Human-Mouse Quantitative Trait Locus Concordance and the Dissection of a Human Neuroticism Locus

Janice M. Fullerton, Saffron A.G. Willis-Owen, Binnaz Yalcin, Sagiv Shifman, Richard R. Copley, Susette R. Miller, Amarjit Bhomra, Stuart Davidson, Peter L. Oliver, Richard Mott, and Jonathan Flint

Background: Exploiting synteny between mouse and human disease loci has been proposed as a cost-effective method for the identification of human susceptibility genes. Here we explore its utility in an analysis of a human personality trait, neuroticism, which can be modeled in mice by tests of emotionality. We investigated a mouse emotionality locus on chromosome 1 that contains no annotated genes but abuts four regulators of G protein signaling, one of which (*rgs2*) has been previously identified as a quantitative trait gene for emotionality. This locus is syntenic with a human region that has been consistently implicated in the genetic aetiology of neuroticism.

Methods: The functional candidacy of 29 murine sequence variants was tested by a combination of gel shift and transient transfection assays. Murine sequences that contained functional variants and exhibited significant cross-species conservation were prioritized for investigation in humans. Genetic association with neuroticism was tested in 1869 high and 2032 low unrelated individuals scored for neuroticism, selected from the extremes of 88,141 people from southwest England.

Results: Fifteen sequence variants contributed to variation in the expression of *rgs18*, the gene lying at the edge of the quantitative trait loci (QTL) interval. There was no evidence of association between neuroticism and single nucleotide polymorphisms (SNPs) lying in the human regions homologous to those of mouse functional variants. One SNP, rs6428058, in a region of sequence conservation 644 kb upstream of RGS18, showed significant association ($p = .000631$).

Conclusions: It is unlikely that a single variant is responsible for the mouse emotionality locus on chromosome 1. This level of underlying genetic complexity means that although cross-species QTL concordance may be invaluable for the identification of human disease loci, it is unlikely to be as informative in the identification of human disease-causing variants.

Key Words: Emotionality, neuroticism, quantitative trait locus, RGS2, RGS18

The heritable personality trait of neuroticism has a substantial genetic correlation with both major depression (1) and generalized anxiety disorder (2). Consequently, successful genetic dissection of this trait is expected to improve our understanding of the molecular basis of at least two common disabling psychiatric disorders. However, although a number of independent linkage analyses have identified a locus on human chromosome 1 that contributes to variation in neuroticism (3–5) and other related traits (6), to date there has been little success in implicating specific gene or genes in this region. One reason for this is likely to be the relatively small effect sizes that underlie human complex trait loci, as found in numerous meta-analyses (7) and evidenced by recent successes in whole-genome association studies (8).

Using DNA pooling to conduct a whole genome association study, we have shown that the heritability of neuroticism is attributable to a large number of loci, each contributing less than 1% of the phenotypic variance (9). This implies that a sample size on the order of tens of thousands will be required for robust genomewide analysis of this trait. Strategies that reduce the number of candidate genes to be tested would in turn lead to a reduction in this extremely large sample size. One such strategy is to focus on homologous loci known to be important in genetically more tractable species such as *Mus musculus* (common house mouse), an approach that has previously proved successful in the genetic dissection of several physiological phenotypes (10–12). However, an appropriate animal model is necessary if the approach is to be useful for human neuroticism.

We use mouse emotionality as that model on the basis of arguments put forward by Gray and colleagues (13–15) that activity and defecation in the open-field arena are indices of an animal's emotional state that can be modulated by anxiolytics, such as benzodiazepines, and involve the same neuroanatomic pathways as trait anxiety in humans (trait anxiety is one manifestation of neuroticism). Gray's argument partly rested on the results of a series of artificial selection experiments in rodents carried out by Hall (16), Broadhurst (17,18) and DeFries (19–21). These experiments demonstrated a common genetic basis for a number of measures of emotionality, including activity and defecation in the open-field arena. Molecular mapping in crosses of inbred rodents have gone on to confirm the existence of pleiotropic loci, influencing both activity and defecation (22–24). One such locus, on the end of mouse chromosome 1, has been consistently identified and may be syntenic with the human neuroticism locus just outlined.

This highly replicated mouse emotionality locus was originally mapped to a 4.8-Mb region spanning 144–148.8 Mb (NCBIm36) on mouse chromosome 1 in outbred heterogeneous stock (HS) mice (25). When the locus was further dissected by genetic association in an outbred stock of MF1 mice, in which linkage disequilibrium is considerably less, we obtained evidence that the 4.8 Mb region contained three independent genetic effects (26), delineated by horizontal black bars in Figure 1. Quantitative complementation was used to show that one of these effects acts through the gene *rgs2,* a regulator of G-protein

From the Wellcome Trust Centre for Human Genetics (JMF, SAGW-O, BY, SS, RRC, SRM, AB, SD, RM, JF), Headington, and Department of Physiology (PLO), Anatomy and Genetics, University of Oxford, Oxford, United Kingdom.

Address reprint requests to Jonathan Flint, M.D., Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Headington, Oxford OX3 7BN, United Kingdom. E-mail: jf@well.ox.ac.uk.

Received August 10, 2007; revised September 28, 2007; accepted October 17, 2007.

0006-3223/08/$34.00
doi:10.1016/j.biopsych.2007.10.019

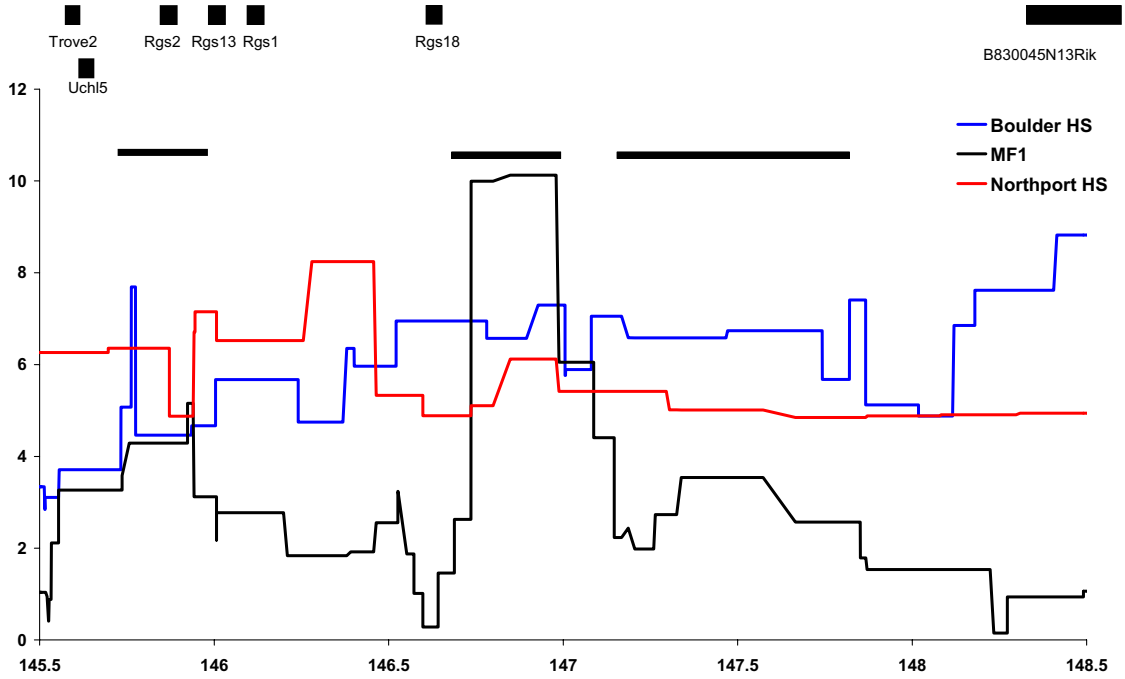

Figure 1. Genetic mapping of EMO (emotionality) in the Boulder heterogeneous stock (HS) (blue line), Northport HS (red line), and MF1 (black). Genes are shown as black boxes. The vertical axis gives the negative logarithm of the *p* value of the genetic association. The horizontal axis is the position on chromosome 1, using a megabase sale. Coordinates are based on build 36 of the mouse genome. The 95% confidence intervals of quantitative trait loci found in the MF1 are shown as black lines.

signaling (RGS); however, the molecular basis for the remaining two loci has not yet been determined (26). Both these effects are located in an intergenic region, upstream of another member for the RGS gene family, *rgs18*; one is contained within a 200-Kb 95% confidence interval between 146.65 and 146.85 Mb (Build 36), and a second 700-Kb interval between 147.15 and 147.87 Mb.

We have estimated that the 200-Kb quantitative trait loci (QTL) interval includes approximately 600 potentially causative sequence variants (27). To reduce this number, we exploited the mosaic structure of the mouse genome to implement a test of the functional candidacy of each sequence variant in the MF1 (28). In heterogeneous stocks, probabilistic reconstruction of the ancestral haplotypes can be used to estimate QTL effects from each progenitor strain (29). This strain distribution pattern can then be compared with the way assayed sequence variants partition the founder strains. In cases in which the strain distribution pattern of the sequence variants is inconsistent with the QTL effect, that region of sequence can be excluded from consideration as the location of the causative variant. We have called this technique a merge analysis (28).

In this study, we sought to establish first whether any of the sequence variants identified by the merge analysis are functional and to arrive at a list of potential quantitative trait nucleotides (in other words, the molecular bases of quantitative variation in emotionality) at least at this locus. We used two in vitro indices of functionality: we asked whether the sequence variants give rise to a difference in protein binding and whether the sequence variants influence gene expression (using *rgs18*, the closest gene to the 95% confidence interval). These experiments were designed to reduce the 600 potentially causative sequence variants to a small set of candidate functional regions. It is important to emphasize that our experiment was designed to detect regions of the mouse genome where allelic variation has an effect on function, rather than simply to detect regions that had an effect on protein binding or gene expression.

We then asked whether any of the functional regions from the mouse experiments contribute to variation in human neuroticism. To answer this question, we used human sequence variants at the homologous loci and tested for genetic association between these sequence variants and variation in neuroticism scores. Because the locus-specific contribution effect was expected to be small, we tested the extremes of a large sample of 88,141 people from southwest England who completed an Eysenck personality questionnaire (5,9). This sample had power to detect effects contributing to less than 1% of the phenotypic variation (depending on the allele frequency of the variant). In this way, we could investigate whether mouse quantitative trait nucleotides could be used to guide the search for human quantitative trait nucleotides.

Methods and Materials

Genetic Mapping of EMO

The phenotypic construct of EMO, or rodent "emotionality," in the anxiogenic open-field arena has previously been described (25). In brief, EMO is a composite phenotype of two inversely correlated mouse open-field anxiety-related behaviors; ambulation and defecation. Three outbred mouse populations, including two heterogeneous stock lines generated from eight known inbred founder strains (Northport [30] and Boulder HS [31]), and an outbred line with unknown progenitors (MF1) (26), were subject to behavioral testing and genotyping of a previously described QTL influencing EMO on mouse chromosome 1 (25). Genetic loci contributing to variation in EMO were mapped via progenitor mapping using the HAPPY algorithm (29). Strain