Sequence Variants of the Brain-Derived Neurotrophic Factor (BDNF) Gene Are Strongly Associated with Obsessive-Compulsive Disorder

Diana Hall,1 Alefiya Dhilla,2 Anna Charalambous,1 Joseph A. Gogos,2 and Maria Karayiorgou1

1The Rockefeller University, Human Neurogenetics Laboratory, and 2Columbia University, College of Physicians & Surgeons, Department of Physiology and Cellular Biophysics and Center for Neurobiology & Behavior, New York

We evaluated a possible association between the brain-derived neurotrophic factor (BDNF) gene and susceptibility to obsessive-compulsive disorder (OCD) by genotyping a number of single-nucleotide polymorphisms (SNPs) and one microsatellite marker from the extended BDNF locus in 164 triads with OCD. Extensive background linkage disequilibrium was observed at this locus. Single-locus transmission-distortion tests revealed significant evidence of association with the disease for all the BDNF gene markers tested, including a Val66Met variation affecting the sequence of the proBDNF protein. Analysis of multi-SNP haplotypes provided similar results. Haplotype transmission comparisons in this and previous studies point to a functionally distinct BDNF haplotype uniquely marked by the rare Met66 allele, which is undertransmitted and likely confers a protective effect in OCD and other psychiatric disorders.
Figure 1  

**a.** Exon-intron structure in the BDNF locus. The relative position of the markers used in this study is indicated by arrowheads. **b.** Pro-BDNF. Arrowheads indicate known protease cleavage sites involved in processing of the mature form, as well as of the secreted proform. The relative position of the Val66Met variation is also indicated.
SNPs (SNP 1 [rs988748] and SNP 2 [rs2049046]) located in the region immediately upstream of exon 3 (fig. 1a). Second, we tested one SNP (SNP 3 [rs6265]) that determines a valine–methionine substitution at amino acid position 66 of the preprotein (fig. 1b), which is highly conserved across species. Finally, in the course of our work, Sklar et al. (2002) and Neves-Pereira et al. (2002) reported independently an association between various SNPs and one microsatellite from the BDNF locus and bipolar disorder. To facilitate cross referencing with these studies, we chose to genotype one SNP from the Sklar et al. study (a20, coded as “SNP 4” in our study, located in the 3′-UTR region [fig. 1a]), as well as the microsatellite marker from the Neves-Pereira et al. study (referred to as “BDNF(CA),” located in intron 5, immediately upstream of coding exon 6 [fig. 1a]). Each SNP was genotyped by use of an allele-specific restriction-site assay. Primers for all SNPs were designed from the sequence available at the UCSC site. Samples were digested for 3 h with the enzymes BsmAI, BstAI, NlaIII, and Eco01091 for SNPs 1, 2, 3, and 4, respectively.

The proband sample consisted of 164 individuals (78 males, 86 females) who met lifetime criteria for OCD, according to the Diagnostic and Statistical Manual, 4th edition (DSM-IV) (American Psychiatric Association 1994). Of the sample, ∼85% were recruited through advocate newspapers, with the remaining 15% referred by local treating clinicians. The study was approved by the Rockefeller University institutional review board, and all subjects provided written informed consent prior to their participation in the study. Subjects or the parents of minors were interviewed by specially trained doctoral-candidate clinicians, who used the Diagnostic Interview for Genetic Studies (DIGS) (Nurnberger et al. 1994). To ensure the accurate detection of DSM-IV OCD, detailed screen, symptom, and impairment questions ascertaining all DSM-IV OCD criteria were added, as described in detail elsewhere (Sobin et al. 1999). Patients with OCD onset following the onset of another psychiatric disorder (e.g., major depressive disorder [MDD]) were not excluded from this study. However, patients with multiple psychiatric disorders whose DSM-IV OCD was never the focus of treatment or never the cause of primary distress were excluded from these analyses. Analyses of demographic and clinical characteristics suggested that the sample is quite heterogeneous and represents a wide range of onset ages, severity levels, and clinical phenomenology (Sobin et al. 1999). To account for the clinical heterogeneity of our sample, as well as for the age-dependent effects of BDNF-deficiency observed in animal models, our analysis included two age-at-onset liability classes: patients with childhood-onset (CO) (<18 years) OCD and patients with adult-onset (AO) (≥18 years) OCD. The CO sample included 38 children with DSM-IV OCD (mean interview age 12.4 ± 3.3 years; 82% >10 years; 50% male) and 82 adults, age ≥19 years, who reported OCD onset prior to age 18 years (mean interview age 31.2 ± 8.8 years; 46% male). The AO sample included 44 adults, age ≥19 years, who reported OCD onset at age ≥18 years (mean interview age 25.3 ± 8.0 years; 48% male). A comparison of the phenotypic characteristics of the CO and AO subsets of patients found that CO OCD was characterized by a relatively aggressive course and a greater number of obsessions and compulsions, unrelated to the amount of time in illness, which indicates that CO OCD is a more severe form of illness (Sobin et al. 2000). The two groups do not have any statistically significant differences in the frequency of comorbid disorders, including other anxiety disorders, bipolar disorder, phobias, eating disorders, tics, or Tourette syndrome (although a trend for an increased prevalence of eating disorders, tics, and Tourette syndrome in the CO sample was observed) (data not shown).

The single-locus test for excess of transmission of alleles from heterozygous parents to affected offspring (transmission/disequilibrium test [TDT] [Spielman et al. 1993]) revealed significant evidence of association to the disease for all the markers tested (table 1). The strongest evidence was obtained at SNPs 1, 3, and 4 for the total sample (p = .0005). Two-locus tests performed by use of a likelihood-based method, from the program TRANSMIT v2.5.2 (Clayton 1999), provided similar results. It is interesting that a highly significant association was observed only when the CO subset or the total sample was considered. By contrast, none of the loci analyzed were significant at the P = .05 level in the subset of families with AO. Although not statistically significant, the transmissions of the alleles in the families with AO revealed identical distortions to those observed in the sample with CO, therefore contributing to the overall highly significant association in the entire sample. Our inability to observe statistical significance could be due to the overall effect of the allele being modest in the subset with AO or to the relatively small sample size of the subset with AO. Larger samples may be required to obtain conclusive results.

We also observed significant evidence for association at the multiallelic marker BDNF(CA), using the program ETDT v1.9 (Sham and Curtis 1995) (table 2). Three alleles showed significant transmission distortions; allele 2 and allele 5 appeared to be undertransmitted, whereas allele 3 is overtransmitted; the allelicwise global test calculated over 6 df showed P = .009 in the entire sample, with the subset with CO demonstrating the most striking transmission distortion. Comparison with the Neves-Pereira et al. (2002) study shows that the same allele is overtransmitted in both studies.

To investigate further the presence of a specific allelic combination associated with an increased risk of OCD, we performed multi-SNP haplotype analysis, using
Table 1

<table>
<thead>
<tr>
<th>SNP LOCATIONa</th>
<th>Np</th>
<th>AAa</th>
<th>Single Locus (n = 120)</th>
<th>Two Loci</th>
<th>Single Locus (n = 44)</th>
<th>Two Loci</th>
<th>ENTIRE SAMPLE (n = 164)</th>
<th>Single Locus (95% CI)d</th>
<th>Two Loci</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>−44645</td>
<td>C:G</td>
<td>nc</td>
<td>0.003i</td>
<td>.08i</td>
<td>0.005; .50 (.34–.74)</td>
<td>.004k</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>−43675</td>
<td>A:T</td>
<td>nc</td>
<td>0.02b</td>
<td>.02c</td>
<td>0.012; .66 (.48–.90)</td>
<td>.001l</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>196</td>
<td>G:A</td>
<td>Val/Met</td>
<td>.001k</td>
<td>.009ma</td>
<td>.08m</td>
<td>.005; .44 (.26–.74)</td>
<td>.004ma</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>11757</td>
<td>C:G</td>
<td>nc</td>
<td>0.0025p</td>
<td>.08p</td>
<td>.005; .51 (.34–.75)</td>
<td>.004pa</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE.—Significant associations are shown in italics.

a SNP location calculated from the ATG for BDNF.
b Np nucleotide change.
c AA = amino acid substitution; nc = noncoding region.
d Approximate relative risk; data based on distribution of alleles on transmitted and nontransmitted chromosomes.
e C:G transmission ratios 51:25.
f C:G transmission ratios 21:11.
g Two-locus analysis (SNP 1 and SNP 2).
h T:A transmission ratios 76:50.
i T:A transmission ratios 26:19.
j Two-locus analysis (SNP 2 and SNP 3).
k G:A transmission ratios 42:17.
m Two-locus analysis (SNP 3 and SNP 4).

TRANSMIT v2.5.2 (Clayton 1999) (table 3). To avoid very small haplotype frequencies, our analysis considered all the families and included only the diallelic polymorphisms. The BDNF gene shows limited haplotype diversity over a region of 56 kb, which is likely adequately captured by the SNPs genotyped here. Specifically, in our data set, we observed 10 haplotypes, of which only 4 were present with probabilities >3% and accounted for the vast majority of the haplotype diversity (98%). Of these, haplotype 2 is significantly overtransmitted to the probands with OCD, and haplotype 4 is undertransmitted. The allelic composition of the haplotypes reflects the single-locus results with a preferentially transmitted CTGC haplotype (P = .007) and an undertransmitted GAAG haplotype (P = .022). The global P value is .003 for the four most common haplotypes and .022 for all possible haplotypes (table 3). It is notable that all four psychiatric disease data sets tested to date for association with BDNF (the OCD data set tested by the current study, two bipolar data sets tested by Sklar et al. [2002], and one bipolar data set tested by Neves-Pereira et al. [2002]) share undertransmission of haplotype 4, which is uniquely marked by the rare A allele of SNP 3 (which results in methionine at codon 66 and was typed in all studies), as well as by allele G of SNP 4 (typed in the current study and the study of Sklar et al. [2002]). On the other hand, haplotype overtransmission is less consistent among the various data sets. This may reflect the presence of more than one functionally equivalent “risk”—or even “neutral”—haplotypes versus one functionally distinct “protective” haplotype that is marked by the rare alleles of SNPs 3 and 4. In that respect, it is of interest that age-at-onset comparisons among the 164

Table 2

<table>
<thead>
<tr>
<th>SAMPLE AND TRANSMISSIONS</th>
<th>CO (n = 120)</th>
<th>AO (n = 44)</th>
<th>ENTIRE SAMPLE (n = 164)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmissionsa</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>CO:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>1</td>
<td>4</td>
<td>44c</td>
</tr>
<tr>
<td>nT</td>
<td>3</td>
<td>11</td>
<td>16c</td>
</tr>
<tr>
<td>AO:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>0</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>nT</td>
<td>1</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Entire Sample:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>1</td>
<td>6e</td>
<td>58e</td>
</tr>
<tr>
<td>nT</td>
<td>4</td>
<td>16e</td>
<td>27e</td>
</tr>
</tbody>
</table>

NOTE.—Significant associations are shown in italics.
a T = transmitted; nT = nontransmitted.
b Allelewise global P value.
c P = .0003.
d P = .0173.
e P = .0008.
f P = .0008.
g P = .0291.
Table 3

Haplotype Transmission Analysis

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>SNP 1</th>
<th>SNP 2</th>
<th>SNP 3</th>
<th>SNP 4</th>
<th>Estimated Probability</th>
<th>O&lt;sup&gt;a&lt;/sup&gt;</th>
<th>E&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Var(O-E)&lt;sup&gt;c&lt;/sup&gt;</th>
<th>x&lt;sup&gt;2&lt;/sup&gt; (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>C</td>
<td>.308</td>
<td>101.38</td>
<td>99.94</td>
<td>38.018</td>
<td>.054 (.816)</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>.467</td>
<td>169.09</td>
<td>151.41</td>
<td>42.525</td>
<td>7.346 (.007)</td>
</tr>
<tr>
<td>3</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.042</td>
<td>11.71</td>
<td>13.73</td>
<td>6.500</td>
<td>.628 (.428)</td>
</tr>
<tr>
<td>4</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>.162</td>
<td>38.78</td>
<td>52.35</td>
<td>19.732</td>
<td>9.332 (.002)</td>
</tr>
</tbody>
</table>

For 4 haplotypes more frequent than .03 (df 4) 15.814 (.003)

Total haplotypes (df 9) 19.463 (.022)

<sup>a</sup> O = observed transmissions of haplotype to affected offspring.
<sup>b</sup> E = expected transmission under Mendelian inheritance.
<sup>c</sup> Var(O-E) = variance of (O-E).
<sup>d</sup> P value calculated for 1 df.

The probands used in the current study indicate that presence of at least one copy of the “protective” haplotype delays onset of the disease by ∼3 years (from 13.2 years to 16.5 years; P = .03).

The limited haplotypic diversity observed over the BDNF locus indicates the presence of high-level linkage disequilibrium (LD) among the transmitted chromosomes. The extent of marker-to-marker disequilibrium is an important factor in interpreting marker-disease association results. We therefore have genotyped six additional markers (rs3763965, rs2352802, rs972096, rs1387145, rs1464896, and rs2140887) (fig. 1a) in the same families and have performed an extensive LD analysis on the nontransmitted chromosomes to investigate the extent of background LD in the BDNF gene and its vicinity (157 kb upstream and 198 kb downstream of the gene). The standard Lewontin D′ (Lewontin 1988) and the square of the correlation coefficient (r<sup>2</sup> or Δ<sup>2</sup>) between adjacent markers was calculated from the program UNPHASED, by use of the EM function for the estimation of uncertain haplotype frequencies, via an EM algorithm (table 4). Both measures range 0–1; however, for the r<sup>2</sup> measure, unless the allele frequencies at the two loci are identical, perfect correlation cannot be achieved. Extremely strong LD was observed between SNPs located within the BDNF gene, as also reported by Sklar et al. (2002). In addition, extremely strong LD was observed between SNPs from the BDNF gene and SNPs located upstream of the BDNF gene, but not ones located downstream (table 4). It should be noted that single-locus analysis for the flanking SNPs did not show any evidence for transmission distortion, with the exception of rs972096 (P = .0002), which is also the marker most highly correlated (r<sup>2</sup>) to the core risk haplotype.

It has been suggested that a contribution of BDNF to vulnerability to depression may be at the core of the recently observed association with bipolar disorder (Lincinio and Wong 2002; Neves-Pereira et al. 2002). Most of the work, to date, that supports such a role of BDNF has been derived from studies that attempt to model depression in animals (with perhaps most intriguing reports that intracerebral administration of BDNF to animals may have antidepressant properties) (Siuciak et al. 1997). However, despite the deficits in the 5-HT system, heterozygous Bdnf<sup>+/H11001</sup>/H11002<sup>+</sup> mice were not more likely to display anxious or depressive-like behaviors (MacQueen et al. 2001). The most frequently occurring comorbid illnesses in patients with OCD are, by far, mood disorders.

Table 4

Background LD in the BDNF Gene and Neighboring Regions

<table>
<thead>
<tr>
<th>Distance (kb)</th>
<th>Frequency&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Marker</th>
<th>rs3763965</th>
<th>rs2352802</th>
<th>rs972096</th>
<th>rs1387145</th>
<th>SNP 1</th>
<th>SNP 2</th>
<th>SNP 3</th>
<th>SNP 4</th>
<th>rs1464896</th>
<th>rs2140887</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>47</td>
<td>rs3763965</td>
<td>.24</td>
<td>.33</td>
<td>.93</td>
<td>.27</td>
<td>.45</td>
<td>.20</td>
<td>.24</td>
<td>.02</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>4.5</td>
<td>22</td>
<td>rs2352802</td>
<td>1.00</td>
<td>.08</td>
<td>.25</td>
<td>.04</td>
<td>.07</td>
<td>.05</td>
<td>.07</td>
<td>.00</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>23</td>
<td>rs972096</td>
<td>.85</td>
<td>.87</td>
<td>.37</td>
<td>.58</td>
<td>.24</td>
<td>.49</td>
<td>.66</td>
<td>.02</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>47</td>
<td>rs1387145</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>.27</td>
<td>.49</td>
<td>.19</td>
<td>.26</td>
<td>.02</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>28</td>
<td>SNP 1</td>
<td>.80</td>
<td>.72</td>
<td>.79</td>
<td>.89</td>
<td>.26</td>
<td>.65</td>
<td>.87</td>
<td>.04</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>42</td>
<td>SNP 2</td>
<td>.82</td>
<td>.66</td>
<td>.89</td>
<td>.82</td>
<td>.97</td>
<td>.16</td>
<td>.23</td>
<td>.01</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>22</td>
<td>SNP 3</td>
<td>.80</td>
<td>.83</td>
<td>.86</td>
<td>.88</td>
<td>.94</td>
<td>.89</td>
<td>.77</td>
<td>.03</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>27</td>
<td>SNP 4</td>
<td>.79</td>
<td>.86</td>
<td>.86</td>
<td>.88</td>
<td>.95</td>
<td>.91</td>
<td>1.00</td>
<td>.04</td>
<td>.00</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Percent of minimum allele frequency calculated from the untransmitted parental chromosomes in the total data set.

Note.—For each pair of markers, the standardized D′ is shown below the diagonal, and the r<sup>2</sup> is shown above the diagonal. D′ values >0.6 are shown in italics.
and, in particular, DSM-IV MDD. Additional mood disorders found in patients with OCD include dysthymia, bipolar I, and bipolar II (Rasmussen and Eisen 1992; Sobin et al. 1999). When patients with OCD are stratified according to the presence or absence of comorbid mood disorders, the statistical evidence weakens in accord with the decrease in the sample size, which excludes comorbid mood disorders as the source of association between BDNF and OCD (data not shown). Therefore, this analysis fails to provide additional support for the speculated BDNF/depression connection. It is, of course, conceivable that partial disruptions in BDNF expression or function may lead to perturbations in central neurotransmission that, depending upon the presence of additional genetic liability, can manifest as OCD, bipolar disorder, MDD, or other related disorders. Alternatively, BDNF-modulated traits or neuropsychological abnormalities other than depression may need to be explored as a likely common basis for the observed association of the BDNF gene with bipolar disorder and OCD.

The nature of protective variation in the BDNF locus and the signaling pathways affected by it need to be identified and explored as a candidate pharmacotherapy target. In light of the extended background LD in this locus, this may require functional studies in animal and cellular models. The presence of strong LD upstream of the BDNF gene does not allow exclusion of other genes that map in that area. For example, the human homolog of mouse Lin-7c is located only 667 bp from rs3763965 (fig. 1) and is an interesting candidate, on the basis of its sequence homology to PSD-95, a gene involved in NMDA receptor–synaptic localization and signaling. Nevertheless, the clustering of positively associated SNPs with variable allele frequencies within the BDNF gene strongly implicates this gene as the source of the association signal from this chromosomal region. Within the gene, the Val66Met substitution seems to be a good candidate for a causative variant, either by affecting the processing of the mature form (thus modulating the growth-promoting Trk pathway) or by affecting the interaction of the secreted and the extracellularly processed pre-pro form with p75 NTR, thus modulating apoptotic signaling through this alternate pathway. Indeed, recent work, which needs to be replicated, suggests that the Val/Met polymorphism impacts intracellular trafficking and activity-dependent secretion of BDNF (Egan et al. 2003). Additional sequencing and functional studies need to address whether other SNPs that affect the expression (activity-dependent or constitutive) or splicing of the BDNF gene may also be implicated.

Acknowledgments

The authors thank all the families who participated in the study. Also, the authors thank Laura Lundy for her help with the phenotypic characterization of the clinical sample. Support was provided, in part, by the Irma Hirschl Trust (to M.K.) and the New York Council Speaker’s Fund (to J.A.G.). Clinical work at the Rockefeller University Hospital General Clinical Research Center was supported by grant M01 RR00102.

Electronic-Database Information

URLs for data presented herein are as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for OCD)
UCSC Genome Bioinformatics, http://genome.ucsc.edu/

References

Licioini J, Wong ML (2002) Brain-derived neurotrophic factor (BDNF) in stress and affective disorders. Mol Psychiatry 7:519