Accurate, large-scale genotyping of 5HTTLPR and flanking SNPs in an association study of depression, anxiety and personality measures

Naomi R Wray[1,5], **Michael R James**[1], **Scott D Gordon**[1], **Troy Dumenil**[1], **Leanne Ryan**[1], **William L Coventry**[1,2], **Dixie J Statham**[1,4], **Michele L Pergadia**[3], **Pamela AF Madden**[3], **Andrew C Heath**[3], **Grant W Montgomery**[1], and **Nicholas G Martin**[1]

[1]Genetic and Molecular Epidemiology Laboratories, Queensland Institute of Medical Research, Brisbane, Queensland 4006, Australia

[2]School of Behavioural, Cognitive and Social Sciences, University of New England, Armidale, New South Wales 2350, Australia

[3]Washington University School of Medicine, Department of Psychiatry, St Louis, Missouri, USA

[4]Faculty of Arts & Social Sciences, University of the Sunshine Coast, Maroochydore, Queensland 4558, Australia

Abstract

Background—5HTTLPR, the length polymorphism repeat in the promoter region of the serotonin transporter gene (5HTT renamed SLC6A4) is one of the most studied polymorphisms for association with a range of psychiatric and personality phenotypes. However, the original 5HTTLPR assay is prone to bias toward short allele calling.

Methods—We designed new assays for the 5HTTLPR suitable for large scale genotyping projects and we genotyped 13 SNPs in a 38kb region around the 5HTTLPR including SNP rs25531, a polymorphism of the 5HTTLPR long allele. Association analysis was conducted for major depression and/or anxiety disorder in unrelated cases (N = 1161) and controls (N = 1051) identified through psychiatric interviews administered to a large population sample of Australian twin families. Participants had been scored for personality traits of neuroticism, extraversion and harm avoidance several years earlier (N ≥ 2643 unrelated individuals).

Results—Using the linkage disequilibrium (LD) between markers we identified a two SNP haplotype proxy for 5HTTLPR; the CA haplotype of SNPs rs4251417 and rs2020934 is coupled with the short allele of 5HTTLPR ($r^2 = 0.72$). We found evidence for association (p=0.0062, after accounting for multiple testing) for SLC6A4 SNPs rs6354 and rs2020936 (positioned in a different LD block about 15.5kb from 5HTTLPR) with anxiety and/or depression and neuroticism,

URL
Ensembl URL for Figure 1: http://www.ensembl.org/Homo_sapiens/exonview?transcript=OTTHUMT00000256114;db=vega

with the strongest association for recurrent depression with onset in young adulthood (OR = 1.55, 95% CI 1.16–2.06).

Conclusions—The associated SNPs are in the same LD block as the VNTR STin2, for which association has previously been reported.

Introduction

Serotonergic neurotransmission impacts on a wide range of behaviours including cognition and emotion (1,2) and drugs targeting serotonin reuptake are clinically effective antidepressants (3). As a result, one of the most studied polymorphisms for association with a broad range of psychiatric and personality phenotypes is the length polymorphism repeat (LPR) in the promoter region of the serotonin transporter gene (5HTT renamed SLC6A4). The 5HTTLPR polymorphism comprises a 43-base pair (4–8) insertion or deletion (long, 'L' with 16 repeat units, or short, 'S' with 14 repeat units, alleles respectively). The S allele (frequency in Caucasians ~0.45 (9)) reduces transcriptional efficiency resulting in decreased SLC6A4 expression and function (10). Association studies and subsequent meta-analyses have shown conflicting results in support (Table 1) (11–15) of an association between the S allele and anxiety, depression and the personality trait neuroticism (a measure of emotional stability that is genetically correlated to both anxiety and depression (16–18)). Conflicting results have dogged candidate gene association studies for many complex disorders, attributable to small sample sizes of both primary and replication studies, heterogenous subject populations, association dependent on environmental conditions such as stressful life events (19), differing instruments for assessment of phenotypic traits and statistical methods e.g.,(20,21). However, an additional problem specific to 5HTTLPR relates to the genotyping assay, which has caused considerable bias towards S allele identification (22,23). Furthermore, association may have been compromised by the presence of an A/G single-nucleotide polymorphism (SNP), rs25531 that lies within the L allele of 5HTTLPR (5,8); the L allele with the rarer G allele of rs25531 (denoted L_G) is functionally equivalent to the S allele, because of changes to the AP2 transcription factor binding site altered by this SNP (5,8).

The aim of this study was to investigate the association between the 5HTTLPR polymorphism including the rs25531 polymorphism and psychiatric and personality phenotypes in a large cohort. As part of our study, we designed new assays for the 5HTTLPR suitable for large scale genotyping projects and we genotyped tagging SNPs in a 38kb region around the 5HTTLPR (Figure 1A) to determine whether any SNP or combination of SNPs could be used as proxies for the difficult and time consuming 5HTTLPR assay. Our study design allows us to examine association in multiple traits within the same cohort and consistency across independent sub-cohorts. Within our study sample we can identify subsets of cases which are predicted to be genetically more homogenous. For example, the relative risk (RR) to first degree relatives is reported to be higher for recurrent early-onset depression (RR ≈ 4–5) (reviewed in (24)) compared to major depression (RR ≈ 2–3). Similarly, although estimates for RR are not available, anxiety co-morbid with depression is also considered to be a genetically more homogeneous group (reviewed in (25)). The optimum balance of sample size versus sample homogeneity cannot be predicted since the true genetic aetiology is unknown. However, the availability of a depth of phenotypic information allows us to investigate associations through case subsets.

Materials and Methods

Samples

All participants were adult twins and their families recruited through the Australian Twin Registry and provided informed consent under study protocols approved by the Queensland Institute of Medical Research and Washington University Human Research Ethics Committees. During the period 1988–1990 study participants from two twin birth cohorts (born 1890–1964, and 1964–1971, respectively) were mailed an extensive Health and Lifestyle Questionnaire ('1989 Questionnaire Survey'(26,27)). This included the shortened revised 48-item Eysenck Personality Questionnaire (EPQ) (28) and a short-form 54-item version of the Temperament and Character Inventory (26,29), the Temperament Personality Questionnaire (TPQ). The EPQ measures four dimensions of personality including neuroticism and extraversion and the TPQ measures three dimensions of temperament including harm avoidance. Between 1992–1994 twins from the older cohort (N=5,995) (30) were interviewed by telephone using the SSAGA (Semi-Structured Assessment for the Genetics of Alcoholism) instrument, a comprehensive psychiatric interview designed to assess psychiatric disorders in adults (31) according to DSM-IIIR and subsequently updated to DSM-IV criteria (32) that was modified for use as a telephone survey instrument in Australia (SSAGA-OZ). Of these participants, N = 4,597 have subsequently provided a blood sample (or rarely a buccal or saliva sample) for genotyping (33). Over the period 1996–9, sibling pairs participating in the 1989 questionnaire survey that were either concordant or discordant for extreme EPQ neuroticism scores (one sibling in the top or bottom decile, the other sibling in the top or bottom quintile, and allowing inclusion of multiple siblings) were recruited (35–37). Participants (N = 2456) completed the shortened Composite International Diagnostic Interview (34) (CIDI), which provides DSM-IV (35) life-time diagnoses of depression (including recurrent major depression) and anxiety disorders. This extreme discordant and concordant (36) design is a cost efficient strategy for obtaining an informative data set for genetic studies (37). Of these participants 2213 provided DNA samples, plus 837 parents. Full details of the recruitment procedure for study, including response rates and incidence of DSM-IV diagnoses for anxiety and depression related disorders are given elsewhere (38–40). Finally, between 2001–2005 some participants from the earlier studies were reinterviewed again using an adaptation of the SSAGA, see (41,42) for details. Participants reported ancestry of all four grandparents and a small number with two or more grandparents with known non-European ancestry were excluded from the analysis. Selection of individuals used in the association analyses are described in the *Association Analysis* section below.

Genotyping

Genomic DNA was extracted from blood (or buccal) samples using standard protocols (43). Samples were plated in 384 well plates in two study sample sets comprising i) MZ and DZ twins from the 1992–2000 SSAGA interview studies and ii) participants and their parents from the CIDI interview. Some 764 DNA samples are included in both sets, providing an opportunity for extensive quality control checking.

5HTTLPR assay—The original assay (44) for the 5HTTLPR used PCR primers in the non-repetitive sequences that flanked the 16 repeat elements, which are each comprised of between 19 and 23 base pairs (bp) (10,44) (Figure 1B). This assay proved less than ideal: the PCR is difficult because of the very high GC content, and the long length of the PCR products. In pilot studies we found that these difficulties cause considerable bias towards S allele identification, since the L allele signal is weak so that heterozygotes are frequently mis-scored as S homozygotes (eg. see Figure 1D of 44). This observation has also been made by others e.g. (22,23). We improved on the assay by redesigning the PCR primers, full

details are provided in the Supplement 1. The final assay comprised a multiplexing scheme of PCR primers L3+R3 and L3+R4 located within the repeat elements (Figure 1 B). This choice of primers provided two PCR products which are semi-independent (L3 is shared) and which could be analysed together on a gel. The primer products were pooled (one DNA sample from each primer product) and run on a gel together (twofold multiplexing). Further, fourfold multiplexing was achieved by loading a second pair of samples on to the same gel (Figure 1C). In operation, a 384-well plate of DNAs can be amplified for both products and pooled in a way to allow duplicate scoring for each individual on two gels. Our pilot experiments showed that this was an efficient and robust system to genotype this difficult polymorphism.

SNP and STin2 genotyping—The genotyping assay for SNP rs25531 [A/G] is described in the Supplement 1. DNA samples were genotyped for 13 additional SNPs (see Figure 1) in the 38kb region surrounding the 5HTTLPR. Our motivation was firstly, to determine whether there is a SNP-based proxy for 5HTTLPR which would be easier to genotype in large sample sets, and secondly, to see if any SNPs in this region show stronger association with any of the phenotypic measures than 5HTTLPR. SNPs were selected to represent the linkage disequilibrium landscape of the region bounded by rs140700 (intron 6) and rs7214991 (promoter) and preference was given to SNPs used in association studies of autism (45, 46). All non-genotyped HapMap (build 35) SNPs with minor allele frequency > 0.1 were represented by single SNP genotyped proxies with $r^2 > 0.8$. Three non-synonymous SNPs were also included: rs28914832 (isoleucine/valine exon 10, also known as SERT Ile425Val), rs6355 (glycine/alanine exon 3, also known as Gly56Ala) and rs1050565 (isoleucine/valine exon 12 of BLMH 13kb from 5HTTLPR).

In pilot experiments we attempted to genotype the 17 bp variable number tandem repeat (VNTR) of SLC6A4 (Figure 1A), known as STin2. However, genotyping results did not pass our quality control checks (including Mendelian error checking) and so STin2 was not genotyped on the full study sample; problems with the assay have been found by others e.g. (22).

Genotype analysis

Our study design allowed us to undertake quality control checks often not possible in association studies (see Supplement 1). As the rare L_G alleles are functionally equivalent to S alleles (4, 8), we constructed a 5HTTLPR+rs25531 (Marker 12) variable with three levels used in the analysis, SS&SL_G, SL_A& L_GL_A and L_AL_A. After exclusion of identifiable errors, genotyping call rates ranged from 97.6% to 99.2% for the 13 SNPs excluding rs25531, but were only 96.9% for 5HTTLPR and 95.4% for 5HTTLPR+rs25531; these latter rates are lower than normal for our laboratory, but nonetheless much improved over the original assay for 5HTTLPR. Tests of deviation from Hardy-Weinberg equilbrium (HWE) and estimates of linkage disequilibrium were calculated in Haploview(47) using the SSAGA sample (a population sample) with one individual randomly selected per family (n = 2341). Within Haploview, the Tagger (48) option was used to force exclusion of 5HTTLPR to see if any one, two or three marker combination could be used to predict 5HTTLPR genotype. Similarly, we used Tagger to predict rs25531 genotype from other genotyped SNPs in those with LL 5HTTLPR genotypes.

Association Analysis

Sum scores of 12 item responses in each EPQ domain resulted in quantitative scores for neuroticism and extraversion. Harm avoidance comprised sum scores of 18 items from the TPQ. Scores for MZ twins were averaged. Sum scores were transformed using the averaged angular transformation (49). Scores used for analysis were standardized residuals

(standardised separately for each sex) after regression of the transformed scores to remove effects of sex and age. This standardisation was conducted using the population sample from the full study sample (n > 20,000), not just those genotyped (see (40)). Association analysis for quantitative traits used one individual per family selecting the individual that deviated most from the population mean, and were conducted in PLINK (50); in order to account for potential bias in the ascertainment of samples for analysis, permutation p-values are reported.

Cases were selected as unrelated individuals with a DSM-IV (51) life-time diagnosis of major depression or an anxiety disorder. Early age of onset for depression was defined as those with self-report age of onset of less than 30 years old. Controls were selected as unrelated individuals, from families in which no siblings who completed the questionnaires (including those not supplying a DNA sample) received a diagnosis of either major depression or of an anxiety disorder from any of the surveys. Details of selection of individuals for analysis within families are provided in Supplement 1. SSAGA controls were only selected if their neuroticism scores were less than the population average. Neuroticism scores of CIDI controls reflected the ascertainment into the study (Table 3). Cases and controls were divided into two independent samples according to the interview instrument. Families who participated in both studies were allocated to the CIDI case-control sample, unless the only anxiety or depression diagnoses for the family were allocated from the SSAGA instrument (2.7% of those measured on both instruments). Thus, the CIDI and SSAGA case control data sets (Table 2) are independent samples from the same population. An additional 133 cases were identified by the SSAGA reinterview. Association analysis was conducted for cases with major depression disorder (Depression), any anxiety disorder (Anxiety), or Anxiety and/or Depression (if they qualify for a diagnosis of any anxiety disorder and/or major depression). Analyses were also conducted for the case subset groups thought to represent genetically more homogeneous groups: Anxiety with Depression (if they qualify for both diagnoses) and recurrent, early age of onset depression (Table 2). Following up on literature reports discussed in the *Introduction*, association analysis of all markers was also conducted for Panic/Agoraphobia and OCD. However, for the other specific anxiety disorders, follow-up analyses were only conducted for SNPs showing association in the primary analyses. Logistic regression association analysis was conducted using PLINK (50) under a model that assumes additivity of allelic effects on the log(risk) scale. Power calculations and details of the permutation test used to determine the empirical significance of the pattern of association observed across data sets and phenotypes are provided in Supplement 1.

Results

The numbers of genotyped individuals for the two independent samples used in the association analysis are shown in Table 2. The ascertainment of the CIDI sample means that both the cases and controls are likely to be more extreme that those of the SSAGA sample, and this is reflected in the summary statistics based on EPQ neuroticism and harm avoidance (Table 3) collected several years earlier.

Linkage disequilibrium around 5HTTLPR

None of the markers showed evidence for Hardy-Weinberg disequilibrium. Figure 2 shows the linkage disequilibrium as both $|D'|$ and r^2 measures; although the $|D'|$ values are high, the corresponding r^2 values are lower reflecting differences in allele frequencies (52). The best prediction of 5HTTLPR is provided by rs4251417 and rs2020934 (Figure 2b), where the CA haplotype is coupled with the S allele and the CG haplotype is coupled with the L allele of 5HTTLPR ($r^2 = 0.72$). No useful prediction could be made of 5HTTLPR-rs25531

or of rs25531 within LL individuals. Genotype data are available from the corresponding author on request.

Association analysis

SNPs rs6354 and rs2020936 (markers 4 and 6) show association ($p < 0.05$) with all of the depression and anxiety phenotypes presented in Figure 1, with the T and A alleles of these SNPs respectively, being more frequent in cases than controls (0.82 vs 0.72, Odds ratio (OR) = 1.19 (1.03 – 1.38)). These SNPs are in high LD ($r^2 = 0.98$) and so represent a single association signal. Although the associations for the two independent data sets are only significant for some analyses the direction of the association is the same, so that the association in the combined data set is significant. The association with Depression is marginally significant, OR = 1.17 (1.01 – 1.26). However, the association with the two subsets of cases thought to represent genetically more homogenous groups (24,25), both show higher association odds ratios: recurrent, early onset depression (Figures 3e and 4), OR = 1.55 (1.16–2.06), and co-morbid group of Anxiety with Depression (Figure 3d and 4), OR = 1.31 (1.02–1.69). Permutation analysis showed that this pattern of association occurred only 62 times in 10,000 permutations across all markers under the null hypothesis, resulting in an empirical p-value which accounts for multiple testing of 0.0062. Odds ratios of association with rs6354 for the different anxiety disorders are presented in Figure 4. These results must be interpreted recognising the constraints on possible diagnostic classes from the questionnaires. Necessarily these results should be considered as hypothesis generating for future studies. The associations for rs6354 and rs2020936 with neuroticism were 0.019 and 0.015 respectively (Figure 3f), but there is no evidence for association between either harm avoidance or extraversion which had phenotypic correlations of 0.66 and −0.25, respectively, with neuroticism. Association ($p < 0.05$) is seen for rs4251417 and harm avoidance. Haplotype analysis of combinations of SNPs revealed no associations more significant than those of the individual genotyped markers.

For the association analysis in the CIDI sample we observed association ($p < 0.05$) for the 5HTTLPR (L allele), most significantly with the Depression cases. We interpret this as a chance result as the association is not replicated in the SSAGA sample for which any trend in association is with the S allele. A detailed investigation is provided in Supplement 1. We find no evidence for association of 5HTTLPR (with or without rs25531) with panic/agoraphobia, OCD, neuroticism, extraversion or harm avoidance (Figure 3 e and f). As a consequence, our results provide no evidence, either way, for the usefulness of genotyping of SNP rs25531 to subclassify the L allele. A full report of allele and genotype frequencies and p-values is available online (Supplement 2).

Discussion

Preferential amplification of one allele over another is recognised as a systematic problem in PCR assays for variable repeat polymorphisms (22). In pilot studies using the standard 5HTTLPR assay (44) we found overamplification of the S allele, as also reported by others (22,23). Yonan et al (23) found that the assay was sensitive to the concentration of $MgCl_2$ and Hardy-Weinberg equilibrium frequencies of genotypes could only be achieved at lower concentrations of $MgCl_2$ and found that a disease association with 5HTTLPR disappeared when the revised assay was used. Although genotyping errors could lead to false detection of association, consistent association with one allele detected across studies would require a mechanism by which a sub-optimal assay systematically biases genotypes of cases versus controls. Such a mechanism is hard to contemplate, but ascertainment, age and storage of DNA samples can often differ between case and control sample sets. Our concern with the 5HTTLPR assay cannot conclusively undermine the meta-analysis results of a large number of studies in which the odds ratio for S vs L alleles was estimated to be 1.12 from 24 studies

(11) as it is possible that optimal genotyping conditions in each study had been achieved. However, it is noteworthy that many of the contributing studies had too few samples to detect anything but extreme deviation from HWE. Furthermore, potential problems with the 5HTTLPR assay are often overlooked in discussions of results of 5HTTLPR associations, e.g., (53).

We have developed a new assay for 5HTTLPR based on two pairs of PCR primers which bind directly to the repeat sequences and which can be multiplexed into a single assay. Further, fourfold multiplexing was achieved by delayed double loading of sample replicates onto the same gel. In operation our protocol took only one hour to genotype 384 different DNA samples, providing replication of two assays for each sample. Genotyping of 13 SNPs within a 38kb region flanking 5HTTLPR has enabled us to examine linkage disequilibrium landscape of the region and identify a two SNP haplotype proxy for it; the CA haplotype of rs4251417 and rs2020934 is coupled with the short allele of 5HTTLPR ($r^2 = 0.72$) (Figure 2). Ideally, a higher r^2 is desirable for SNP proxies (usually an r^2 threshold of 0.8 is set for selection of tagging SNPs), but with large sample sizes genotyping this two SNP proxy could usefully replace the time consuming 5HTTLPR assay in association studies, i.e. a sample size increased by a factor of 1/0.72 is required when genotyping the two SNP proxy to achieve the same power as direct genotyping of 5HTTLPR (54). Others (45,46,55) have genotyped SNPs in the 38kb region surrounding 5HTTLPR as part of association studies for autism or bipolar disorder. However, none of these studies (45) genotyped SNP rs4251417 whose minor T allele is useful in splitting the majority of L alleles that couple with the A allele of rs2020934. Despite the large number of association studies for depression and anxiety related traits that have genotyped 5HTTLPR none (to our knowledge) has genotyped more than a couple of additional SNPs in the region. Unfortunately, rs2020934 has not been genotyped as part of the HapMap project and has not been included on any of the genome-wide SNP platforms. rs4251417 is included on the Illumina 610K and 1M chips, but on its own it is not a useful proxy for 5HTTLPR ($r^2 = 0.06$).

We have conducted an association study of 5HTTLPR, including the rs25531 polymorphism found on the L allele, as well as an additional 13 SNPs in a 38 kb region surrounding the 5HTTLPR using a large population cohort rich in psychiatric and personality phenotypic information. Our case and controls represent two independent study samples. The ascertainment for the sample completing the CIDI questionnaire is likely to have identified more extreme cases and controls and this is quantified through personality measures collected several years prior to administration of the psychiatric interviews (Table 3). Although, we might expect to observe more extreme associations with the CIDI cohort we would expect the direction of the association to be the same in the two cohorts if a true association exists. We found no consistent evidence of association with 5HTTLPR, nor with functional genotype classes of 5HTTLPR with SNP rs25531. However, we did find consistent evidence of association with SNPs rs6354 and rs2020936 in both independent samples. These two SNPs are in high LD ($r^2 = 0.95$) and so represent a single association signal; the LD with 5HTTLPR for these SNPs is $r^2 = 0.01$ and $|D'| = 0.20$. The association was upheld across multiple (correlated) phenotypic measures with larger differences in allele frequencies occurring with more extreme, homogenous phenotypes (although because of the smaller sample sizes, association p-values were not always smaller) e.g. association p-values with SNP rs6354 for Depression and/or Anxiety, co-morbid Anxiety with Depression, and recurrent, early onset depression are 0.021, 0.032 and 0.0027 respectively, reflecting control vs case frequencies of the T allele of 0.78 vs 0.81, 0.83 and 0.85, and generating allelic OR (95% CI) of 1.19 (1.03–1.38), 1.31 (1.02–1.69) and 1.55 (1.16 – 2.06) respectively. The higher OR for subsets of cases which are likely to be genetically more homogeneous provide support that this association is not a false positive. Indeed, permutation testing confirmed this, generating a significance of p = 0.0062, which includes correction for testing of

multiple markers. The association was supported through analysis of the quantitative trait neuroticism with measures available on more than 2600 individuals collected several years prior to the diagnostic psychiatric interviews, p = 0.019 with rs6354, reflecting a standardized difference between homozygotes of $d = 0.2$, with a higher score associated with their T and A alleles respectively. There was no evidence for association with extraversion or harm avoidance with these SNPs, and within the anxiety disorders, no evidence was found for association with OCD, although the number of cases is low.

SNPs rs6354 and rs2020936 are positioned in the 15.5kb region that separates the 5HTTLPR and STin2 (they are 1.3 and 2.2 kb from STin2 respectively). These SNPs are not in LD with 5HTTLPR (Figure 2). Using CEU genotypes downloaded from the HapMap (56) website and haplotype information (57) we deduce that rs6354 and rs2020936 are in the same haplotype block as STin2 but we are unable to deduce the likely coupling of alleles between these polymorphisms (see Supplement 1). As for 5HTTLPR, mixed association results have been reported for STin2. A meta-analysis of 8 association studies of STin2 and major depression(11), reported an OR of 1.33 for the rare allele 9 vs 12, but this failed to reach significance because of small sample size, low frequency of the risk allele and heterogeneity between studies. Their result for allele 10 vs 12 was not significant. However, a recent association study of panic and social anxiety disorder reported significant association with rs140701 ($|D'|=1$, $r^2 \approx 0.2$ with rs6354), and no association with 5HTTLPR (58). Furthermore, in a study of allelic expression of variants within SLC6A4 and its promoter, the most significant correlation was found with two SNPs in intron 1 (rs16965628 and rs2020933, r^2 between them of 0.79) which are not in LD with 5HTTLPR. Taken together these studies provide evidence for a role for SLC6A4, but not necessarily 5HTTLPR, in the etiology of depression and anxiety disorder, although whether these results reflect single or multiple causal variants is unclear.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Hariri AR, Holmes A. Genetics of emotional regulation: the role of the serotonin transporter in neural function. Trends in Cognitive Sciences. 2006; 10:182–191. [PubMed: 16530463]

2. Canli T, Lesch KP. Long story short: the serotonin transporter in emotion regulation and social cognition. Nature Neuroscience. 2007; 10:1103–1109.

3. Papakostas GI, Thase ME, Fava M, Nelson JC, Shelton RC. Are antidepressant drugs that combine serotonergic and noradrenergic mechanisms of action more effective than the selective serotonin reuptake inhibitors in treating major depressive disorder? A meta-analysis of studies of newer agents. Biological Psychiatry. 2007; 62:1217–1227. [PubMed: 17588546]

4. Hu XZ, Oroszi G, Chun J, Smith TL, Goldman D, Schuckit MA. An expanded evaluation of the relationship of four alleles to the level of response to alcohol and the alcoholism risk. Alcohol Clin Exp Res. 2005; 29:8–16. [PubMed: 15654286]

5. Hu XZ, Lipsky RH, Zhu G, Akhtar LA, Taubman J, Greenberg BD, et al. Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder. Am J Hum Genet. 2006; 78:815–826. [PubMed: 16642437]

6. Kraft JB, Slager SL, McGrath PJ, Hamilton SP. Sequence analysis of the serotonin transporter and associations with antidepressant response. Biol Psychiatry. 2005; 58:374–381. [PubMed: 15993855]

7. Nakamura M, Ueno S, Sano A, Tanabe H. The human serotonin transporter gene linked polymorphism (5-HTTLPR) shows ten novel allelic variants. Mol Psychiatry. 2000; 5:32–38. [PubMed: 10673766]

8. Wendland JR, Martin BJ, Kruse MR, Lesch KP, Murphy DL. Simultaneous genotyping of four functional loci of human SLC6A4, with a reappraisal of 5-HTTLPR and rs25531. Mol Psychiatry. 2006; 11:224–226. [PubMed: 16402131]

9. Rasmussen HB, Werge TM. Novel procedure for genotyping of the human serotonin transporter gene-linked polymorphic region (5-HTTLPR)--a region with a high level of allele diversity. Psychiatr Genet. 2007; 17:287–291. [PubMed: 17728667]

10. Lesch KP, Bengel K, Heils A, Sabol SZ, Greenberg BD, Petri S, et al. Association of anxiety-related traits with a polymorphism in the serotonin transoprter gene regulatory region. Science. 1996; 274:1527–1531. [PubMed: 8929413]

11. Lopez-Leon S, Janssens A, Ladd A, Del-Favero J, Claes SJ, Oostra BA, et al. Meta-analyses of genetic studies on major depressive disorder. Molecular Psychiatry. 2008; 13:772–785. [PubMed: 17938638]

12. Schinka JA, et al. Measurement scale does moderate the association between the serotonin transporter gene and trait anxiety: comments on Munafo et al. Molecular Psychiatry. 2005; 10:892–893. [PubMed: 15983622]

13. Willis-Owen SAG, Turri MG, Munafo MR, Surtees PG, Wainwright NWJ, Brixey RD, et al. The serotonin transporter length polymorphism, neuroticism, and depression: a comprehensive assessment of association. Biological psychiatry. 2005; 58:451–456. [PubMed: 16023085]

14. Blaya C, Salum GA, Lima MS, Leistner-Segal S, Manfro GG. Lack of association between the Serotonin Transporter Promoter Polymorphism (5-HTTLPR) and Panic Disorder: a systematic review and meta-analysis. Behavioral and Brain Functions. 2007; 3

15. Lin PY. Meta-analysis of the association of serotonin transporter gene polymorphism with obsessive-compulsive disorder. Progress in Neuro-Psychopharmacology & Biological Psychiatry. 2007; 31:683–689. [PubMed: 17291658]

16. Fanous A, Gardner CO, Prescott CA, Cancro R, Kendler KS. Neuroticism, major depression and gender: a population-based twin study. Psychol Med. 2002; 32:719–728. [PubMed: 12102386]

17. Kendler KS, Gatz M, Gardner CO, Pedersen NL. Personality and major depression: a Swedish longitudinal, population-based twin study. Arch Gen Psychiatry. 2006; 63:1113–1120. [PubMed: 17015813]

18. Jardine R, Martin NG, Henderson AS. Genetic covariation between neuroticism and the symptoms of anxiety and depression. Genet Epidemiol. 1984; 1:89–107. [PubMed: 6544237]

19. Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H. Influence of life stress on depression: moderation by a polymorphism in the 5HTT gene. Science. 2003; 301:386–389. [PubMed: 12869766]

20. Ioannidis JPA, Trikalinos TA, Khoury MJ. Implications of small effect sizes of individual genetic variants on the design and interpretation of genetic association studies of complex diseases. American Journal of Epidemiology. 2006; 164:609–614. [PubMed: 16893921]

21. Kavvoura FK, Ioannidis JPA. Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. Human Genetics. 2008; 123:1–14. [PubMed: 18026754]

22. Kaiser R, Tremblay PB, Roots I, Brockmuller J. Validity of PCR with emphasis on variable number of tandem repeat analysis. Clinical Biochemistry. 2002; 35:49–56. [PubMed: 11937078]

23. Yonan AL, Palmer AA, Gilliam TC. Hardy-Weinberg disequilibrium identified genotyping error of the serotonin transporter (SLC6A4) promoter polymorphism. Psychiatr Genet. 2006; 16:31–34. [PubMed: 16395127]
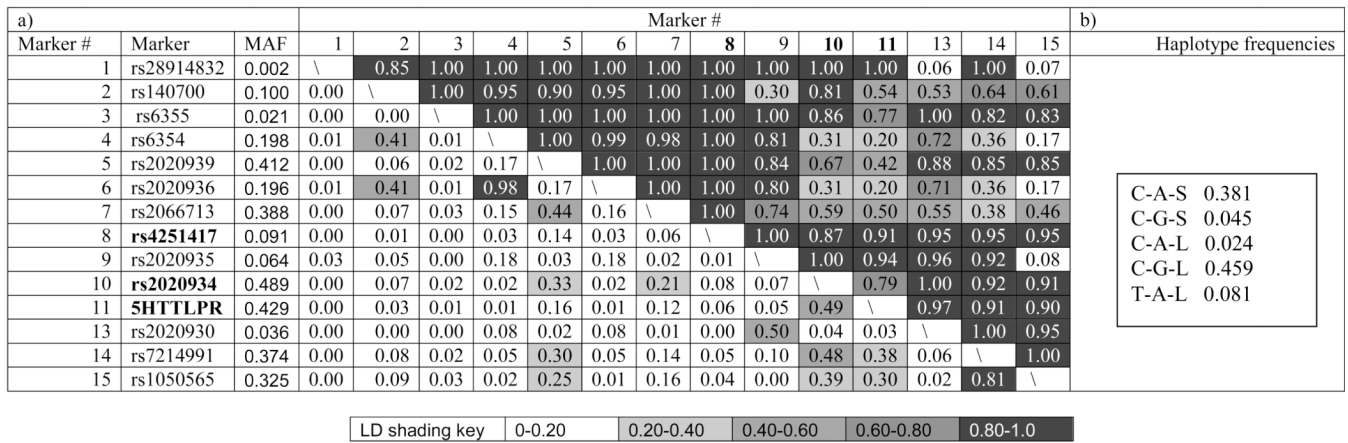
24. Holmans P, Weissman MM, Zubenko GS, Scheftner WA, Crowe RR, DePaulo JR, et al. Genetics of recurrent early-onset major depression (GenRED): Final genome scan report. American Journal of Psychiatry. 2007; 164:248–258. [PubMed: 17267787]

25. Hettema JM. What is the genetic relationship between anxiety and depression? American Journal of Medical Genetics Part C-Seminars in Medical Genetics. 2008; 148C:140–146.

26. Heath AC, Cloninger CR, Martin NG. Testing a model for the genetic structure of personality: a comparison of the personality systems of Cloninger and Eysenck. J Pers Soc Psychol. 1994; 66:762–775. [PubMed: 8189351]

27. Lake RIE, Eaves LJ, Maes HHM, Heath AC, Martin NG. Further evidence against the environmental transmission of individual differences in neuroticism from a collaborative study of 45,850 twins and relatives on two continents. Behavior Genetics. 2000; 30:223–233. [PubMed: 11105396]

28. Eysenck SBG, Eysenck HJ, Barrett P. A revised version of the Pyschoticism scale. Personality and individual differences. 1985; 6:21–30.

29. Cloninger CR, Przybeck TR, Svrakic DM. The Tridimensional Personality Questionnaire: U.S. normative data. Psychol Rep. 1991; 69:1047–1057. [PubMed: 1784653]

30. Bierut LJ, Heath AC, Bucholz KK, Dinwiddie SH, Madden PA, Statham DJ, et al. Major depressive disorder in a community-based twin sample: are there different genetic and environmental contributions for men and women? Arch Gen Psychiatry. 1999; 56:557–563. [PubMed: 10359473]

31. Bucholz KK, Cloninger CR, Dinwiddie DH, Hesselbrock VM, Nurnberger JL, Reich T, et al. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report of the reliability of the SSAGA. Journal of Studies on Alcoholism. 1994; 55:149–158.

32. American-Psychiatric-Association. Diagnostic and statistical manual of mental disorders. revised 3 ed.. Washington, DC: Author; 1987.

33. Macgregor S, Lind P, Bucholz K, Hansell N, Madden P, Richter M, et al. Associations of ADH and ALDH2 gene variation with self report alcohol reactions, consumption and dependence: an integrated analysis. Human Molecular Genetics. 2008 Online November 7 2008.

34. Composite International Diagnostic Interview version 2.1. Geneva: World Health Organization; 1997.

35. Diagnostic and Statistical Manual of Mental Disorders - Fourth edition. Washington DC: American Psychiatric Association; 1994.

36. Risch N, Zhang H. Extreme discordant sib pairs for mapping quantitative trait loci in humans. Science. 1995; 268:1584–1589. [PubMed: 7777857]

37. Purcell S, Cherny SS, Hewitt JK, Sham PC. Optimal sibship selection for genotyping in quantitative trait locus linkage analysis. Hum Hered. 2001; 52:1–13. [PubMed: 11359064]

38. Kirk KM, Birley AJ, Statham DJ, Haddon B, Lake RI, Andrews JG, et al. Anxiety and depression in twin and sib pairs extremely discordant and concordant for neuroticism: prodromus to a linkage study. Twin Res. 2000; 3:299–309. [PubMed: 11463151]

39. Birley AJ, Gillespie NA, Heath AC, Sullivan PF, Boomsma DI, martin NG. Heritability and nineteen-year stability of long and short EPQ-R neuroticism scales. Personality and individual differences. 2006; 40:737–747.

40. Wray NR, Birley AJ, Sullivan PF, Visscher PM, Martin NG. Genetic and phenotypic stability of measures of neuroticism over 22 years. Twin Research and Human Genetics. 2007 submitted.

41. Saccone SF, Pergadia ML, Loukola A, Broms U, Montgomery GW, Wang JC, et al. Genetic linkage to chromosome 22q12 for a heavy-smoking quantitative trait in two independent samples. American Journal of Human Genetics. 2007; 80:856–866. [PubMed: 17436240]

42. Hansell NK, Agrawal A, Whitfield JB, Morley KI, Zhu G, Lind PA, et al. Long-term stability and heritability of telephone interview measures of alcohol consumption and dependence. Twin Research and Human Genetics. 2008; 11:287–305. [PubMed: 18498207]

43. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Research. 1988; 16:1215. [PubMed: 3344216]

44. Heils A, Teufel A, Petri S, Stöber G, Riederer P, Bengel D, et al. Allelic variation of human serotonin transporter gene expression. J Neurochem. 1996; 66:2621–2624. [PubMed: 8632190]

45. Kim SJ, Cox N, Courchesne R, Lord C, Corsello C, Akshoomoff N, et al. Transmission disequilibrium mapping at the serotonin transporter gene (SLC6A4) region in autistic disorder. Molecular Psychiatry. 2002; 7:278–288. [PubMed: 11920155]

46. Ramoz N, Reichert JG, Corwin TE, Smith CJ, Silverman JM, Hollander E, et al. Lack of evidence for association of the serotonin transporter gene SLC6A4 with autism. Biological Psychiatry. 2006; 60:186–191. [PubMed: 16616719]

47. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21:263–265. [PubMed: 15297300]

48. de Bakker PIW, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nature Genetics. 2005; 37:1217–1223. [PubMed: 16244653]

49. Freeman MF, Tukey JW. Transformations Related to the Angular and the Square Root. Annals of Mathematical Statistics. 1950; 21:607–611.

50. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–575. [PubMed: 17701901]

51. DSM-IV. Diagnostic and Statistical Manual of Mental Disorders - Fourth edition. Washington DC: American Psychiatric Association; 1994.

52. Wray NR. Allele frequencies and the 12 measure of linkage disequilibrium: Impact on design and interpretation of association studies. Twin Research and Human Genetics. 2005; 8:87–94. [PubMed: 15901470]

53. Canli T, Lesch KP. Long story short: the serotonin transporter in emotion regulation and social cognition. Nat Neurosci. 2007; 10:1103–1109. [PubMed: 17726476]

54. Pritchard JK, Przeworski M. Linkage disequilibrium in humans: Models and data. American Journal of Human Genetics. 2001; 69:1–14. [PubMed: 11410837]

55. Mansour HA, Talkowski ME, Wood J, Pless L, Bamne M, Chowdari KV, et al. Serotonin gene polymorphisms and bipolar I disorder: focus on the serotonin transporter. Ann Med. 2005; 37:590–602. [PubMed: 16338761]

56. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007; 449:U851–U853.

57. Zaboli G, Jonsson EG, Gizatullin R, De Franciscis A, Asberg M, Leopardi R. Haplotype analysis confirms association of the serotonin transporter (5-HTT) gene with schizophrenia but not with major depression. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics. 2008; 147B:301–307.

58. Strug LJ, Suresh R, Fyer AJ, Talati A, Adams PB, Li W, et al. Panic disorder is associated with the serotonin transporter gene (SLC6A4) but not the promoter region (5-HTTLPR). Mol Psychiatry. 2008
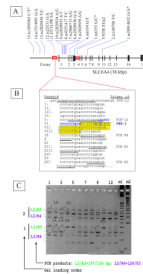
Figure 1.
SLC6A4 (5HTT) gene locus, polymorphisms and assays. (A) Genomic locus map showing exons 1–15 (solid & open boxes; open boxes are exons transcribed only in some splicing variants), the Length Polymorphic Region (5HTTLPR; striped box), the VNTR (STin2, striped box) and SNPs (rs#) with (minor allele listed first). * indicates a non-synonymous exonic SNP. The exon/intron numbering is based on the most recent Ensembl update (SLC6A4-001; OTTHUMT00000256114, June, 2007). (B) Sequence of the 5HTTLPR with 16 copies of repeat unit plus flanking unique genomic sequence. The shaded and boxed sequence (of repeats VI–VIII) indicates the 43bp deleted in the short (S) version. The dash in the sequence cc-ccc indicates the removal of one c residue according to recent resequencing data. Within the deleted sequence is an [a/g] SNP (rs25531). Underlined sequences show primers used to obtain various PCR products; primer names are shown to the right. The sequence PEX-1 (bold italics) is used in a special primer-extension reaction to assay rs25531 as explained in the text and Supplement 1. (C) A 3.5% agarose gel with 4-fold multiplexed samples: DNAs subjected to PCR with either L3/R3 or L3/R4 primers. DNA samples from each amplification were pooled in pairs and loaded on to 24 lanes of the gel (half is displayed) in two loadings separated in time, so that 96 DNAs were genotyped on one gel. The sizes of the Long and Short alleles are shown at bottom (primers have 10nt 5'-extensions, not shown, which add 20 bp to product size); m1, m2=marker lanes (50 bp ladder) for each loading. Representative examples of genotypes are: LS (het, note the weak L band) Lane-1 Loading-2 both L3R3 and L3R4. LL: Lane-1 Loading-1 both L3R3 and L3R4; SS: Lane-3 Loading-1 L3R3; Lane-9 Loading-1 L3R3.

a)									Marker #							b)
Marker #	Marker	MAF	1	2	3	4	5	6	7	**8**	9	**10**	**11**	13	14	15	Haplotype frequencies
1	rs28914832	0.002	\	0.85	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.06	1.00	0.07	
2	rs140700	0.100	0.00	\	1.00	0.95	0.90	0.95	1.00	1.00	0.30	0.81	0.54	0.53	0.64	0.61	
3	rs6355	0.021	0.00	0.00	\	1.00	1.00	1.00	1.00	1.00	1.00	0.86	0.77	1.00	0.82	0.83	
4	rs6354	0.198	0.01	0.41	0.01	\	1.00	0.99	0.98	1.00	0.81	0.31	0.20	0.72	0.36	0.17	
5	rs2020939	0.412	0.00	0.06	0.02	0.17	\	1.00	1.00	1.00	0.84	0.67	0.42	0.88	0.85	0.85	
6	rs2020936	0.196	0.01	0.41	0.01	0.98	0.17	\	1.00	1.00	0.80	0.31	0.20	0.71	0.36	0.17	
7	rs2066713	0.388	0.00	0.07	0.03	0.15	0.44	0.16	\	1.00	0.74	0.59	0.50	0.55	0.38	0.46	
8	**rs4251417**	0.091	0.00	0.01	0.00	0.03	0.14	0.03	0.06	\	1.00	0.87	0.91	0.95	0.95	0.95	
9	rs2020935	0.064	0.03	0.05	0.00	0.18	0.03	0.18	0.02	0.01	\	1.00	0.94	0.96	0.92	0.08	
10	**rs2020934**	0.489	0.00	0.07	0.02	0.02	0.33	0.02	0.21	0.08	0.07	\	0.79	1.00	0.92	0.91	
11	**5HTTLPR**	0.429	0.00	0.03	0.01	0.01	0.16	0.01	0.12	0.06	0.05	0.49	\	0.97	0.91	0.90	
13	rs2020930	0.036	0.00	0.00	0.00	0.08	0.02	0.08	0.01	0.00	0.50	0.04	0.03	\	1.00	0.95	
14	rs7214991	0.374	0.00	0.08	0.02	0.05	0.30	0.05	0.14	0.05	0.10	0.48	0.38	0.06	\	1.00	
15	rs1050565	0.325	0.00	0.09	0.03	0.02	0.25	0.01	0.16	0.04	0.00	0.39	0.30	0.02	0.81	\	

Haplotype frequencies (b):
C-A-S 0.381
C-G-S 0.045
C-A-L 0.024
C-G-L 0.459
T-A-L 0.081

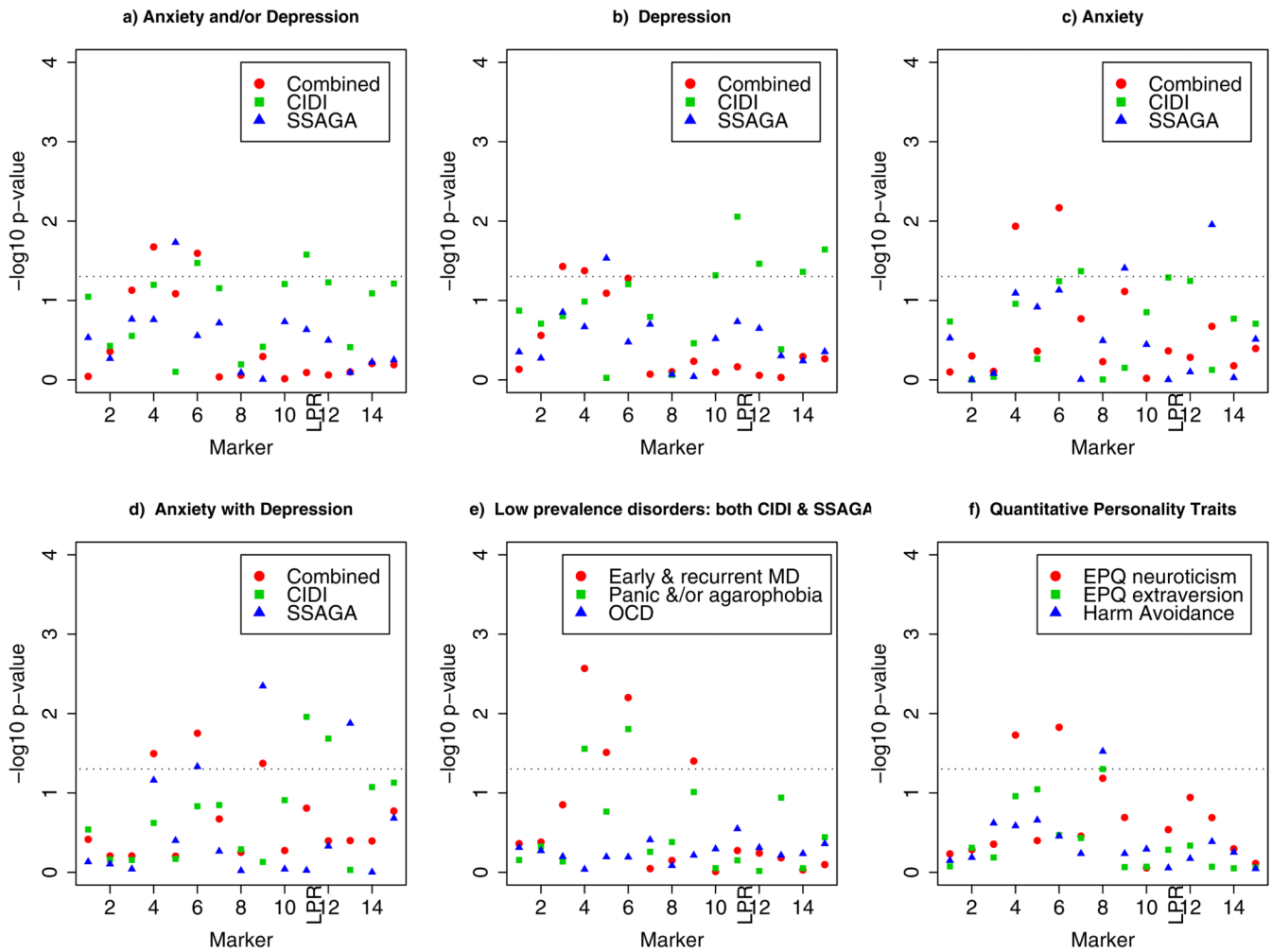| LD shading key	0-0.20	0.20-0.40	0.40-0.60	0.60-0.80	0.80-1.0

Figure 2.

a)Allele frequencies and linkage disequilibrium of genotyped markers. $|D'|$ above the diagonal (\) and r^2 below the diagonal.

b) Haplotype frequencies of rs4251417-rs2020934-5HTTLPR

Figure 3.
Association analysis results. Above the dotted horizontal line represents p < 0.05. Marker numbers as in Figure 2. a–d, results for anxiety and depression disorders with graphs representing a different balance of phenotypic homogeneity and sample size. For each SNP, when the circle representing the combined data set analysis is at the top (highest −log$_{10}$(p-value)) of the 3 symbols, then both the SSAGA and CIDI samples show a trend in association in the same direction. In contrast, when the combined analysis result has the lowest −log$_{10}$(p-value) of the three analyses, the associations in the SSAGA and CIDI samples are with different alleles.
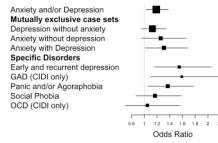
Figure 4.
Forrest plot of odds ratio and their 95% confidence intervals for different case groups for rs6354 showing increased frequency of allele T in cases than controls (frequency in controls 0.78).
NB. The case groups are combined from both the SSAGA and CIDI questionnaires and reflect the diagnostic classes achievable from them.

Table 1

Summary of meta-analyses of association studies of 5HTTLPR in major depression, anxiety disorders and neuroticism

Phenotype	Comparison	Association (95% CI)	No. Studies	Description	Reference
Major Depression	S vs L SS vs LL	OR = 1.11 (1.04–1.19) OR = 1.39 (1.21–1.61)	24	3752 cases 5707 controls	(11)
Neuroticism	SS or SL vs LL	d = −0.22 (−0.44 − −0.01)	12	3247 individuals	(12)
Neuroticism	S vs L	NS: p = 0.81 NS : p = 0.47	1[a]	768 individuals selected from the 5% tails of the neuroticism distribution from a population sample of 88,142 4175 selected for extreme neuroticism from population sample of 20,921	(13)
Panic Disorder	S vs L	OR = 0.91 (0.80–1.03)	10	1025 cases 1568controls	(14)
OCD	SS vs LL SL vs LL S vs L	OR = 1.21 (1.10, 1.45) OR = 0.79 (0.64, 0.98) OR = 1.01 (0.88, 1.03)	13	1242 cases 2203 controls	(15)
Harm Avoidance	SS or SL vs LL	d = −0.05 (−0.66 − 0.56)	13	2521 individuals	(12)

[a]Large single study published subsequent to the meta-analysis of neuroticism.

Abbreviations: CI: Confidence interval; S, Short allele; L, long allele, OR: odds ratio ; d standardized mean difference; NS = not significant OCD: Obsessive Compulsive Disorder

Table 2

Number of genotyped individuals contributing to the association analyses.

Diagnosis (indenting denotes subsetting)	Diagnoses identified from the SSAGA interview administered between 1992 and 2000 (or 2001–2005) and the CIDI interview administered between 1996 and 1999.			Quantitative personality traits measured in questionnaires administered between 1988 and 1989.	
	SSAGA	CIDI	Combined	Trait	Unrelated Individuals
control	724	327	1051[a]	EPQ-Neuroticism	2651
Anxiety and/or Depression	691	470	1161[a]	EPQ-Extraversion	2643
Anxiety with Depression[b]	81	184	265	Harm Avoidance	2645
Depression	587	366	953		
Recurrent early onset	83	129	212		
Depression without Anxiety[b]	506	182	688		
Anxiety	185	288	473		
Generalised Anxiety Disorder		129	129		
Social Phobia	58	105	163		
Panic/Agoraphobia	141	108	249		
Obsessive Compulsive Disorder		79	79		
Anxiety without Depression[b]	104	104	208		

[a] A total data set of 1051 controls and 1161 cases, all unrelated.

[b] Mutually exclusive case sets, so numbers sum to Anxiety and/or Depression cases. Note that these classifications are limited by the diagnoses that could be allocated by the questionnaires.

Table 3

Description of the diagnosis classes in terms of the standardised personality measures, age and gender in independent SSAGA and CIDI subsets.

| | Personality questionnaires administered between 1988–1989 | | | | | | | | | age | | Gender |
| | EPQ-Neuroticism | | | EPQ-Extraversion | | | Harm Avoidance | | | | | |
	N	mean	sd	N	mean	sd	N	mean	sd	age mean	age sd	%female
SSAGA interview administered between 1992 and 2000 (or 2001–2005)												
control	724	−0.82	0.53	711	0.24	0.97	714	−0.55	0.85	41.9	12.4	64.7
Anxiety and/or Depression	618	0.31	1.00	617	0.02	1.05	613	0.17	1.01	38.9	10.1	70.2
Anxiety with Depression	73	0.83	1.08	73	0.02	1.11	73	0.54	1.11	39.0	8.4	78.1
Anxiety	164	0.63	1.13	164	−0.11	1.07	163	0.47	1.10	40.3	10.8	75.0
Depression	527	0.28	0.98	526	0.07	1.04	523	0.13	0.99	38.4	9.6	69.8
CIDI interview administered between 1996 and 1999												
control	321	−1.45	0.44	318	0.23	0.94	321	−0.83	0.76	34.9	11.9	61.4
Anxiety and/or Depression	449	0.82	1.34	440	−0.09	1.07	445	0.55	1.17	35.1	10.5	65.5
Anxiety with Depression	178	1.12	1.13	174	−0.31	1.03	176	0.87	1.02	35.7	10.8	70.2
Anxiety	274	1.01	1.20	268	−0.22	1.05	261	0.76	1.06	35.7	10.9	64.6
Depression	353	0.82	1.35	346	−0.10	1.07	350	0.55	1.19	34.9	10.3	68.6