Letters to the Editor

Am. J. Hum. Genet. 75:151–152, 2004

On the Significance of Linkage Studies of Complex Traits

To the Editor:

We read with interest the recent article by Wang et al. (2004) reporting linkage of premature myocardial infarction (MI) to a locus on 1p34-36. The authors have recruited a large sample of families with premature coronary artery disease (CAD), as detected by catheterization, revascularization, or MI. Such large-scale approaches will be crucial to the identification of genetic variation underlying complex traits, including atherosclerotic CAD and MI, the leading killer of men and women in the Western world. We commend the authors for undertaking such an important study.

Although their article reports the nominal LOD score of 11.68 for linkage of premature MI to 1p34-36 and a corrected pointwise *P* value of .00011, we note that the genomewide significance of the linkage statistic is not clear. We have some methodological concerns regarding the initial emphasis on results presented in their table 2 and figure 1:

1. The finding of 11 independent $-\log_{10}P$ values >3.13 in the multipoint linkage approach (regions identified in single-point and multipoint W2 analyses overlap significantly) raises questions regarding the assertion that such a threshold corresponds to a LOD score >2.2, as proposed by Lander and Kruglyak (1995). A LOD score >2.2 should occur once in a maximally informative genomewide scan, under the null hypothesis that no genetic linkage is present. It seems unlikely that 11 true genetic loci influencing a complex phenotype would be detected in a single study. It is more likely that the asymptotic *P* value statistic generated by the authors' modified Haseman-Elston regression model is inflated.

2. The marked attenuation of the multipoint *P* value of $<10^{-12}$ to 10^{-4} on pointwise permutation testing at the 1p34-36 locus suggests that the nominal asymptotic *P* values are inflated. It is possible that the method of linkage analysis may have inflated the *P* value estimates. In particular, the treatment of the dichotomous MI phenotype as a continuous variable may not be appropriate. The assumption of equal variances required for a quantitative trait may not be valid for different numbers of affected and unaffected individuals in each family.

3. The authors' attempt to correct the pointwise empirical *P* value estimates for the number of markers tested is quite important for establishing the significance of their findings. However, the attempt may be inadequate to account for the testing of multiple markers. The authors refer to the simulation analyses by Wiltshire et al. (2002), which explored the influence of experimental deviations from the Lander-Kruglyak assumptions of completely informative linkage analyses. We are uncertain whether the empirical genomewide *P* value estimates derived from the Wang et al. (2004) data correspond to the same nominal LOD-score thresholds identified in the Wiltshire et al. (2002) study using simulated data. Permutation testing of the Wang et al. (2004) data by use of the Wiltshire approach might provide greater confidence regarding the genomewide significance of the study findings, but this approach admittedly represents a significant computational burden.

4. Genomewide linkage analyses at 10 cM may not extract maximally the identity-by-descent information for the sample under study. A fine-mapping study at higher density across a region of interest may show a change in the maximum-LOD-score estimate. An increase in the LOD-score estimate with better information extraction might be reassuring, but a fall in the LOD score may signal a false-positive finding. We would encourage the authors to perform and publish the results of a higher-density map.

5. The study cohort was recruited on the basis of a composite definition of premature CAD. Was the broader CAD phenotype (including MI) the primary phenotype prespecified in the linkage analysis? The reported MI linkage analysis represents a subgroup of the subjects enrolled; the "less-restrictive" CAD phenotype was also tested and revealed no suggestive or significant linkage evidence. Could the authors clarify their original primary analysis and whether additional subgroups were analyzed? A true empirical *P* value would also account for the multiple phenotypes tested.

On review of the study, the declaration of a finding of genomewide significance may not be as strongly supported as suggested by the authors. The results of this linkage analysis do not contain much overlap with those

of similar analyses, and this certainly could result from differences in phenotype definition, environmental exposures, or study design (Pajukanta et al. 2000; Francke et al. 2001; Broeckel et al. 2002; Harrap et al. 2002; Chiodini and Lewis 2003). Replication of linkage analyses for complex cardiovascular traits has often proven challenging, and the difficulty in achieving replication for MI underscores the many difficulties in the conduct and interpretation of such linkage analyses (Altmuller et al. 2001).

Identifying genetic factors underlying linkage peaks in this and related studies of MI will require considerable expenditure of resources and should proceed on the basis of the strongest possible evidence. We encourage the systematic comparison of available and accruing linkage data across studies in various CAD phenotypes, including continued assessment of the most appropriate linkage methods.

CHRISTOPHER NEWTON-CHEH,[1,2,4] MARTIN LARSON,[3,4] SEKAR KATHIRESAN,[1,4] AND CHRISTOPHER O'DONNELL[1,4]

[1]*Cardiology Division, Massachusetts General Hospital,* [2]*Division of Preventive Medicine, Brigham and Women's Hospital, and* [3]*Section of Epidemiology and Preventive Medicine, Boston University School of Medicine, Boston; and* [4]*National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA*

References

Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M (2001) Genomewide scans of complex human diseases: true linkage is hard to find. Am J Hum Genet 69:936–950

Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ, Comuzzie AG, Blangero J, Nurnberg P, Reis A, Riegger GA, Jacob HJ, Schunkert H (2002) A comprehensive linkage analysis for myocardial infarction and its related risk factors. Nat Genet 30:210–214

Chiodini BD, Lewis CM (2003) Meta-analysis of 4 coronary heart disease genome-wide linkage studies confirms a susceptibility locus on chromosome 3q. Arterioscler Thromb Vasc Biol 23:1863–1868

Francke S, Manraj M, Lacquemant C, Lecoeur C, Lepretre F, Passa P, Hebe A, Corset L, Yan SL, Lahmidi S, Jankee S, Gunness TK, Ramjuttun US, Balgobin V, Dina C, Froguel P (2001) A genome-wide scan for coronary heart disease suggests in Indo-Mauritians a susceptibility locus on chromosome 16p13 and replicates linkage with the metabolic syndrome on 3q27. Hum Mol Genet 10:2751–2765

Harrap SB, Zammit KS, Wong ZY, Williams FM, Bahlo M, Tonkin AM, Anderson ST (2002) Genome-wide linkage analysis of the acute coronary syndrome suggests a locus on chromosome 2. Arterioscler Thromb Vasc Biol 22:874–878

Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247

Pajukanta P, Cargill M, Viitanen L, Nuotio I, Kareinen A, Perola M, Terwilliger JD, Kempas E, Daly M, Lilja H, Rioux JD, Brettin T, Viikari JS, Rönnemaa T, Laakso M, Lander ES, Peltonen L (2000) Two loci on chromosomes 2 and X for premature coronary heart disease identified in early- and late-settlement populations of Finland. Am J Hum Genet 67: 1481–1493

Wang Q, Rao S, Shen G-Q, Li L, Moliterno DJ, Newby LK, Rogers WJ, Cannata R, Zirzow E, Elston RC, Topol EJ (2004) Premature myocardial infarction novel susceptibility locus on chromosome 1P34-36 identified by genomewide linkage analysis. Am J Hum Genet 74:262–271

Wiltshire S, Cardon LR, McCarthy MI (2002) Evaluating the results of genomewide linkage scans of complex traits by locus counting. Am J Hum Genet 71:1175–1182

Address for correspondence and reprints: Dr. Christopher J. O'Donnell, National Heart, Lung, and Blood Institute's Framingham Heart Study, 73 Mount Wayte Avenue, Suite 2, Framingham, MA 01702. E-mail: codonnell@nih.gov

0002-9297/2004/7501-0020$15.00

Am. J. Hum. Genet. 75:152–154, 2004

Reply to Newton-Cheh et al.

To the Editor:

Newton-Cheh et al. (2004 [in this issue]) raise the issue of some methodological concerns in our genomewide-scan study that identified significant linkage on chromosome 1p34-36 for premature myocardial infarction (MI). We would like to systematically address their concerns. First, we do explicitly report that the genomewide significance for the chromosome 1p34-36 linkage as $P < .05$ ($P = .030$–$.038$), derived from Wiltshire et al. (2002). This point was duly emphasized in the abstract and the "Results" and "Discussion" sections (Wang et al. 2004). Second, with respect to the high number of loci with asymptotic P values (pP) that were suggestive of linkage, we performed permutation tests and reported empirical P values. As we reported, only the chromosome 1p34-36 locus fulfilled the criteria of genomewide significance. Third, MI is a dichotomous phenotype. Either patients have an MI or do not have this acute ischemic event. As reported by Altmuller et al. (2001), studies of 101 genomewide scans in 31 different diseases revealed that quantitative "intermediate" traits did not have any advantages over dichotomous traits for linkage analysis. Furthermore, several methodological investigations indicated that, in practice, treatment of ordinal (or binary) data as continuous with standard linear models for genetic mapping of categorical traits is feasible, with marginal

Table 1

Five Genomewide Scans for CAD

Study	Population	No. of Families	Mean Age (years)	Locus/Loci	Analysis Programs
Pajukanta et al. 2000	Finnish	156	<55	2q21 and Xq23	MAPMAKER/SIBS, and SIBPAIR in the ANALYZER package
Francke et al. 2001	Mauritian	99	47	16p13	Genehunter 2.1
Broeckel et al. 2002	European	513	52	14q32	SOLAR
Harrap et al. 2002	Australian	61	62	2q36	MAPMAKER/SIBS
Wang et al. 2004	American	428	44	1p34-36	SAGE

loss of statistical powers (Hackett and Weller 1995; Rao and Xu 1998; Rao and Li 2000; Rebai 2000). Fourth, we did indeed use permutation testing, and it was a significant computational burden. In addition, the results of genomewide significance fully incorporated the criteria as set forth by Wiltshire et al. (2002). Fifth, the fine-mapping study is ongoing; we agree that the 10-cM analysis will not maximally extract the identity-by-descent information. Since two markers at the chromosome 1p34-36 locus showed asymptotic multipoint P values of $<10^{-12}$, it is unlikely that further fine mapping will be informative for confirming the finding of this positive linkage. Beyond fine mapping, the most convincing evidence of linkage will be the replication of our findings in an independent population. Sixth, our primary analysis was for both premature MI and coronary artery disease (CAD). No additional subgroups were analyzed by phenotype.

Our population has the most restrictive demographic features of any cohort with CAD to date (Pajukanta et al. 2000; Francke et al. 2001; Broeckel et al. 2002; Harrap et al. 2002; Wang et al. 2004) (table 1). (Not only were our patients the youngest, probably representing the most aggressive phenotype, but the exclusion of other important risk factors, such as insulin-dependent diabetes mellitus or significant hypercholesterolemia, also helped define a cohort with the highest likelihood of genetic underpinning.) Although it is true that our findings of a chromosome 1p locus are not concordant with those of the other four genomewide scans (table 1), our population of young Americans is unique, and it is noteworthy that three of the other four studies (Pajukanta et al. 2000; Francke et al. 2001; Harrap et al. 2002) had a relatively small number of families. Furthermore, each of the studies used a different statistical method implemented in separate programs, and none of the loci recorded to date have been replicated by the other studies.

In summary, we reported a locus with high genomewide significance for linkage of premature MI through use of state-of-the-art analysis, including permutation testing. We look forward to identification of the specific gene(s) underlying this linkage peak and replication of this linkage in an independent population.

QING WANG, SHAOQI RAO, AND ERIC J. TOPOL
Department of Cardiovascular Medicine, Center for Cardiovascular Genetics, and Department of Molecular Cardiology, Lerner Research Institute, The Cleveland Clinic Foundation, Cleveland

References

Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M (2001) Genomewide scans of complex human diseases: true linkage is hard to find. Am J Hum Genet 69:936–950

Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ, Comuzzie AG, Blangero J, Nurnberg P, Reis A, Riegger GA, Jacob HJ, Schunkert H (2002) A comprehensive linkage analysis for myocardial infarction and its related risk factors. Nat Genet 30:210–214

Francke S, Manraj M, Lacquemant C, Lecoeur C, Lepretre F, Passa P, Hebe A, Corset L, Yan SL, Lahmidi S, Jankee S, Gunness TK, Ramjuttun US, Balgobin V, Dina C, Froguel P (2001) A genome-wide scan for coronary heart disease suggests in Indo-Mauritians a susceptibility locus on chromosome 16p13 and replicates linkage with the metabolic syndrome on 3q27. Hum Mol Genet 10:2751–2765

Hackett CA, Weller JI (1995) Genetic mapping of quantitative trait loci for traits with ordinal distributions. Biometrics 51: 1252–1263

Harrap SB, Zammit KS, Wong ZY, Williams FM, Bahlo M, Tonkin AM, Anderson ST (2002) Genome-wide linkage analysis of the acute coronary syndrome suggests a locus on chromosome 2. Arterioscler Thromb Vasc Biol 22:874–878

Newton-Cheh C, Larson M, Kathiresan S, O'Donnell C (2004) On the significance of linkage studies of complex traits. Am J Hum Genet 75:151–152 (in this issue)

Pajukanta P, Cargill M, Viitanen L, Nuotio I, Kareinen A, Perola M, Terwilliger JD, Kempas E, Daly M, Lilja H, Rioux JD, Brettin T, Viikari JSA, Rönnemaa T, Laakso M, Lander ES, Peltonen L (2000) Two loci on chromosomes 2 and X for premature coronary heart disease identified in early- and late-settlement populations of Finland. Am J Hum Genet 67: 1481–1493

Rao S, Li X (2000) Strategies for genetic mapping of categorical traits. Genetica 109:183–197

Rao S, Xu S (1998) Mapping quantitative trait loci for ordered categorical traits in four-way crosses. Heredity 81:214–224

Rebai A (2000) Comparison of methods for regression interval

mapping in QTL analysis with non-normal traits. Genet Res 69:69–74

Wang Q, Rao S, Shen G-Q, Li L, Moliterno DJ, Newby LK, Rogers WJ, Cannata R, Zirzow E, Elston RC, Topol EJ (2004) Premature myocardial infarction novel susceptibility locus on chromosome 1P34-36 identified by genomewide linkage analysis. Am J Hum Genet 74:262–271

Wiltshire S, Cardon LR, McCarthy MI (2002) Evaluating the results of genomewide linkage scans of complex traits by locus counting. Am J Hum Genet 71:1175–1182

Address for correspondence and reprints: Dr. Eric J. Topol, Department of Cardiovascular Medicine, F25, The Cleveland Clinic Foundation, Cleveland, OH 44195. E-mail: topole@ccf.org

0002-9297/2004/7501-0021$15.00