Editorial

On the interpretation of genetic association studies

Sekar Kathiresan\textsuperscript{a,b,c,1}, Christopher Newton-Cheh\textsuperscript{a,b,c,1}, Robert E. Gerszten\textsuperscript{a,b,*}

\textsuperscript{a} Cardiology Division and Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
\textsuperscript{b} Program in Medical and Population Genetics, Broad Institute of Harvard University and Massachusetts Institute of Technology, Cambridge, MA, USA
\textsuperscript{c} National Heart, Lung, and Blood Institute’s Framingham Heart Study, Framingham, MA, USA

This editorial refers to “Association of RANTES G-403A gene polymorphism with increased risk of coronary arteriosclerosis” by E. Simeoni et al. on page 1438 and “Association of hypo-responsive toll-like receptor 4 variants with risk of myocardial infarction” by K. Edfeldt et al. on page 1447.

Genetic association studies seek to relate variation in human DNA sequence with a disease or trait. Compared with linkage analysis, the association study design provides greater power to detect common genetic variants conferring susceptibility to complex phenotypes such as atherosclerosis and myocardial infarction (MI). In a case-control study, a common and convenient association study design, the frequency of a harmful genetic variant is expected to be greater among cases than controls (a protective variant is less frequent in cases). Though the case-control design may be simple, interpreting the results of these genetic association studies has been far less straightforward.

In this issue, the results of two case-control association studies are reported. Simeoni et al.\textsuperscript{2} investigated the relationship between polymorphisms in four candidate genes and coronary atherosclerosis detected at coronary angiography. They report that a variant in the C–C chemokine Regulated Upon Activation, Normal T-cell Expressed and Secreted (RANTES) gene, G-403A, is associated with coronary atherosclerosis. Specifically, the A allele frequency was higher in 2694 cases with coronary atherosclerosis compared to 530 controls free of coronary atherosclerosis ($p = 0.041$). Meanwhile, Edfeldt et al.\textsuperscript{3} studied the relationship between two coding polymorphisms in the Toll-like receptor 4 (TLR4) gene and MI. They report that among the 1368 men in the study, the frequency of the co-segregating 299Gly and 399Ile alleles was higher among survivors of a first MI than in controls (uncorrected $p = 0.004$).

These two studies are extensions of exciting animal work implicating inflammatory pathways in atherogenesis. It was recently demonstrated that antagonism of RANTES receptors substantially reduces plaque burden in pro-atherogenic murine models.\textsuperscript{4} Furthermore, hyperlipidaemic mice lacking myeloid differentiation protein-88 (MyD88), which transduces cell signalling events downstream of the Toll-like receptors, have a similarly marked reduction in atherosclerosis.\textsuperscript{5}

The exciting pre-clinical data notwithstanding, these important translational studies must be evaluated entirely on the merits of the genetic findings. A central problem has been the general lack of replication of genetic associations reported to date.\textsuperscript{6} A recent examination of 301 published studies of 25 different reported associations found that less than half of all the reported associations had strong evidence of replication.\textsuperscript{7} False positive reports, false negative replication studies, or true variability in association among different populations have been proposed to explain the inconsistency in the literature. However, in pooled analyses of all follow-up studies, eight of the associations yielded a significant replication of the initial report, with modest genetic effect sizes (with odds ratios generally between 1.1 and 2.0). Thus, the authors concluded that although there are abundant false positive associations in the literature, many real associations lurk among the data.
Table 1: Assessing whether a reported genetic association is real

<table>
<thead>
<tr>
<th>I. Reasons for apparent significant association</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) True positive</td>
</tr>
<tr>
<td>(i) Variant is causal</td>
</tr>
<tr>
<td>(ii) Variant is in linkage disequilibrium with causal variant</td>
</tr>
<tr>
<td>(b) False positive</td>
</tr>
<tr>
<td>(i) False positive due to multiple testing</td>
</tr>
<tr>
<td>(ii) False positive due to systematic genotyping error</td>
</tr>
<tr>
<td>(iii) False positive due to population stratification or other confounder</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>II. Reasons for lack of replication</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) False positive</td>
</tr>
<tr>
<td>(i) Original report is a false positive</td>
</tr>
<tr>
<td>(b) False negative</td>
</tr>
<tr>
<td>(i) Phenotypes differ across studies</td>
</tr>
<tr>
<td>(ii) Study populations differ in genetic or environmental background</td>
</tr>
<tr>
<td>(iii) Replication study is underpowered</td>
</tr>
</tbody>
</table>

Are these reported associations real? How is a reader to assess the validity of a reported association between a genetic variant and a phenotype? Table 1 outlines several considerations in addressing these questions. Two key issues include: (1) What are the possible reasons for an apparent association between a polymorphism and a phenotype, and (2) What are the reasons for lack of replication of a reported association? The two current reports highlight some of the issues relevant to answering these questions.

**True positive**

The RANTES-403A allele and the co-segregating TLR4 299Gly/399Ile alleles may directly cause increased coronary atherosclerosis and increased risk of MI, respectively. Alternatively, these associations may exist because of linkage disequilibrium (LD) between these variants and causal variants nearby. LD is the non-random correlation between alleles at a pair of neighbouring polymorphic sites. For example, Asp299Gly and Thr399Ile of TLR4 are in tight LD: the 299Gly allele occurs almost exclusively in the presence of 399Ile. Recent empirical studies have shown that most of the genome exists in block-like regions of strong LD. For individuals of European ancestry, these regions have on average been shown to extend by 22 kilobases (kb) of DNA. Without a comprehensive assessment of the LD patterns at the RANTES or TLR4 loci, it is difficult to be certain whether the reported variants are causal or in LD with the causal variant. Both scenarios would represent a true positive finding.

**False positive**

A false-positive association may arise for several reasons (Table 1). A failure to account for the number of hypotheses tested is a major concern in many association studies. A staggering number of genetic variants exist in the human genome. Eleven million single nucleotide polymorphisms (SNPs) with >1% minor allele frequency are estimated to exist. In addition, for each genetic variant, a number of comparisons may be tested including those of allele frequencies (for the RANTES G-403A variant, frequency of A versus G), genotype frequencies (e.g., GG versus AG versus AA), and combinations of genotypes in specific genetic models (e.g., dominant: AG + AA versus GG). Lastly, many phenotypes are available for study: coronary atherosclerosis at catheterisation, acute coronary syndrome, MI, etc. Thus, the standard p-value threshold of 0.05, which yields on average one false-positive report for every 20 independent tests, is too liberal. Simeoni et al. attempt to account for multiple hypothesis testing by performing a Bonferroni correction for the 4 variants tested in their study. It is not clear, however, whether the authors have accounted for the testing of multiple genetic models of causation or for the subphenotypes examined. Admittedly, the testing of genetic models and subphenotypes are not independent of the primary tests and thus, additional Bonferroni correction for these non-independent tests is likely to be too conservative. In order to account for multiple testing of correlated hypotheses, a number of statistical methods such as false discovery rate and permutation testing have been developed.

Systematic genotyping error can also produce a false-positive result. For example, individuals heterozygous at a variant are more difficult to assign genotypes on many genotyping platforms than are homozygotes. If loss of heterozygous genotypes is concentrated among the cases or among the controls, as can occur if controls are genotyped separately from cases, then tests of some genetic models may produce false-positive associations. The deficiency of RANTES G-403A heterozygotes among the controls in the study by Simeoni (resulting in failure of a test of Hardy-Weinberg equilibrium) may have produced just such a result. The authors were appropriately concerned about this possibility and re-designed the genotyping assay in an attempt to bring the genotypes into Hardy-Weinberg equilibrium (HWE) and, when this failed, tested for association using a test of trend that does not require HWE. Re-genotyping using a technology different from that originally employed could provide reassurance that the results are not due to systematic genotyping error.

Another potential contributor to a false-positive association is confounding by differences in genetic background, termed population stratification. Population stratification refers to differences in allele frequencies between cases and controls due to differences in ancestry between the two groups rather than association of genes with disease. A recent analysis of data from 11 case-control association studies suggested that population stratification may be a source of spurious association in well-designed studies despite matching on self-reported ancestry. Genotyping a set of unlinked markers in both cases and controls may allow investigators to test for the presence of stratification and if found to control for the effect of stratification. The power to detect such confounding is often inadequate if only a few dozen
markers are tested or if the markers are not informative of ancestry. Population stratification was not explicitly tested for in either study.

Given the concern for false-positive findings in the many association studies being performed today, replication in another sample is crucial to establish the significance of a reported association. Replication involves a new patient sample which confirms that the same allele confers risk to the same phenotype in the same direction as the original report (harmful or protective). As mentioned above, associations may not be replicated if the original association is a false positive (a true-negative finding in the replication sample).

False negative

A lack of replication may also be falsely negative, possibly due to differences in phenotypes across studies, differences in the genetic or environmental background of the populations studied or the lack of adequately powered sample sizes (Table 1). While Edfelt et al. found that the co-segregating 299Gly/399Ile variants confer a higher risk of MI, the initial report by Kiech et al.,15 associated the 299Gly variant with decreased atherosclerosis as measured by carotid intima-media thickness in 810 Italians. In 183 French cases with acute coronary syndromes compared with 216 controls, Ameziane et al.,16 reported that the 299Gly was associated with a significantly lower risk of acute coronary syndrome. Meanwhile, Yang et al.17 studied 1400 Caucasian subjects who underwent coronary angiography and related Asp299Gly genotype to the number of vessels with >50% coronary stenosis. No differences were noted in 299Gly carrier frequency in patients with 0, 1, 2, or 3 coronary arteries with >50% stenosis.

This discrepancy in direction of effect of the 299Gly allele could be due to differences in phenotype definition. Given the range of phenotypes studied including MI, carotid atherosclerosis evident on ultrasonography, acute coronary syndrome (a combination of MI and unstable angina), and coronary stenosis, it is difficult to make direct comparisons among the studies, although aspects of each phenotype are correlated. It is conceivable that an allele might decrease the risk for an atherosclerosis phenotype like carotid intima-media thickness while increasing the risk of a clinical event like MI as suggested by Edfelt et al. However, the lack of allele-direction-phenotype replication of any prior study leaves the ultimate fate of the current finding uncertain.

The ancestry of source populations may differ between studies and this difference may cause a lack of replication. Unmeasured differences in environmental and genetic background in diverse populations may influence the effect of a given variant. This does not seem to be an adequate explanation for the heterogeneity in the TLR4 variant finding. All four studies are drawn from a European population. Certainly, the hypothesis that environmental or non-TLR4 genetic differences explain the discordant findings has not been tested.

Lastly, and perhaps most importantly, most association studies are powered to detect only the strongest genetic associations. Common variants that contribute to common diseases are likely to be modest in their genetic effects. Moreover, the effect size found in an initial study is typically much greater than that ultimately found in pooled analyses of several well-powered studies, a phenomenon known as the Winner’s Curse.7,18 The failure of smaller-sized studies to confirm an association may thus be falsely negative. Thus, studies may need to have sample sizes in the 1000–10000 range to have adequate power to replicate initial findings that are truly positive.

In summary, Simeoni et al. have put forth the hypothesis that the A allele of the RANTES G-403A variant is associated with coronary atherosclerosis. In contrast to earlier studies of related but different phenotypes, Edfelt et al. report that the 299Gly allele at the TLR4 locus is associated with an increased risk of MI. The strength of both associations is modest. Additional limitations include multiple testing, the possibility of genotyping error, and the lack of testing for population stratification. While tantalising pre-clinical studies have drawn our attention to these pathways, only further attempts at replication will tell whether these associations with human disease are real.

Acknowledgement

The authors thank Drs. Joel N. Hirschhorn and Christopher J. O’Donnell for their helpful comments on the manuscript.

References


