoter. Transcriptional activity is shown as relative luciferase activity on the y axis. The transcriptional activity of the reporter gene only (vector) was set arbitrarily to 1. Western blot analysis with anti-MEF2A rabbit polyclonal antiserum showed that both WT and mutant MEF2A (Δ7aa) were successfully expressed in transfected HeLa cells (shown in the box; C, vector only as negative control; the MEF2A antibody detected two bands as previously reported). The data shown were from two independent experiments in triplicate, and are expressed as mean ± S.E. WT = wild type MEF2A; WT/Δ7aa = coexpression of both wild type and mutant MEF2As; Δ7aa = the 7 amino acid deletion of MEF2A. (With permission from Wang et al. [36]).

**IMPLICATIONS FOR THE FUTURE**

Back in 1985, Brown and Goldstein were awarded the Nobel Prize for their discovery of an autosomal-dominant inheritance of the low-density lipoprotein receptor accounting for familial hypercholesterolemia, and leading to the discovery of statins. In 1996, these investigators published a controversial editorial in Science entitled “Heart Attacks: Gone With the Century? (45). They wrote “... better definition of genetic susceptibility factors—may well end coronary disease as a major public health problem early in the next century.” It appears unlikely that we will be rid of CAD in the imminent years ahead, owing not only to its pervasiveness but the ongoing diabesity epidemic. On the other hand, MI is much more heritable and appears to be a more ideal, restrictive phenotype as a complex trait that is making it an attractive discovery target for determination of its genetic basis. If most individuals who carry MI-susceptible genes can be recognized at an early age, prevention, lifestyle factors, and personalized drug approaches could be implemented to markedly reduce the toll of MI in the future. Indeed, it does seem possible that through genomic definition of MI we will see a time in the years ahead that MI is an unusual event albeit not “gone with the century.” Yet, limiting the MI events has the capacity to transform CAD to a far less life or
death and disabling condition. With the rapid and explosive information on the genomics root cause of MI coming forward, we can look forward to such advances in the next decade. Undoubtedly, a not-so-distant future Simon Dack lecture will be dedicated to the overwhelming triumph of reduction in the incidence and toll of MI.

Reprint requests and correspondence: Dr. Eric J. Topol, Desk F25, 9500 Euclid Avenue, Cleveland Clinic Foundation, Cleveland, Ohio 44195. E-mail: topole@ccf.org.

REFERENCES


