Meta-Analysis of 4 Coronary Heart Disease Genome-Wide Linkage Studies Confirms a Susceptibility Locus on Chromosome 3q

Benedetta D. Chiodini, Cathryn M. Lewis

**Objective**—In coronary heart disease (CHD), 4 independent, genome-wide screens have now been published on Finnish, Mauritian, European, and Australian families. Results from these studies are inconclusive. We performed a meta-analysis to identify genetic regions that show evidence for susceptibility genes across studies.

**Methods and Results**—The rank-based genome-scan meta-analysis (GSMA) method was applied to the 4 CHD genome-wide linkage studies. The strongest evidence for linkage was found on chromosomes 3q26–27 (P = 0.0001) and 2q34–37 (P = 0.009). Analysis weighted by study size confirmed linkage in these regions (3q26–27, P = 0.0002; 2q34–37, P = 0.014).

**Conclusions**—The genetic regions 3q26–27 and 2q34–37 might contain susceptibility genes for CHD. Linkage to the 3q26-qter region has previously been shown in type 2 diabetes mellitus, metabolic syndrome, cholesterol concentration in LDL size fractions, and renal function in hypertensive subjects. The 2q34–37 region lies close to the type 2 diabetes NIDDM1 locus. Both of these regions harbor several candidate genes involved in the homeostasis of glucose and lipid metabolism. These results are particularly intriguing, given the growing evidence of an association between CHD risk and metabolic abnormalities, such as insulin resistance, type 2 diabetes, abdominal obesity, and dyslipidemia. (Arterioscler Thromb Vasc Biol. 2003;23:1863-1868.)

**Key Words:** meta-analysis ■ coronary heart disease ■ genome scan ■ linkage analysis ■ genetics

Coronary heart disease (CHD) includes several disorders determined by different levels of atherosclerosis, plaque instability, and coronary thrombosis: myocardial infarction, unstable angina, and various levels of asymptomatic arterial occlusions. CHD is the primary cause of mortality in developed countries, and it is predicted to become the major cause of morbidity and mortality word-wide within the next 2 decades.1

CHD is a multifactorial disease, determined by both genetic and environmental factors. Common CHD risk factors are hypertension, hyperlipidemia, obesity, type 2 diabetes (T2DM), and tobacco consumption. However, some populations are more susceptible than others to CHD, even when they share very similar environmental exposures.2 Family history of CHD is a strong and independent risk factor for myocardial infarction in both sexes and has a synergistic effect with other cardiovascular risk factors.3,4 According to the Swedish Twin Registry, death from CHD after the cotwin died before 55 years of ischemic heart disease was 3.8 in male, dizygotic twins.5 The heritability of death caused by CHD was estimated to be 0.57 and 0.38 in males and females, respectively.6 Despite strong evidence for a genetic factor contribution to CHD, only a few genes have been confirmed.

Genome-wide linkage studies are widely used to identify chromosomal regions that might contain susceptibility genes for complex diseases. For mendelian diseases, results have been easily replicated across studies. However, a complex disease might have many low- to moderate-risk genes involved, with a low probability that the same results would be found in many different genome-wide screens. Altmüller et al7 reviewed 101 whole-genome scans in 31 different complex human diseases and concluded that most studies did not show significant linkage and that results from the same disease were often inconsistent. In CHD, 4 independent, genome-wide screens have now been published, but results from these studies were inconclusive.8–11

Meta-analysis can be used to explore the lack of replication, and the combination of evidence from multiple studies might prove critical to localizing genes in common, complex human diseases.7 Meta-analysis cannot overcome complete genetic heterogeneity between studies, but it can deal with low power in individual scans. It can confirm evidence for regions highlighted in at least 1 scan or identify new regions
where weak but consistent evidence for linkage is seen across studies. The standard epidemiologic methods for meta-analyses are inappropriate, because maximum LOD scores (which assess significant linkage) might occur up to 30 cM from a susceptibility gene, so meta-analysis methods must assess linkage across a region and not merely at discrete points. Novel methods for meta-analysis of linkage studies are therefore required, eg, the genome-scan meta-analysis (GSMA) method. Under the null hypothesis of no linkage, therefore required, eg, the genome-scan meta-analysis (GSMA) method. Under the null hypothesis of no linkage, there will be randomly distributed across the genome. The GSMA assesses whether the strongest evidence for linkage across scans clusters in the same genomic region. In the GSMA, the genome is divided into bins of 30 cM. Within a study, the highest LOD score within each bin is identified, and the bins are then ranked. The autosomes and X chromosome comprise 124 bins, so the bin containing the most significant result will have rank 124, regardless of the precise LOD score obtained. Ranks for each bin are then added across studies, and the bins with the highest summed ranks indicate regions with the strongest evidence for linkage. The probability values for each bin can then be determined. The GSMA method is applicable to diverse study designs and is not restricted by different phenotypes, family structures, markers, or analysis methods across studies. The GSMA uses genome-wide LOD scores (not original genotype data), and relevant information can be extracted from published articles. We have applied the GSMA to the 4 published CHD genome screens to determine whether a degree of statistical support can be achieved for any chromosomal region.

**Methods**

The first CHD genome-wide linkage scan was performed in the genetically isolated population of Finland; 2 study samples and the pooled data set provided evidence for linkage to premature CHD on chromosome 2q21.1–22 and chromosome Xq23–26. A genome scan in 77 Mauritian families with premature CHD showed suggestive linkage on chromosomes 10q23, 16p13-pter, and 3q27. In a study of 513 European families, significant linkage to myocardial infarction was found on chromosome 14. The fourth study tested for linkage to acute coronary syndromes in 61 Australian families; this study identified 1 suggestive locus on chromosome 2q36–37 and 2 other potential loci on chromosomes 3q26–27 and 20q11–13. The design and analysis of each CHD genome scan are given in the Table. Studies show considerable diversity in clinical criteria, sample size, number of markers, and analysis methods, but these do not affect the GSMA. The studies used different analysis programs and methods, but the rank-based statistic of the GSMA makes any difference in LOD score calculation methods across studies irrelevant. The range of phenotypes will affect the genes that could be detected (see Discussion). The different numbers of markers used will not affect the GSMA, because each study had a sufficiently high density of markers, which were evenly distributed across the genome. The GSMA takes no account of the relative sample size of studies, but a weighted analysis was also performed to account for the wide range in sample sizes.

In the GSMA, autosomes and chromosome X are divided into 124 bins of ~30 cM. For each study, the bin location of each marker used was determined, and then the highest LOD score from each bin was identified. The bins were ranked in descending order (where rank 124 designates the bin with the most significant result). The ranks for each bin were then summed across studies. Bins that achieve high summed ranks therefore show most evidence for linkage, and probability values for each bin (P for unweighted analysis) were calculated from the GSMA distribution function. Data for the GSMA were extracted from the published results from graphs or tables. The Marshfield genetic maps (http://www.cshl.org/cshl-genethon-map.html) were used to determine marker locations. The Finnish study reported only markers with maximum LOD score (MLS) >1.0, for either single-point or multipoint analysis. We could therefore assign LOD scores for only 24 bins, and the remaining 100 bins were given equal ranks of (124–24)/2. Simulation showed that the GSMA distribution function provides a good fit for the CHD summed-rank data, despite the high number of tied ranks in the Finnish study, and missing data provide conservative probability values. The Mauritian and European articles presented graphs that showed the multipoint MLS results for all chromosomes. The Australian article presented results as the maximum LOD score per chromosome, and the authors contributed additional information for multipoint LOD scores across the genome. In cases where the initial genome search was followed up with additional markers or families, these data were excluded from the GSMA. Such results bias the

### Summary of Genome Searches Included in GSMA

<table>
<thead>
<tr>
<th>Study Authors (year)</th>
<th>Finnish</th>
<th>Mauritian</th>
<th>European (western)</th>
<th>Australian (white)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pajukanta et al (2000)</td>
<td>Premature CHD, age: &lt;55 years for males, &lt;65 years for females</td>
<td>Premature CHD, age: &lt;52 years</td>
<td>MI and/or premature CAD, age: &gt;=59 years</td>
<td>ACS, age: &lt;70 years (average 61.9)</td>
</tr>
<tr>
<td>Francke et al (2001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Inclusion criteria for probands and affected siblings:
- Premature CHD; age: <55 years for males, <65 years for females
- MI and/or premature CAD; age: >=59 years
- ACS, age: <70 years (average 61.9)
- MI, myocardial infarction; CAD, coronary artery disease.

<table>
<thead>
<tr>
<th>Study population</th>
<th>Finnish</th>
<th>Mauritian</th>
<th>European</th>
<th>Australian</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of families</td>
<td>156</td>
<td>77</td>
<td>513</td>
<td>61</td>
</tr>
<tr>
<td>No. of independent affected sibling pairs</td>
<td>208</td>
<td>126</td>
<td>618</td>
<td>61</td>
</tr>
<tr>
<td>Relative weighing factor</td>
<td>0.99</td>
<td>0.77</td>
<td>1.71</td>
<td>0.54</td>
</tr>
<tr>
<td>No. of autosomal markers</td>
<td>303</td>
<td>403</td>
<td>394</td>
<td>400</td>
</tr>
</tbody>
</table>

**Test Statistic output**
- MLS (single-point)
- LOD (multipoint)
- LOD (multipoint)
- MLS (multipoint)

| Analysis program | MAPMAKER/SIBS, SIBPAIR | GENEHUNTER | SOLAR | MAPMAKER/SIBS |

ACS indicates acute coronary syndrome; MI, myocardial infarction; CAD, coronary artery disease.

*Defined as >=50% stenosis of 1 to 3 coronary arteries at a young age, as verified by coronary angiography or verified myocardial infarction. †Defined as previous diagnosis of MI, history of revascularization procedures, treatment for angina pectoris. ‡Documented percutaneous transluminal coronary angioplasty (PTCA), or bypass surgery (CABG). §Defined as myocardial infarction or unstable angina.
method, because denser genotyping within a bin increases the highest LOD score achieved, even under no linkage.

The CHD studies varied widely in sample size, from 618 affected sibling pairs in the European study to 61 in the Australian study. Larger studies will have greater power to detect true linkage effects and should therefore have a greater contribution to the meta-analysis. This can be accounted for in the GSMA by a weighting function, so that the ranks reflect the relative contribution of each study. Weights were calculated from the square root of the number of affected sib pairs for each study and then scaled to give an average weight of 1.0. The optimal weighting function is not clear, but simulations have shown good performance from this function. The weights for the CHD scans were 1.71, 0.99, 0.77, and 0.54 for the European, Finnish, Mauritian, and Australian studies, respectively. Ranks for each study were multiplied by the relevant weights and then added to obtain a weighted, summed rank; probability values (Pw) for the weighted analysis were obtained by simulation.

When a CHD susceptibility locus lies close to a bin boundary, information for linkage might be "diluted" between adjacent bins. To determine whether the placement of bin boundaries was having a critical effect on results, adjacent bins were combined into 60-cM bins for the 2 possible combinations: bins 1+2, 3+4, etc, and then 2+3, 4+5, etc. The GSMA was then reapplied with 62 and 39 bins, respectively. Furthermore, to explore the effect of any single study on the overall unweighted results, each scan in turn was omitted, and the GSMA was applied to the remaining 3 studies.

Results

The GSMA results are presented as the summed ranks in each bin for unweighted and weighted analyses (Figures 1 and 2, respectively). A single point is plotted for the summed rank of each bin across the genome. Approximate 90%, 95%, and 99% confidence limits were calculated from the probability distribution (unweighted analysis) or by simulation (weighted analysis). Points above the 95% confidence limit have a probability <5%
of occurring within any bin. The notation “bin 3.7” refers to the seventh bin on chromosome 3 (Figure 3).

In the unweighted analysis (Figure 1), 3 bins obtained a summed rank >410, the value that would be exceeded in 1% of bins under the null hypothesis of no linkage. This result indicates that some of the regions above the 1% threshold might contain true susceptibility loci.

The 3q26–28 chromosomal region (bin 3.7) was the most significant result ($P_U < 0.0001$), followed by chromosome 2q34–37 (bin 2.9, $P_U = 0.009$). This bin was ranked highest only in the Finnish study, but interestingly, the adjacent distal bin (2.10) achieved the highest rank in the Australian study.

The adjacent bin 3.8 also obtained a summed rank exceeding the 1% threshold ($P_U = 0.008$). Significant results in 2 flanking bins are common, because linkage peaks in a complex trait might extend across 2 bins. Chromosome 14q24–32 (bin 14.3) also gave a significant summed rank ($P_U = 0.025$). The unweighted GSMA based on 60-cM bins confirmed the results for chromosomes 3, 2, and 14 and identified no novel regions. The results from analyzing 3 of the 4 studies also confirmed the findings on chromosome 3 and 2, with no additional bins achieving $P < 0.01$.

When the ranks were weighted by study size, the distal region of chromosome 3q remained the most significant finding (bin 3.7, $P_W = 0.0002$; bin 3.8, $P_W = 0.011$). Chromosome 2q34–37 was also significant (bin 2.9, $P_W = 0.014$). Chromosome 14q24–32 was no longer significant (bin 14.3, $P_W = 0.056$), but evidence for linkage increased in chromosomal region 8q22–24 (bin 8.5, $P_W = 0.049$). In the weighted analysis of 60-cM bins, the chromosome 3q region showed the strongest evidence for linkage, with other significant results at bins 8.4 to 8.5 ($P_W = 0.05$) and at bins 14.3 to 14.4 ($P_W = 0.01$).

**Discussion**

We performed a meta-analysis of the 4 published genome screens for CHD conducted in Finnish, Mauritian, European, and Australian families. The major problem in developing a meta-analysis technique for genome scans is the diversity of studies. The 4 CHD genome scans differed in population sample size, statistical analysis methods, and phenotype definition. In addition, the method for reporting results was not consistent between the 4 studies, with only partial results available for the Finnish study. The missing data reduced the overall power of the study to locate genes. When the Finnish study was omitted from the analysis, the regions on chromosomes 2 and 3 remained significant, and no new genetic regions were identified. The major strength of the GSMA is its application to a diversity of study designs. With regard to the phenotype investigated, Finnish and Mauritian sib pairs were selected for premature CHD, whereas in the European scan, all probands were affected with myocardial infarction; Australian sib pairs were older and affected with acute coronary syndromes. Nevertheless, all of these different traits share a common pathogenic origin, ie, the atherosclerotic process, in which both genes and environment are likely to be involved. Therefore, it is likely that although different forms of CHD might involve unique susceptibility genes, common genes for atherosclerosis will exist, and our study will have high power to identify them. The differences in the populations investigated in terms of clinical phenotype and ethnicity could in part explain the lack of consistent results across the CHD genome-wide linkage studies. Clearly, the GSMA will detect only those genes that are relevant across populations. Most of the studies are small and will have low power to detect linkage to each of the multiple susceptibility genes for CHD. This is a particular problem, considering the low sibling recurrence risk for CHD ($\lambda_s = 3.8$).

Application of the GSMA to the 4 CHD scans showed strong evidence of linkage to chromosome 3q26–27 (bin 3.7) in the unweighted ($P = 0.0001$), weighted ($P = 0.0002$), and combined analyses after removing 1 study at a time. Notably, bin 3.7 did not rank highest in any of the 4 genome-wide screens, achieving only weak evidence for linkage in each of
the 4 studies. The gene in this region might be involved in the pathogenesis not only of CHD but also of the cardiovascular dysmetabolic syndrome, a clustering of metabolic abnormalities (dyslipidemia, insulin resistance, obesity, and high blood pressure) known to seriously increase the risk of cardiovascular mortality and morbidity.\textsuperscript{16} Chromosome 3q26–27 has already shown linkage to T2DM,\textsuperscript{17–19} the metabolic syndrome,\textsuperscript{20} cholesterol concentration in LDL size fractions,\textsuperscript{21} and renal function in hypertensive subjects.\textsuperscript{22} Other more extensive CHD genome-wide screens are in progress and are likely to be published over the next year. The GSMA can then be extended, but the current study provides strong evidence for linkage on chromosome 3, which warrants immediate follow-up.

The chromosome 3q26-qter region harbors several candidate genes (see Figure 3) for CHD and the cardiovascular dysmetabolic syndrome, which have already been tested for association and linkage in CHD. One candidate gene is solute carrier family 2 of the facilitated glucose transporter (GLUT2). GLUT2 (3q26.1–26.2) influences pancreatic sen-sorship and the temporal release of insulin in response to glucose and could play an etiologic role in non–insulin-dependent diabetes mellitus (NIDDM).\textsuperscript{2} Two case-control studies\textsuperscript{23,24} reported a significant association between GLUT2 allelic variants and NIDDM.

The catalytic polypeptide of phosphoinositide 3-kinase (PIK3CA, 3q26.3) is a key component of the insulin signaling cascade and is a logical candidate for a diabetogene. A missense variant, Met326His, in the p85 regulatory subunit, has been tested for association with both NIDDM and insulin resistance, with 1 study\textsuperscript{25} showing association with insulin resistance.

Apolipoprotein D (APOD, 3q26.2-qter) is a protein component of HDL with a putative role in the cholesterol-transport pathway. Although its role in metabolism has yet to be defined, apolipoprotein D is likely to be a multiligand, multifunctional transporter. The APOD gene harbors several polymorphisms, which might be associated with lipoprotein metabolism.\textsuperscript{26} NIDDM,\textsuperscript{27} obesity, and cardiovascular risk.

Adiponectin (APM1, 3q27) is an adipocyte-secreted protein that regulates energy homeostasis and glucose and lipid metabolism, modulates endothelial function, and has an inhibitory effect on vascular smooth muscle cell proliferation. Adiponectin could be a link between obesity and related atherosclerosis and might have anti-inflammatory effects on cellular components of the vascular wall. It might also act as an antidiabetic hormone, by directly regulating glucose metabolism and insulin sensitivity in vitro and in vivo. Many studies have shown an association between APM1 allelic variants and adiponectin levels,\textsuperscript{28–31} NIDDM,\textsuperscript{28,30–33} obesity,\textsuperscript{29,34,35} insulin resistance,\textsuperscript{29,30,34,35} and CHD.\textsuperscript{32,33} The somato-tostatin gene (SST) and the protein phosphatase-1 regulatory (inhibitor) subunit 2 gene (PPP1R2) are 2 further potential candidate genes situated in this region.

The second strongest evidence for linkage occurred on chromosome 2q34–37 (bin 2.9; \(P_e = 0.014\)). This region lies close to the locus NIDDM1, which has been linked to T2DM in Mexican-Americans\textsuperscript{46} and Chinese Huns.\textsuperscript{37} Chromosome 2q36–37 harbors the insulin receptor substrate-1 (IRS-1) and the calpain-10 (CAPN-10) genes, which are promising candidate genes for both CHD\textsuperscript{38} and metabolic abnormalities.\textsuperscript{39–41}

In conclusion, the 3q26-qter showed the strongest evidence for linkage in a meta-analysis of the 4 CHD genome-wide screens. This result is intriguing, given the existence of several candidate genes involved in the homeostasis of glucose and lipid metabolism and growing evidences of the crucial association between CHD risk and metabolic abnormalities, such as insulin resistance, T2DM, abdominal obe-sity, and dyslipidemia. Further studies of this region should be undertaken.

Acknowledgments

We thank the original authors of the genome-wide linkage scans in CHD, particularly Stephen Harrap and Philippe Froguel, who contributed unpublished, genome-wide LOD scores, and Sheila Fisher, Jonathan Berg, and Maria Grazia Franzosi for valuable discussions of CHD and the GSMA.

References


Meta-Analysis of 4 Coronary Heart Disease Genome-Wide Linkage Studies Confirms a Susceptibility Locus on Chromosome 3q
Benedetta D. Chiodini and Cathryn M. Lewis

Arterioscler Thromb Vasc Biol. 2003;23:1863-1868; originally published online August 28, 2003;
doi: 10.1161/01.ATV.0000093281.10213.DB
Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2003 American Heart Association, Inc. All rights reserved.
Print ISSN: 1079-5642. Online ISSN: 1524-4636

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://atvb.ahajournals.org/content/23/10/1863

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Arteriosclerosis, Thrombosis, and Vascular Biology can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Arteriosclerosis, Thrombosis, and Vascular Biology is online at:
http://atvb.ahajournals.org//subscriptions/