Neuregulin 1 Genotype and Schizophrenia

Marcus R. Munafo¹,², Angela S. Attwood², and Jonathan Flint³

¹Department of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol BS8 1TU, UK; ²Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK

The neuregulin 1 (NRG1) gene has been the subject of considerable excitement within the psychiatric genetics literature since it was originally identified as a potential susceptibility locus for schizophrenia. Here we provide an update of our first meta-analysis of this association. Case-control and family-based genetic association studies of the NRG1 gene in healthy control groups and clinically diagnosed schizophrenia patients were included. We repeated the search strategy in our earlier meta-analysis for studies published between December 31, 2005, and September 30, 2007, and updated the results of our original meta-analysis accordingly. Superficially, the results of our updated meta-analysis are consistent with those in our previous report, although it is notable that the strength of evidence as based on our haplotype analysis has weakened over this period. The evidence for association of the SNP8NRG221533 polymorphism continued to be nonsignificant. We discuss a number of problems in the interpretation of a disparate and inconsistent gene-disease association literature, including the difficulties associated with determining what constitutes replication across studies which vary in their methods, marker sets employed, phenotype definition, and other study characteristics.

Key words: schizophrenia/neuregulin 1/NRG1/genetic association/meta-analysis

Introduction

The neuregulin 1 (NRG1) gene, located at 8p12-p12, has been the subject of considerable excitement within the psychiatric genetics literature over the last few years since it was originally identified as a potential susceptibility locus for schizophrenia following a linkage study carried out in Iceland.¹ The gene is large, consisting of at least 15 alternative transcripts spanning approximately 1.4 Mb.² In the original association study of NRG1 and schizophrenia,¹ a haplotype of 5 single-nucleotide polymorphisms (SNPs) and 2 microsatellite markers was significantly overrepresented in cases compared with healthy controls. Of the markers typed and included in the haplotype, one (SNP8NRG221533, a T>C SNP) was the single most significant variant within this haplotype and was suggested to be a potential haplotype-tagging SNP.

Since the initial report, there have been numerous attempts to replicate this association, with many focusing on the original 7-marker high-risk haplotype (referred to as HAPICE), as well as other sequence variants within the gene. Given work in other domains of the psychiatric genetics literature, where initial excitement following reports of gene-disease associations is frequently followed by failure to replicate and a gradual decline in the strength of evidence for association³,⁴ we were motivated in 2006 to assess the strength of evidence for the putative association between NRG1 and schizophrenia in a meta-analysis of published studies.⁵

In this meta-analysis, we observed no evidence for association of the SNP8NRG221533 marker with schizophrenia and observed a negative association between date of publication and strength of association, consistent with evidence that initial reports of genetic association typically provide a substantially greater effect size estimate than in those studies which subsequently attempt to replicate this association.⁶ However, a complicating feature of the NRG1 schizophrenia literature is that multiple haplotype combinations have been tested across studies, frequently using overlapping but not identical marker sets. We therefore also combined individual haplotype P values, using the most promising haplotypes identified in individual studies, and hence obtained evidence that variation in NRG1 may be associated with risk of schizophrenia. Unfortunately, this approach is extremely nonconservative, because it takes the “best case” result from each individual study, and does not allow any overall estimate of effect size to be determined.

Here we provide an update of our original meta-analysis, and discuss problems with the interpretation of the NRG1 schizophrenia literature, and more generally

¹To whom correspondence should be addressed; tel: +44-117-9546841, fax: +44-117-9288588, e-mail: marcus.munafo@bristol.ac.uk.

© The Author 2007. Published by Oxford University Press on behalf of the Maryland Psychiatric Research Center. All rights reserved.
For permissions, please email: journals.permissions@oxfordjournals.org.
We repeated the search strategy described previously. Briefly, for the haplotype analysis, we combined the lowest (ie, most highly significant) individual haplotype P value for the strongest reported haplotype association with schizophrenia case status, while a secondary outcome measure was the allele frequency odds ratio (OR) for the SNP8NRG221533 polymorphism.

Inclusion and exclusion criteria, data extraction methods, and statistical analysis were the same as reported previously. Briefly, for the haplotype analysis, we combined the lowest (ie, most highly significant) individual haplotype P values reported in individual studies by first transforming them to Z statistics and then combining them using a modified Stouffer method, either unweighted or weighted by sample size. The selection of haplotypes for inclusion was based solely on the strongest result reported in each study, despite this being a nonconservative method because there are at present no agreed techniques for determining the most appropriate criteria for comparison at the level of the gene (as opposed to individual variant).

Methods

Case-control and family-based genetic association studies of the NRG1 gene in healthy control groups and clinically diagnosed schizophrenia patients were included. We repeated the search strategy described previously for studies published between December 31, 2005, and September 30, 2007. Unlike in our previous meta-analysis, the principal outcome measure was the haplotype P value for the strongest reported haplotype association with schizophrenia case status, while a secondary outcome measure was the allele frequency odds ratio (OR) for the SNP8NRG221533 polymorphism and schizophrenia case status.

Inclusion and exclusion criteria, data extraction methods, and statistical analysis were the same as reported previously. Briefly, for the haplotype analysis, we combined the lowest (ie, most highly significant) individual haplotype P values reported in individual studies by first transforming them to Z statistics and then combining them using a modified Stouffer method, either unweighted or weighted by sample size. The selection of haplotypes for inclusion was based solely on the strongest result reported in each study, despite this being a nonconservative method because there are at present no agreed techniques for determining the most appropriate criteria for comparison at the level of the gene (as opposed to individual variant).

Results

We identified 8 new studies through our search strategy published subsequent to our original meta-analysis, comprising 5 case-control studies and 3 family-based studies. One study reported 1-sided P values for haplotype analyses, which were converted to 2-sided values for the purposes of the present analysis. Two studies did not include the SNP8NRG221533 variant in their analysis, while a further one included this SNP but did not report data in a format that enabled their inclusion in our meta-analysis.

A further study was identified but excluded because it reported data on a sample included in another report. Two further studies investigated gene × gene interaction effects involving NRG1 in samples already included in our original meta-analysis and were therefore excluded. The characteristics of included studies are described in table 1.

Combining the most highly significant haplotype P values of these newly identified studies with those in our previous meta-analysis indicated a significant association of NRG1 with schizophrenia in both the unweighted (P = .026) and weighted (P = .036) analyses. There was evidence of significant between-study heterogeneity (χ²(25) = 28.22, P = .042). Haplotype P values for individual studies by year of publication are presented in figure 1.

As in our original meta-analysis, we observed no evidence of association of the SNP8NRG221533 polymorphism with schizophrenia case status using either a fixed-effects model (OR = 1.04, 95% confidence interval [CI] = 0.99, 1.10, P = .10) or a random-effects model (OR = 1.02, 95% CI = 0.95, 1.10, P = .61). There was evidence of significant between-study heterogeneity (χ²(23) = 54.45, P < .001).

Discussion

Superficially, the results of our updated meta-analysis are consistent with those in our previous report, although it is notable that the strength of evidence as based on our haplotype analysis has weakened over this period. The evidence for association of the SNP8NRG221533 polymorphism with schizophrenia...
polymorphism with schizophrenia continued to be non-significant. However, a number of issues were highlighted by our update. The significant between-study heterogeneity observed for SNP8NRG221533 is of potential interest because it suggests that any effects of this variant may operate differently in distinct populations, but we were not able to ascertain any likely cause of this heterogeneity, on the basis of individual study characteristics such as sample ancestry or sex distribution.

One is that the combination of \( P \) values from haplotype analyses is insufficiently conservative and likely to overestimate the strength of evidence. This problem is further compounded by the use of nonconservative methods in published studies, such as failure to correct for multiple testing or the reporting of 1-sided significance levels. For example, inclusion of measures to deal with multiple testing unsurprisingly reduces the significance of findings; after applying a permutation test to their data, Thomson et al.\(^\text{10}\) found that evidence for association fell from a nominal level of \( P = .00014 \) to \( P = .024 \). In our meta-analysis, we have only been able to use the nominal significance level achieved in haplotype analyses for consistency across studies. A more consistent and systematic use of appropriate correction methods within primary studies would afford more conservative results, which would in turn reduce the problems associated with combining \( P \) values in a meta-analysis.

A further problem is the use of disparate marker sets in haplotype analyses. While replication at the level of the gene has been recommended,\(^\text{18}\) the use of multiple marker sets, and multiple haplotype combinations of markers, has the potential to increase the risk of Type I error. This is potentially compounded by the use of multiple phenotypes; eg, in our updated search, we identified a number of studies that investigated \( NRG1 \) in relation to other phenotypes such as medication response\(^\text{19}\) and childhood-onset schizophrenia.\(^\text{20}\) Such studies are clearly potentially valuable, but they do not represent replications of the primary gene-disease association. Despite this, they may be regarded as supporting the original gene-disease association. There is a need for clarity and consensus regarding what constitutes a replication, both at the level of genotype and at the level of phenotype, as discussed in a recent National Cancer Institute-National Human Genome Research Institute working group report.\(^\text{21}\)

Finally, the widely observed pattern of a negative correlation between the strength of evidence for gene-disease association and the date of publication, the so-called “winner’s curse,”\(^\text{4}\) is apparent in the \( NRG1 \) schizophrenia literature not only for evidence for individual markers but also for haplotype analyses. The true strength of association may not become apparent until many years after the date of the initial positive report. Unfortunately, there appears to be a strong primacy effect in the perception of published gene-disease associations, where the first published study exerts a disproportional influence over the imagination of the scientific community.

While the \( NRG1 \) gene remains a promising susceptibility gene for schizophrenia, evidence to date for this association can only at best be regarded as preliminary. More comprehensive and consistent analysis of variation at this locus, ideally in multiple studies with large case and control samples, is necessary before its involvement in the etiology of schizophrenia can be considered to be confirmed. The development of standardized reporting guidelines for association studies, and the use of conservative methods for testing association in haplotype analyses, should also be encouraged.

**Funding**

National Alliance for Research on Schizophrenia and Depression (Young Investigator Award to M.R.M.); Wellcome Trust to Flint (to J.F.).

**Acknowledgments**

The authors are grateful to Dr Pippa Thomson for releasing data in a format that enabled their inclusion in our meta-analysis.

**References**

4. Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG, Ioannidis JP. Establishment of genetic associations for


