PROCEEDINGS OF THE
SIXTH INTERNATIONAL CONGRESS
ON TWIN STUDIES

Rome: 28-31 August 1989

1. Twin Registers and Clinical Studies
Risk Factor Variability and Coronary Heart Disease

K. Berg

Institute of Medical Genetics, University of Oslo; and Department of Medical Genetics, Ullevål Hospital, Oslo, Norway

Abstract. Present attempts to identify genes contributing to coronary heart disease (CHD) risk focus on "candidate genes". With respect to CHD this could be any gene whose protein product is directly or indirectly involved in atherogenesis, thrombogenesis or thrombolysis/fibrinolysis. Genes that exhibit associations with absolute risk factor levels may be referred to as "level genes" to distinguish them from "variability genes", which are genes involved in establishing the framework within which environmental influences may cause risk factor variation. In a series of persons recruited from the Norwegian Twin Panel, confirmatory evidence for level gene effect with respect to apolipoprotein B (apoB) concentration was found with an XbaI polymorphism in DNA at the apoB locus corresponding to residue 2,488 in the mature protein. Evidence for variability gene effect with respect to apoB as well as body mass index emerged with DNA variants in the 3' part of the apoB gene. Level gene effect with respect to apolipoprotein A-I (apoA-I) and high density lipoprotein (HDL) cholesterol as well as apparent variability gene effect with respect to total and LDL cholesterol were detected with a DNA polymorphism at the cholesteryl ester transfer protein (CETP) locus. The first example of interaction between normal genes in determining risk factor level was uncovered in analysis of the apolipoprotein E (apoE) polymorphism and a restriction fragment length polymorphism (RFLP) at the low density lipoprotein receptor (LDLR) locus. An LDLR gene identified by presence of a PvulI restriction site eliminated completely the well known effect of the apoE4 allele on cholesterol level. Finally, in families where high Lp(a) lipoprotein level (a well established risk factor for CHD) segregated as a Mendelian trait, very close linkage with an RFLP at the plasminogen locus was established and DNA variation at the LPA locus reflecting varying numbers of a structure homologous to the "kringle IV" region of plasminogen was uncovered.

Key words: Hyperlipidemia, Body mass index, Lipoproteins, Atherosclerosis, Coronary heart disease, Candidate genes, DNA polymorphisms
CANDIDATE GENES

Life style and dietary habits as well as genes contribute to the population variation in risk factors for coronary heart disease (CHD). Present attempts to identify genes contributing to CHD risk focus on “candidate genes”. With respect to CHD, there are several candidate genes, such as any gene whose protein product is:
- involved in lipoprotein structure, lipoprotein metabolism or lipid metabolism;
- involved in thrombogenesis, thrombolysis or fibrinolysis;
- involved in regulation of blood flow in coronary arteries;
- involved in regulation of blood pressure;
- involved in reverse cholesterol transport;
- present in atherosclerotic lesions;
- involved in the regulation of growth of atherosclerotic lesions;
- involved in the early development of coronary arteries.

Although the term had not been coined at that time, the candidate gene approach was used in the 1970s when the association between Lp(a) lipoprotein [1] and CHD was detected [2], and when the associations between the low density lipoprotein (LDL) allotypic Ag(x) variation and lipids were uncovered [3]. In this paper, some of our recent studies on candidate genes with respect to CHD risk factors will be briefly reviewed. Most of the studies were done on people recruited from the Norwegian Twin Panel [5].

LEVEL GENES AND VARIABILITY GENES

Marker genes exhibiting direct association with absolute risk factor levels may conveniently be referred to as “level genes”. However, genes may be of importance for CHD risk factors not only by contributing to absolute risk factor levels, but also by (partly) determining the framework within which environmental, life style or nutritional factors can cause risk factor variation [5,8-10]. Such genes may be referred to as “variability genes” to distinguish them from “level genes”. The method we have developed [4,5,24] to detect variability gene effects employs monozygotic (MZ) twins. Since MZ twins have identical genes, any difference between the two members of a pair in a quantitative biological parameter must be caused by environmental, life style or nutritional factors. A gene affecting variability should therefore be detectable by comparing the within-pair difference in a quantitative parameter between MZ pairs who have and MZ pairs who lack the gene under study. If a variability gene has a permissive effect, greater within-pair difference should be observed in MZ pairs possessing than in MZ pairs lacking that gene, whereas the opposite would be true for a gene with a restrictive effect. The method offers a unique possibility to analyze gene-environment interactions.

The rationale for applying this method to CHD risk factors is that a person’s response to atherogenic stimuli may be as relevant to his or her risk to develop atherosclerosis, as absolute risk factor levels. It has been known for many years...
that strain differences exist in various animal species, with respect to response to fat intake, and terms such as hyporesponders and hyperresponders have been introduced to describe phenotypes which are almost certainly genetically determined. Katan and his coworkers [15,18,19] have recently demonstrated that hypo- and hyperresponders to dietary cholesterol exist in man and that these traits persist at least over several years. They have calculated that if the mean response to a certain dietary cholesterol load is 0.58 mmol/l, then the 16% of subjects who are least susceptible will experience a cholesterol rise of only 0.29 mmol/l or less whereas the 16% of subjects who are most susceptible to diet will have a rise of 0.87 mmol/l or more. Thus, important individual differences with respect to response to dietary cholesterol are present in man and it seems plausible that hypo- and hyperresponders exist in the same way as they do in animals. Presumably, genetic determinants of these traits exist also in man. The newly demonstrated difference between individuals in response to fat intake makes it particularly desirable to introduce more dynamic approaches to the genetics of CHD risk factors than analysis of marker gene associations with absolute lipid or apolipoprotein levels.

EFFECTS OF GENES AT THE APOLIPOPROTEIN B LOCUS ON APOLIPOPROTEIN B LEVEL AND VARIABILITY

The Ag(x) allotypic LDL variation has been shown to reside in apolipoprotein B (apoB) rather than in other parts of the LDL particle [7]. Associations between absence of the Ag(x) antigen and high levels of cholesterol and triglycerides were clearly demonstrated in the mid-1970s [3].

The Ag(x) polymorphism exhibits strong association with an Xbal restriction site polymorphism in DNA at the apoB locus corresponding to residue 2,488 in the mature protein. Law et al [22] found significantly lower cholesterol and triglyceride levels in people who lacked this polymorphic restriction site than in other people and we have confirmed the association between the Xbal polymorphism and lipids [6]. The association between lipid levels and each of the two polymorphisms are in excellent agreement with the association between the polymorphisms themselves. Thus, genes at the apoB locus appear to have level gene effects. Since the polymorphic Xbal site reflects a silent third base mutation, this RFLP cannot cause structural differences between apoB from different persons. Therefore, the lipid association of this DNA variation most likely reflects linkage disequilibrium with functionally important domains.

A restriction fragment length polymorphism (RFLP) detectable with the enzyme EcoRI in the coding sequence of the apoB gene corresponding to amino acid 4,154 in the mature protein appears to express variability gene effect. The with-in-pair difference in apoB level was in homozygotes for absence of the restriction site almost twice as high as in homozygotes for presence of the site [11]. This difference was statistically significant despite the small number of homozygotes for absence of the restriction site (Table 1).
Table 1 - Mean within-pair difference in age- and sex-adjusted apolipoprotein B level ($\Delta$ apoB) in healthy Norwegian MZ twin pairs homozygous in an EcoRI restriction site polymorphism (2 alleles) at the apoB locus corresponding to amino acid 4,154 in mature apoB, or in a polymorphism (3 alleles scored) reflecting varying numbers of a 30-base pair repeat in the 3' flanking area of the apoB gene [extracted from reference 11]

<table>
<thead>
<tr>
<th>Homozygous genotype</th>
<th>Polymorphism detectable with EcoRI</th>
<th>Polymorphism in hypervariable 3' flanking region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of pairs</td>
<td>$\Delta$ apoB (mg/dl)</td>
</tr>
<tr>
<td>1-1</td>
<td>104</td>
<td>8.7*</td>
</tr>
<tr>
<td>2-2</td>
<td>5</td>
<td>16.8*</td>
</tr>
<tr>
<td>3-3</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

* $t = 2.47$, $P = 0.02$; ** $t = 2.53$, $P = 0.01$.

In the test system used, we also scored the MZ twins with respect to three "genes" in a polymorphism arising from varying numbers of a 30-base pair repeat, outside the 3' part of the apoB gene. The differences between 2-2 and 3-3 homozygous pairs in this polymorphism with respect to within-pair variation in apoB level was significant (Table 1). The EcoRI RFLP is in the 3' part of the gene and the hypervariable area is in the 3' flanking area [7]. The two DNA polymorphisms exhibit very strong allelic association. The apparent variability gene effect of each RFLP is in agreement with the strong allelic association between the two RFLPs. There is no reason to believe that any of the two RFLPs themselves has variability gene effect, but the data do suggest that there are areas in the 3' part of the apoB gene that affect apoB variability.

POSSIBLE EFFECT OF GENES AT THE APOLIPROPROTEIN B LOCUS ON BODY MASS INDEX

The genes contributing to predisposition or resistance to obesity in man are unknown. However, Rajput-Williams et al [28] found that certain haplotypes reflecting closely linked RFLPs at the apoB locus were associated with obesity, and LDL allotypes reflecting genetic apoB variation appear to be associated with leanness or fatness in swine (dr. Jan Rapacz, personal communication).

We have found suggestive evidence for variability gene effect with respect to body mass index at the apoB locus [11]. In the allotypic Ag(x) system of LDL, sex- and age-adjusted within-pair difference in body mass index was significantly lower in MZ twin pairs of phenotype Ag(x+) than in MZ twin pairs of phenotype Ag(x−) (Table 2). In the EcoRI polymorphism corresponding to amino acid 4,154 in mature apoB, we found that MZ twin pairs who were homozygous for absence of the restriction site had a higher mean within-pair difference in body mass index
than did pairs who were homozygous for presence of the site (Table 3). Variability gene effect seems to be expressed also by the polymorphism of the hypervariable region in the 3' flanking area of the apoB gene. The findings with the two DNA polymorphisms indicate that domains with variability gene effect on body mass index are present in the 3' part of the apoB gene. Thus, this area of the gene may influence variability of quantitative lipoprotein parameters as well as body mass index. The exact site of the Ag(x) variation in the apoB gene is unknown.

Table 2 - Sex- and age-adjusted within-pair difference in body mass index (Δ body mass index) in MZ Norwegian twin pairs according to Ag(x) phenotype [extracted from reference 11]

<table>
<thead>
<tr>
<th>Ag(x) phenotype</th>
<th>No. of pairs</th>
<th>Δ body mass index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag(x-)</td>
<td>92</td>
<td>1.89</td>
</tr>
<tr>
<td>Ag(x+)</td>
<td>60</td>
<td>1.35</td>
</tr>
</tbody>
</table>

\[t = 2.14; P = 0.03.\]

Table 3 - Sex- and age-adjusted within-pair difference in body mass index (Δ body mass index) in MZ Norwegian twin pairs according to genotype in an EcoRI site polymorphism at the apolipoprotein B locus (apoB genotype), corresponding to amino acid 4,154 in the mature protein [extracted from reference 11]

<table>
<thead>
<tr>
<th>ApoB genotype</th>
<th>No. of pairs</th>
<th>Δ body mass index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>101</td>
<td>1.63</td>
</tr>
<tr>
<td>1-2</td>
<td>46</td>
<td>1.69</td>
</tr>
<tr>
<td>2-2</td>
<td>5</td>
<td>2.98</td>
</tr>
</tbody>
</table>

\[t = 1.98, P = 0.05\] for comparison between the two categories of homozygotes.

**RISK FACTORS AND GENETIC VARIATION AT THE CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) LOCUS**

Reverse cholesterol transport is the least understood part of lipid metabolism. Cholesteryl ester transfer protein (CETP) is believed to be the most important component in this lipid transport system. DNA representing the CETP gene has been cloned and RFLPs have been uncovered in the gene. Studying a polymorphism detectable with the restriction enzyme TaqI, we found a striking association between genotypes in this CETP polymorphism and apolipoprotein A-I (apoA-I) as...
well as high density lipoprotein (HDL) cholesterol levels, absence of the restriction site being associated with high levels [20]. The difference was significant also with log transformed values [20]. This strong association suggests that genetic variation at the CETP locus may be important for levels of CHD risk factors or protective factors (Table 4). The association appears to be limited to non-smokers. If confirmed, this would be another example of interaction between genotype and lifestyle factors.

Table 4 - Mean sex- and age-adjusted apolipoprotein A-I (apoA-I) and HDL cholesterol levels in people of different genotypes in the TaqI B polymorphism in DNA at the CETP locus [extracted from reference 20]

<table>
<thead>
<tr>
<th>Genotype</th>
<th>No. of persons</th>
<th>ApoA-I (mg/dl)</th>
<th>HDL cholesterol (mmol/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>46</td>
<td>139²</td>
<td>1.27</td>
</tr>
<tr>
<td>1-2</td>
<td>76</td>
<td>147ᵇ</td>
<td>1.39</td>
</tr>
<tr>
<td>2-2</td>
<td>24</td>
<td>158</td>
<td>1.48ᶜ</td>
</tr>
<tr>
<td>Total</td>
<td>146</td>
<td>146</td>
<td>1.37</td>
</tr>
</tbody>
</table>

² Significance of difference from 2-2 group: P = 0.005.
ᵇ Significance of difference from 2-2 group: P = 0.05.
ᶜ Significance of difference from 1-1 group: P = 0.03.

We also have evidence for variability gene effect at the CETP locus on total and LDL cholesterol [14]. The difference between the two categories of homozygotes was significant also with log transformed values. This may suggest that genetic variants at the CETP locus affect the body’s response to atherogenic stimuli. At present it is not known how the same CETP allele (detected because of absence of the TaqI B restriction site) can cause increased apoA-I and HDL cholesterol levels and at the same time exert a restrictive effect on LDL and total cholesterol variability. The results of our studies of CETP variants indicate that components involved in reverse cholesterol transport may play important roles in determining risk factor levels and variability.

GENE-GENE INTERACTION AND RISK FACTORS

Studying an RFLP at the low density lipoprotein receptor (LDLR) locus, we have found that normal alleles at that locus are associated with cholesterol levels, homozygotes for absence of a PvuII restriction site having a significantly higher mean cholesterol level than heterozygotes (the number of homozygotes for presence of the restriction site was too small to permit meaningful comparisons) [26]. This confirms previous suggestions from the study of LDLR function parameters which indicated that normal LDLR activity is partly genetically determined, probably by a small number of normal alleles [23].
As part of our search for gene-gene interactions, we have analyzed the effect of the apolipoprotein E4 (apoE4) allele in one random member of each of 156 twin pairs which have been scored with respect to the LDLR polymorphism [27]. In the total series, the effect of the E4 allele on age- and sex-adjusted total and LDL cholesterol was as expected, people possessing the allele having higher values than those lacking it. The two groups were further subdivided according to genotype in the LDLR polymorphism. The effect of the apoE4 allele could then be seen only in people who were homozygous for absence (genotype A2A2) of the PvuII restriction site at the LDLR locus (Tables 5-6). Thus, a gene expressed as presence of the restriction site appears to eliminate the effect of the apoE4 allele [27]. The effect of LDLR genes on total and LDL cholesterol was significant only in people possessing the apoE4 allele. A number of analyses were conducted on this sample, and the possibility of a chance occurrence had to be considered. Therefore, a second series consisting of 239 unrelated persons was examined. The effect of LDLR alleles on total and LDL cholesterol as well as the interaction between LDLR and apoE genes were confirmed [Pedersen & Berg, in preparation]. This appears to be the first example of interaction between normal genes in determining the level of a CHD risk factor.

Table 5 - Interaction between apolipoprotein E (apoE) alleles and normal alleles at the low density lipoprotein receptor (LDLR) locus expressed as a PvuII restriction site polymorphism, in determining age- and sex-adjusted total cholesterol [extracted from reference 27]

<table>
<thead>
<tr>
<th>ApoE4 allele</th>
<th>Mean total cholesterol (mmol/l) in people with LDLR genotype</th>
<th>In total series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>A2A2 7.06* (N = 31) A1A1 or A1A2 5.87 (N = 15) In total series 6.67 (N = 46)</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>6.08* (N = 69) A1A1 or A1A2 5.84 (N = 41) In total series 5.99 (N = 110)</td>
<td></td>
</tr>
</tbody>
</table>

* t = 3.90, P < 0.001.

Table 6 - Interaction between apolipoprotein E (apoE) alleles and normal alleles at the low density lipoprotein receptor (LDLR) locus expressed as a PvuII restriction site polymorphism, in determining age- and sex-adjusted LDL cholesterol [extracted from reference 27]

<table>
<thead>
<tr>
<th>ApoE4 allele</th>
<th>Mean LDL cholesterol (mmol/l) in people with LDLR genotype</th>
<th>In total series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>A2A2 5.07* (N = 31) A1A1 or A1A2 4.05 (N = 15) In total series 4.74 (N = 46)</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>4.16* (N = 69) A1A1 or A1A2 3.98 (N = 41) In total series 4.09 (N = 110)</td>
<td></td>
</tr>
</tbody>
</table>

* t = 4.03, P < 0.001.
RECENT PROGRESS IN THE STUDY OF LP(A) LIPOPROTEIN

Lp(a) lipoprotein [1] is a well established risk factor for coronary heart disease [2,16,29]. Total sequencing, by research groups in San Francisco and Chicago, of cDNA representing the gene for the polypeptide carrying the Lp(a) antigen (the Lp(a) polypeptide chain) revealed extensive homology with plasminogen [25]. This suggests that interference with thrombolytic or fibrinolytic processes may (at least in part) explain the increased susceptibility to develop premature CHD in people with a genetically determined high level of Lp(a) lipoprotein [12].

Although the homology with plasminogen causes difficulties in studying the LPA gene, we have recently identified a quantitative DNA variation in the gene, most likely reflecting varying numbers between people, of a structure that is homologous to the “kringle IV” region of plasminogen [21].

In families where a single gene determining a high Lp(a) lipoprotein level segregates, we have shown very close genetic linkage between the gene for high level of Lp(a) lipoprotein and an RFLP at the plasminogen locus. Others have shown close linkage between Lp(a) type scored by double immunodiffusion and plasminogen [30], as well as between isoforms of the Lp(a) polypeptide chain and plasminogen [17]. Thus, it must be concluded that all three categories of Lp(a) variations scored reflect variation at a single locus: the LPA locus [12].

CONCLUDING REMARKS

The above and other progress in the study of the effect of genes on CHD risk factor levels and variability suggest that genes contributing to predisposition to early coronary heart disease will be identified [13]. It is likely that it will be an individual’s combination of level genes and variability genes that determines the total genetic risk. A broad spectrum of candidate genes should be examined but it seems plausible that the final outcome will be that analysis with respect to a relatively small number of genes will detect the greater part of the total genetic risk.

In order to utilize in an optimal way existing and future possibilities to identify individuals with increased CHD risk, a high-risk strategy based on family-oriented preventive medicine is called for.

Acknowledgments. Work in the author’s laboratory is supported by the Norwegian Council on Cardiovascular Disease, the Norwegian Research Council for Science and the Humanities, and Anders Jahres Foundation for the Promotion of Science.
REFERENCES


Correspondence: Professor Kåre Berg, Institute of Medical Genetics, University of Oslo, P.O. Box 1036 Blindern, 0315 Oslo 3, Norway.