Molecular Psychiatry (2003) 8, 546–557

www.nature.com/mp

ORIGINAL RESEARCH ARTICLE

Evidence that a single nucleotide polymorphism in the promoter of the G protein receptor kinase 3 gene is associated with bipolar disorder

TB Barrett[1,2], RL Hauger[1,2], JL Kennedy[3], AD Sadovnick[4], RA Remick[5], PE Keck[6], SL McElroy[6], M Alexander[1,2], SH Shaw[1,7] and JR Kelsoe[1,2]

[1]*Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA;* [2]*Department of Psychiatry, San Diego VA Healthcare System, La Jolla, CA, USA;* [3]*Department of Psychiatry, University of Toronto, Toronto, Ont., Canada;* [4]*Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada;* [5]*Department of Psychiatry, St Paul's Hospital, Vancouver, BC, Canada;* [6]*Department of Psychiatry, University of Cincinnati, Cincinnati, OH, USA;* [7]*Illumina, Inc., San Diego, CA, USA*

In a genome-wide linkage survey, we have previously shown evidence suggesting that the chromosome 22q12 region contains a susceptibility locus for bipolar disorder (BPD). Two independent family sets yielded lod scores suggestive of linkage at markers in this region near the gene G protein receptor kinase 3 (GRK3). GRK3 is an excellent candidate risk gene for BPD since GRK3 is expressed widely in the brain, and since GRKs play key roles in the homologous desensitization of G protein-coupled receptor signaling. We have also previously shown GRK3 expression to be induced by amphetamine in an animal model of mania using microarray-based expression profiling. To identify possible functional mutations in GRK3, we sequenced the putative promoter region, all 21 exons, and intronic sequence flanking each exon, in 14–22 individuals with BPD. We found six sequence variants in the 5′-UTR/promoter region, but no coding or obvious splice variants. Transmission disequilibrium analyses of one set of 153 families indicated that two of the 5′-UTR/promoter variants are associated with BPD in families of northern European Caucasian ancestry. A supportive trend towards association to one of these two variants (P-5) was then subsequently obtained in an independent sample of 237 families. In the combined sample, the P-5 variant had an estimated allele frequency of 3% in bipolar subjects, and displayed a transmission to non-transmission ratio of 26 : 7.7 ($\chi^2 = 9.6$, one-sided *P* value = 0.0019). Altogether, these data support the hypothesis that a dysregulation in GRK3 expression alters signaling desensitization, and thereby predisposes to the development of BPD.
Molecular Psychiatry (2003) **8,** 546–557. doi:10.1038/sj.mp.4001268

Keywords: Bipolar disorder; Mania; Depression; Linkage disequilibrium; Transmission disequilibrium test; SNP; Chromosome 22; Genetic association; G protein receptor kinases; Receptor desensitization

Bipolar disorder (BPD) is a severe psychiatric disorder that affects approximately 1% of the world's population. A variety of family epidemiological data have long argued for a strong genetic component. Twin studies have generally reported a concordance rate in monozygotic twins of approximately 60% compared to a rate in dizygotic twins of 15%.[1] First-degree relatives of probands have a 7% risk for BPD, and a two-fold increase in risk for unipolar disorder.[2–4] While the mode of inheritance remains unclear, segregation analyses of family data have generally suggested a complex mode of inheritance, although several recent large samples have reported evidence for autosomal dominant major gene effects.[3,5]

More than 10 genome-wide linkage scans have been published, and many additional scans which are in progress have reported some findings for selected chromosomal regions.[6] A number of chromosomal loci have now been reported in multiple studies including: 4p, 5q, 10q, 12q, 13q, 16p, 18p, 18q, 21q, 22q. Some of these regions have met the proposed rigorous criteria for statistical significance and replication.[6,7] A meta-analysis of 11 whole-genome linkage scans found the strongest evidence for susceptibility loci for BPD on 13q and 22q.[8]

We first reported evidence for linkage to the proximal portion of 22q in a study of 13 families.[9] A maximum lod score of 2.5 was obtained at D22S303 (position (pos.) 20.1 Mb per NCBI). We subsequently expanded this set to 20 families and reported a

Correspondence: JR Kelsoe, MD, Department of Psychiatry, 0603, University California, San Diego, La Jolla, CA 92093-0603, E-mail: jkelsoe@ucsd.edu.
Received 3 April 2002; revised 10 July 2002; accepted 13 August 2002

genome-wide maximum lod score of 3.8 at D22S278 (pos. 33.2 Mb) at 22q13.[10] However, lod scores $\geq$1.0 spanned a region of 32 cM (according to the Marshfield sex-averaged map) covering most of 22q, and appeared to include possibly two distinct regions for which there was suggestive evidence of linkage. The second putative peak in lod scores was located more proximally at 22q12.1, where a maximum lod score of 2.2 was obtained at D22S419 (pos. 22.7 Mb). The genetic and physical distances between these markers are consistent with a reported increased rate in recombination on 22q such that the 10.5 Mb between D22S419 and D22S278 is equivalent to 15 cM (Sanger Centre web site, Marshfield Map). In a recent follow-up study of a second independent set of 32 families (to be reported separately), we obtained the maximum evidence for linkage on chromosome 22 at precisely this same 22q12.1 locus. The combined analysis of these 52 families produced a peak NPL score for chromosome 22 of 2.45 at D22S925 (pos. 22.5 Mb) with an apparent 1-lod support interval centered at this marker and spanning 8–10 cM.

Support for this region also comes from the NIMH Genetics Initiative for Bipolar Disorder consortium.[11] In a study of 97 families, they reported a peak lod score for chromosome 22 of 2.5 at the marker D22S533 (pos. 22.6 Mb). Under the disease model that gave the maximum score at D22S533, lod scores of 0.3 and 0.7 were obtained at flanking markers $\sim$8 cM centromeric and $\sim$7 cM telomeric to D22S533. These studies on three independent family sets (and a total of 149 families) thus support a distinct peak in evidence for linkage in the 22q12.1 region.

An attractive candidate gene in this region is G protein receptor kinase 3 (GRK3) located just 40 kb telomeric to the marker D22S419. In the continuing presence of high agonist concentrations, signaling of a G protein-coupled receptor (GPCR) is rapidly terminated via a process termed homologous desensitization. During the initial phase of homologous desensitization, a G protein receptor kinase (GRK) phosphorylates serines and/or threonines in the activated GPCR's C-terminus or in the third intracellular loop.[12–16] There are only seven known members of the GRK family of genes, but hundreds of neurotransmitter GPCRs. GRK3 is widely expressed in the brain including regions of the limbic system[17,18] and has been demonstrated to phosphorylate multiple GPCRs critical for brain neurotransmission including adrenergic, muscarinic, histaminic, dopamine, and corticotropin-releasing factor (CRF) receptors.[12–16,19–21]

Support for the role of the GRK3 gene also comes from a recent animal model study we have conducted.[22] The behavioral hyperactivity induced by amphetamines in animals and man mimics aspects of mania, and hence has been suggested to model components of the pathophysiology of mania. Using an oligonucleotide-based microarray expression profiling method, we found, in rats, a 14-fold increase in GRK3 mRNA in the prefrontal cortex 24 h following administration of methamphetamine. In fact, GRK3 displayed the largest change in the expression of any of the 8000 genes examined in the prefrontal cortex. These data argue that GRK3 plays a substantial role in the brain's homeostatic response to dopamine and possibly other neurotransmitters.

The biological function, anatomical distribution, and chromosomal location of GRK3 led us to prioritize it as a positional candidate in this region of 22q. For these reasons, we have conducted a detailed examination of GRK3 in bipolar subjects in search of potential pathogenic mutations, and have examined several resulting single nucleotide polymorphisms (SNPs) for association to BPD.

Methods

Subjects

Two independent sets of families were analyzed. Probands were first drawn from set 1 described below for mutation screening. Parent–proband triads from set 1 were then used for association analysis. Subsequently, set 2 was examined for association.

Set 1: UCSD and NIMH triads A total of 153 nuclear families with 227 affected children were drawn from the UCSD and NIMH BP pedigrees. Of these families, 47 families were ascertained from the general North American population as part of the UCSD collaborative linkage study.[10] These families were collected at UCSD and at two collaborating sites, the University of British Columbia and the University of Cincinnati. Families were ascertained through a proband with either bipolar I (BP I) or bipolar II disorder (BP II), and were required to have a minimum of two additional members affected with either bipolar disorder or recurrent major depression. These families were derived from two independent linkage sets. The first (UCSD 20) included 20 families and 164 members and was used for our recently reported genome scan.[10] The second set (UCSD 32) includes 32 families now being studied as a replication set.

Informed consent procedures approved by each institution's institutional review board were employed. After informed consent was obtained, each family member was directly interviewed using the Structured Clinical Interview for DSM-III-R (SCID; in the case of the first set of 20 families) or the SCID for DSM-IV (in the case of the second set of families). As described previously,[10] information from the interview, other family members and medical records were reviewed by a panel of experienced clinicians in order to determine a consensus, best-estimate diagnosis. Blood was obtained from each subject for the establishment of lymphoblastoid cell lines and preparation of DNA.

DNA was obtained for an additional 106 families from the NIMH Genetics Initiative for Bipolar Disorder first wave pedigree collection.[11] This sample was collected as part of a four-site collaboration whose ascertainment and diagnostic procedures have been described previously.[11]

A total of 153 families were drawn from these three pedigree collections. In all families, the affected offspring were diagnosed with either BP I or BP II. Families were considered to be of a single ancestral ethnic/geographic origin if three or more grandparents were reported to be of similar ancestral origin. On this basis, 130 families were classified as northern European Caucasian. For 23 families, two or more grandparents were reported to have been of Ashkenazi Jewish, Mediterranean, Hispanic, African-American, Asian, or native American origin. Of the 130 northern European Caucasian families, 64 consisted of parent–offspring triads, 65 consisted of parents plus two affected siblings, and one consisted of parents plus three affected siblings. Altogether thereby, a total of 197 triads were extracted from the 130 Caucasian pedigrees.

Set 2: University of Toronto triads A total of 275 triads ascertained through probands with BP I or BP II disorders from Toronto and across Central Canada[23] were analyzed. Subjects were interviewed using SCID-I[24] and diagnoses were made using DSM-IV criteria by a 'best estimate' procedure. In all, 237 families, with 242 affected offsping, were Caucasian of northern European ancestry, mostly from Great Britain, 33 families were Caucasian of southern European or Mediterranean ancestry, mostly Italian and Jewish, and five families were not Caucasian.

Sequencing, mutation screening and sequence analysis To enrich the sample for subjects likely to contain a functional mutation in the GRK3 gene, families showing the most strongly positive lod scores for the markers D22S419 and D22S315 were selected from the UCSD family set and the 57-family NIMH subset which we previously used for a reanalysis of chromosome 22.[10] From the 12 families with the highest lod scores, 14 bipolar subjects were selected and sequenced over all the regions described below. An additional eight subjects from six families were sequenced just in the promoter region.

To identify variants, we sequenced the promoter region, exons, flanking intronic region, and 5′-UTR in subjects with BPD. The GRK3 gene spans 170 kb and has 21 exons ranging in size from 52 to 163 bp. Using genomic sequence from the Sanger Centre, we designed PCR primers to amplify ~500 bp genomic DNA fragments centered on each of the 21 exons and including approximately 200 bp of flanking intronic sequence (see Table 1 for positions of amplified fragments). (Primers and reaction conditions are available at http://psychgenes.ucsd.edu.) Primers were also designed to amplify 1.5 kb in the 5′ promoter region immediately upstream of exon 1 and 1.5 kb of the 3′-UTR immediately downstream of the coding portion of exon 21, each in overlapping segments. Fragments were sequenced bidirectionally using the Perkin Elmer Big Dye fluor-ddNTP sequencing kit and an ABI 377 sequencer, according to the manufacturer's recommendations. Minor modifications were used for sequencing the GC-rich region of the promoter (annealing temperature of 54°C and addition of 5% DMSO). Using these methods, we obtained full bidirectional sequence for all fragments examined except for a 38-bp region located 3′ to P-6 (ntds.1497–1534 in Figure 2). Sequencing gels were tracked and data extracted using ABI sequence analysis software. Chromatogram files were then transferred to a Sun Unix workstation for assembly into contigs using the Pred/Phrap/Consed suite of programs and SNPs were identified using Polyphred[25] and by visual inspection. All promoter and exon sequence was visually scanned to evaluate sequence quality, confirm SNPs, and check for possible false negatives (missed SNPs). Likewise, any regions of reduced sequence quality (<30 on the Pred/Phrap scale, or approx. error rate of 1 : 1000) were visually inspected and resequenced if necessary. All regions where sequence was visually inspected showed no discrepancy with Polyphred calls.

Table 1 Regions of GRK3 sequenced to identifiy variants

Exon number	*Position of first bp of the exon*[a]	*Exon length (bp)*[b]	*Relative pos. of first bp of seq*[c]	*Length of the sequenced genomic segment (bp)*
1	1	113	−1696	2124
2	39337	77	−149	1272
3	79552	74	−219	508
4	96536	102	−109	381
5	98590	75	−247	511
6	102699	62	−249	428
7	107256	52	−242	433
8	109397	92	−110	643
9	113776	100	−114	1160
10	120130	79	−284	530
11	122497	131	−204	549
12	125149	95	−202	496
13	130067	108	−144	416
14	138470	67	−112	502
15	139069	101	−313	514
16	144858	67	−147	415
17	146028	96	−155	520
18	149368	163	−191	520
19	153205	137	−185	489
20	156244	114	−219	513
21	157249	162	−176	1734

[a]The position of the first bp and the length of each of the 21 exons of GRK3 is indicated relative to the first bp of the coding region of exon 1. Positions were determined according to the Sanger Centre web site SUPERLINK_C22 in which system the first bp of the coding region of exon 1 is designated bp 9532042.
[b]The length in bp of the coding region of each of the 21 exons is listed. The lengths of the 3′-UTR and 5'-UTR are not known.
[c]For each exon-spanning segment sequenced, the distance in bp from the exon to the first base sequenced is indicated as well as the length of the genomic segment sequenced.

To compare human to mouse promoter sequence, murine genomic sequence which terminated 227 bp 5′ to exon 1 was downloaded from the Celera Discovery System (>GA_x5J8B7W5ADD, http://www.celera.com) and spliced to a murine sequence from the NCBI database (mmtrace sequence >gnl|ti|19424054 G10P638823RD9.T0) extending from 481 bp 5′ of exon 1 through exon 1. Human GRK3 sequence was downloaded from the Sanger Centre web site. Alignment of the human and mouse putative promoter regions was performed using NCBI's ''Blast 2 Sequences'' with the settings: match 2, mismatch −1, gap open 6, gap extension 0.4, × _dropoff 50, expect 10.0, word size 7.

Genotyping

Single base substitution SNPs selected for analysis were genotyped by the TaqMan allele specific assay method (Applied Biosystems) according to the manufacturer's protocols. For each site, primer pairs flanking the site to be interrogated were selected for PCR amplification of fragments of <150 bp (primers, probes, and reaction conditions are available at http://psychgenes.ucsd.edu). Two dual labeled probes centered on the SNP and differing in sequence by the 1-bp polymorphism of the SNP site itself were designed. The probes were labeled with 5′ reporter fluors FAM or TET and 3′ quencher TAMARA. Sensitivity and specificity for allelic discrimination were tested over a wide range of primer and probe concentrations on the DNA samples whose allele type was known by sequencing. Concentrations and cycling parameters were chosen for genotyping that produced clustered values for heterozygotes which separated from homozygotes by >4 standard deviations. Samples which gave ambiguous calls were retyped. Accuracy of typing was checked by retyping all 7 SNPs on 180 samples of triad set 1 and the SNP P-5 on 450 samples of triad set 2. No incorrect calls were detected.

We could not develop TaqMan reagents for a 5′-UTR single base deletion variant (P-6). Instead we typed it by standard size-based methods used for microsatellite genotyping. An FAM-labeled forward and unlabeled reverse primer pair were used to amplify a 228-bp genomic fragment spanning the variant. (Primers: fwd FAM-cgaccccggccggctacag, rev tcggccagcacagcctccag. Reaction mix: FailSafe PCR Enzyme Mix (Epicentre Technologies, Madison, WI, USA), buffer G, 0.25 μM primers, 40 ng genomic DNA in a 25 μl reaction. Thermal cycling: 82°C for 10 min, 98°C for 3 min, then four cycles of 98°C for 60 s, 71°C for 30 s less 0.5°C per cycle, 72°C for 45 s, followed by 38 cycles of 98°C for 60 s, 67°C for 45 s, 72°C for 30 s.) The 1-bp deletion was detected by size discrimination on a sequencing gel. All genotypes were read in a machine-assisted fashion using ABI software and confirmed by two independent human readers.

Transmission disequilibrium test (TDT) analysis

TDTs were carried out using the program Transmit.[26] Transmitted and non-transmitted alleles are counted from each heterozygotic parent to affected offspring. For triads in which a parental genotype is missing, alleles or haplotypes are estimated. Additional offspring can be included in the analysis without the association statistic being inflated by the presence of linkage by using the 'robust' empirical estimate of variance.[26]

Results

Identification of sequence variants

We sequenced 1500 bp of the putative promoter region, all 21 exons and approximately 200 bp of flanking intronic region, plus 1500 bp of the 3′-UTR in 14 bipolar subjects. We also sequenced the promoter region in an additional eight bipolar subjects. No SNPs were identified in the coding region. Seven SNPs were identified in intronic regions flanking exons (Figure 1 and Table 2), but none of these is anticipated to affect splicing nor generate a new splice site. Only one of these intronic SNPs (A630) was a 'high frequency' SNP, found by genotyping to have an allele frequency of 36% in triad set 1. Each of the six other intronic SNPs was observed in only a

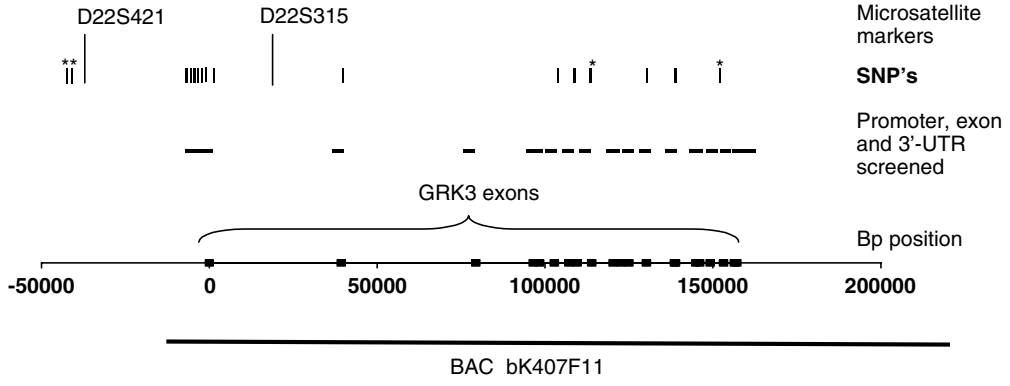

Figure 1 GRK3 mutation screening. The positions of the 21 exons of GRK3 are indicated as blocks on the map, and the corresponding segments sequenced are indicated by the bars above. Positions for the SNPs identified by sequencing are indicated by bars and the four high-frequency SNPs are marked by asterixes.

Table 2 SNPs identified in GRK3

SNP name	*SNP*	*Relative Position*[a]	
A486a	taattgccttcttac c/t gctccctcatagccc	−29635	
A486b	ttatctaaaggtacc t/c ctctcaccaagcttc	−29428	
P-1	gggattttaaaacat t/c aatttccagcttagg	−1329	
P-2	gcttaggtcttccta a/g gactgcctgcgtgtc	−1305	
P-3	tgaaaaccccaggca t/g cggggtggagtgatg	−1196	
P-4	gagctcttccatcag a/g gggctccctaagggc	−900	
P-5	gtggccaccccgagg g/a gagggcgaccgtaga	−382	
P-6	gcgggggcgcgcgga g/- gggggggctgccccg	−116	
First coding bp of exon 1		1	
In01-1	cgggcgatgcggggc c/g ggcctcgtcgttcca	253	140 bp 3′ of exon 1
In02-1	Ggatttttacatgaa c/a tcagcgttgctgaat	39644	230 bp 3′ of exon 2
In05-1	tagtggcaaacaggg a/c agacttgtctgtgtc	102482	210 bp 5′ of exon 6
In08-1	gaggcgggtggatca c/t gaggtcaggagatca	109691	200 bp 3′ of exon 8
A630	gaatatagtcaagac c/t tgactaaacgactgg	114181	305 bp 3′ of exon 9
In13-1	gtgatctttaagcca t/c gtgctaatctggaag	130299	130 bp 3′ of exon 13
In14-1	aactgtattcttgtc t/c gtatggtaattgtca	138566	29 bp 3′ of exon 14
A665	cgaggaacttagcac a/g acatgtcacttgtca	150011	480 bp 3′ of exon 18

[a]Map positions were derived from the Sanger Centre web site according to the SUPERLINK_C22. The first bp of exon 1 is SUPERLINK_C22 position 9532042. The position of each of the SNPs identified in this project is indicated relative to the first bp of exon 1 of GRK3. A486a and A486b have been identified by the SNP Consortium as rs1467387 and rs576895 respectively. A665 has ben identified by Celera as CV2190064. Eight of the SNPs were found in only one of the subjects sequenced—P-3 and P-4 in c1402, a Caucasian; In01-1 and In14-1 in c0105, a Caucasian; and In02-1, In05-1, In08-1, and In13-1 in c1216, a native American.

single chromosome. Four were found in one subject and two in another subject. Because of their low frequency, we did not genotype these SNPs.

The GRK3 promoter has yet to be functionally identified. However, based on the following three points, we reasoned that the promoter for GRK3 lies in the genomic region immediately 5′ to the first coding exon 1. (1) The region is GC rich (75% GC over a 500-bp region just upstream of exon 1 with a maximum of 94% GC for 104 bp). (2) The region meets the definition for a CpG island (http://www.sanger.ac.uk). 3. The analogous 5' region of the GRK2 gene has been demonstrated to have promoter activity by expression vector assays and has been shown by RNA protection assays to contain the transcription start site.[27] We identified 6 SNPs within the 1500-bp putative GRK3 promoter region (Table 2 and Figure 2). Four of these SNPs, P-1, -2, -5, and -6, were observed on more than one chromosome and were selected for genotyping.

In addition, as part of an independent ongoing project to examine linkage disequilibrium (LD) across a broader region in the vicinity of GRK3 (which will be reported elsewhere), we identified a second high-frequency SNP in the 3′ region of GRK3 plus 2 high-frequency SNPs 28 kb upstream of GRK3 by direct sequencing of genomic DNA. The 3′ SNP has also been identified by Celera (Celera Discovery System) and the two 5′ high-frequency SNPs have also been identified in the NCBI dbSNP database (http://www.ncbi.nlm.nih.gov)). We genotyped these four high-frequency flanking SNPs and the four promoter SNPs in set 1 of the bipolar triads. Two of the promoter SNPs were subsequently genotyped in triad set 2.

Alignment of the human and mouse GRK3 promoter regions

To determine if the variants in the putative promoter region affect evolutionarily conserved regulatory motifs, we compared the human to mouse sequence 5′ to and including exon 1 (see Figure 3). Exon 1 is highly conserved: 92% identity over 113 bp. A stretch of 35 bp located at human –1640 bp (with respect to the ATG of exon 1) and at mouse –1078 is 86% identical. Over the intervening sequence, stretches of moderate alignment are evenly distributed: there are seven short blocks, from 9 to 12 bp long, of complete identity and ~500 bp total of ≥70% identity. A GA-rich region lies 5′ to P-5 (67 bp with 96% GA in human and 73 bp with 100% GA in mouse). A GC-rich region lies 3' to P-5 and extends to exon 1 (251 bp with 89% GC in human and 119 bp with 87% GC in mouse). Of note, all six SNPs in the putative promoter region fall in sequence of low to no identity. Only for P-2 is the variant base conserved between human and mouse. Both P-5 and P-6 lie in regions of no human to mouse homology. P-6 is a single G deletion at the base of a 28-nucleotide potential hairpin structure that comprised 12 aligning G–C pairs.

Searches run on the web site Transcription Element Search System (TESS, http://www.cbil.upenn.edu)[28]

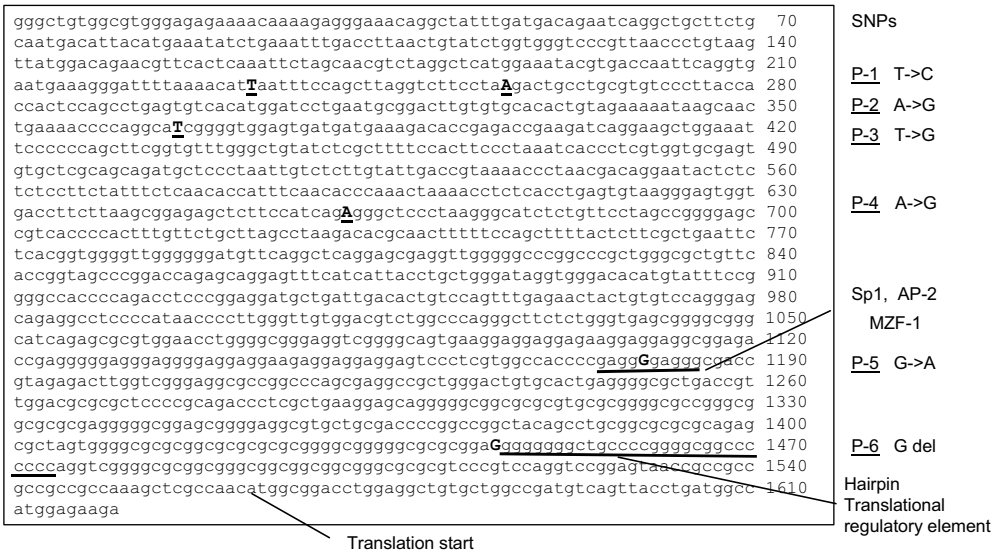

Figure 2 SNPs in the GRK3 promoter. The six promoter region SNPs (capitalized base) identified and the corresponding base change or single base deletion are indicated. Three transcription factors predicted to bind to the P-5 site are listed, the hairpin repeat immediately 3′ to P-6 is underlined, and the ATG translation start site of exon 1 is indicated.

predicted more than eight potential transcriptional binding factors that might recognize the sequence at each of the variant positions. For P-5, binding is predicted for SP1, AP-2, MZF-1, CAC-binding_protein, PEB1, Olf-1, MalT, and Ttk_88 K.

TDT analysis

Table 3 shows results for triad set 1 of transmission disequilibrium analyses done using Transmit on genotype data for the four promoter SNPs selected for testing. The number of times a marker was transmitted or not transmitted from a heterozygous parent to an affected child is shown. For P-1 there were 20 transmissions and six non-transmissions observed, and for P-5 there were 15 transmissions and four non-transmissions observed. Transmitted and non-transmitted alleles are counted from each heterozygotic parent to affected offspring. Fractional values are contributed by a small number of triads for which one parental genotype was missing. χ^2 statistics were derived using the 'robust' option for estimation of variance, which uses an empirically derived variance–covariance matrix that corrects for the presence of more than one affected offspring in some triads.[26] Promoter SNPs were tested under the *a priori* hypothesis that the rare variant is a risk allele and will be associated with an excess of transmissions. Therefore, the significance of χ^2 is assessed by one-sided *P* values. Two SNPs had *P* values less than 0.05. For P-1 $\chi^2 = 5.6$, *P* value = 0.018, and for P-5 $\chi^2 = 6.0$, *P* value = 0.014. Empirical *P* values computed using 100,000 replications were similar to those derived from the 'robust' χ^2 statistic; however, Transmit failed to produce *P* values for some of the data points, and therefore only the 'robust' values are listed in the tables. These *P* values are not corrected for multiple testing. Also, because of the strong LD among markers, these tests are not fully independent.

Haplotypes could be unambiguously assigned for all triads except one. Only six haplotypes for these four SNPs were observed (see Table 3), indicating a high degree of LD among the markers as anticipated by their close proximity. As they are all in tight LD, it is not possible to determine from these triads which of the three SNPs, P-1, P-5, or P-6, is most likely to be functionally relevant, or to exclude the possibility that the functional SNP is some other nearby variant not yet identified.

In order to attempt to define the extent of the region of association, we also genotyped four additional SNPs. One of these was an intronic SNP 3′ of exon 9 (A630) identified as part of the studies described above. Another SNP in intron 18 (A665) and two others, approximately 28 kb upstream of exon 1 (A486a and A486b); were identified as part of a study of LD across a larger genomic region and will be reported separately. The analyses of these four flanking SNPs, located 28 kb upstream and 114 and 150 kb downstream, considered individually and as bi-allelic haplotypes, were uniformly negative (Table 4). This provides some evidence to restrict the region of association to the approximately 140 kb containing the 5′ end of the gene.

These data observed for the promoter region SNPs in the first triad set suggest a significant association with susceptibility for disease. To verify this we genotyped two of these SNPs in a second independent set of triads recruited at the University of Toronto from eastern Canada. Notably, these data (Table 5) show a strong trend toward association for P-5 among northern European Caucasians. For P-1 all excess transmissions were accounted for by the P-1/5

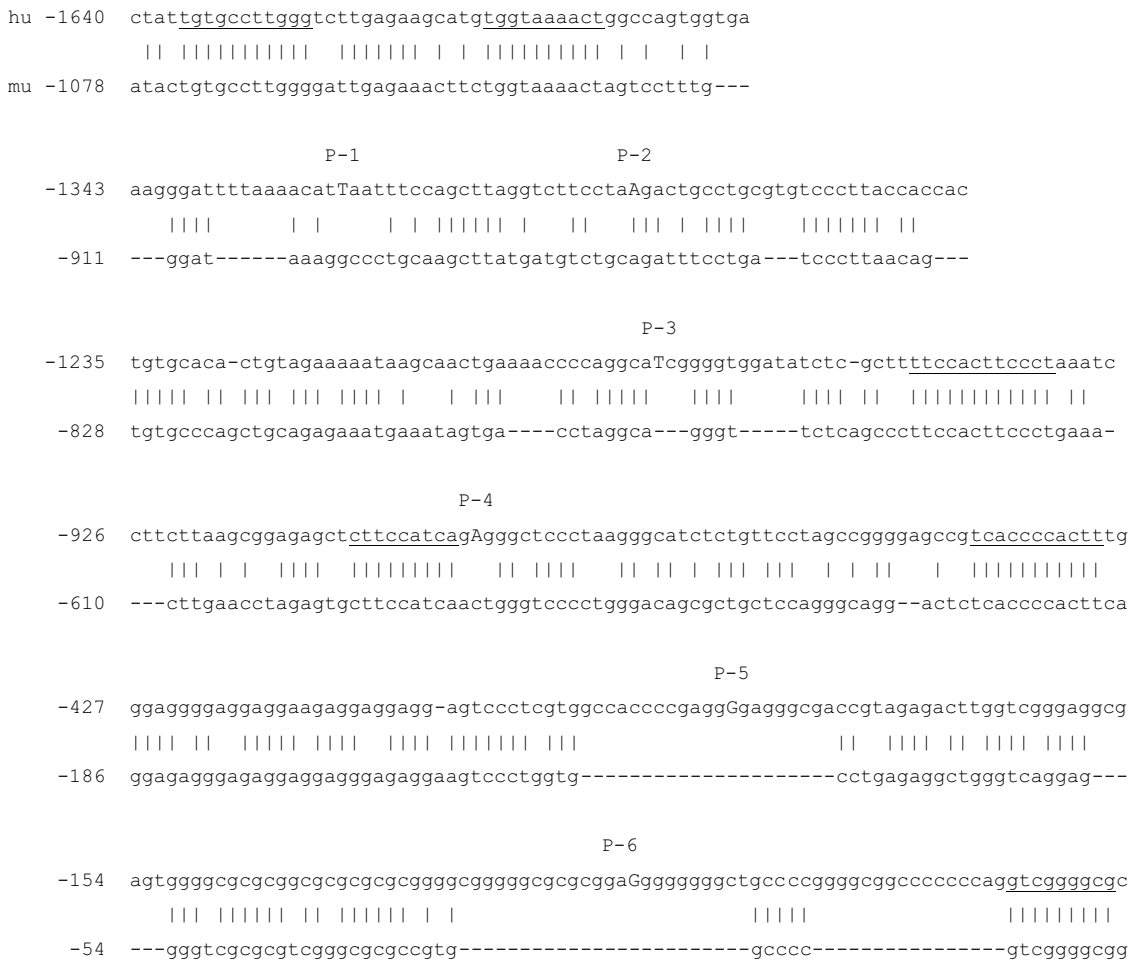

Figure 3 Alignment of human and mouse GRK3 putative promoter regions. In each block of aligned sequence human is on the top, mouse on the bottom. Bp positions are indicated relative to the first coding base of exon 1. Alignment performed as described in Methods was anchored at the 3′ end by exon 1 (not shown), and at the 5' end by a 35 bp region of 86% similarity contained in the first block of sequence shown. Below this block of sequence five blocks of sequence containing the 6 promoter region SNPs are shown. Stretches of 9 or more consecutive identical bp are underlined.

haplotypes. χ^2 analysis of the combined data set is shown in Table 6. For P-5 a total of 26 transmissions and 7.7 non-transmissions were observed when the two sets were combined for a χ^2 of 9.6 and an associated P value of 0.0019. A global test incorporating all three observed haplotypes was also significant ($\chi^2 = 10.0$, $P = 0.019$). Although the number of Caucasian triads not of northern European origin is small and the number of non-Caucasian triads is also small, the data suggest that there is no trend toward excess transmission of any of the promoter markers in either of these groups.

To examine haplotype structure across this region, we determined LD among all pairs of SNPs. As anticipated by their physical proximity, the two 5′ SNPs (A486a and A486b) were in statistically significant disequilibrium ($D' = 0.415$, $P = 0.0000$), as were the four promoter SNPs (D' values for pairs ranged from 0.74 to 1.00) as well as the two 3' SNPs (A630 and A665, $D' = 0.764$). Of note, while D' values were not significant between the 5′ SNPs and the promoter SNPs, there was strong LD observed between the promoter SNPs and the 3′ SNPs. Seven of the eight promoter : 3′ SNP pairs had D' values of 0.71–1.00, indicating preservation of haplotype structure across nearly the full length of GRK3.

Association of P-5 with the evidence for linkage at 22q12.1

If the variant P-5 is indeed associated with disease, we would then expect that P-5 would be present in those families contributing to the linkage evidence that we observed at D22S419, the microsatellite immediately centromeric to GRK3. Of the 153 families from which triads were drawn for triad set 1, 73 have been analyzed for linkage as part of the three data sets (UCSD 20, UCSD 32, and NIMH) described above. We examined the lod scores for subsets of these 73 families. Only seven of these families had the P-5

Table 3 TDT data on four promoter SNPs in the UCSD/NIMH families (153)[a]

	Caucasian					*Non-caucasian*
	Northern European			*Non-northern European*		
	Allele frequency		*130*			*8*
SNP	*BP*	*ctrl*	*T:N*	χ^2	*P value*	*T:N*	*T:N*
P-1	0.051	0.015	20:6	5.6	0.018	0:1	3:3
P-2	0.015	0.018	6.1:7.1	0.1	0.8		
P-5	0.038	0.010	15:4	6.0	0.014		1:3
P-6	0.033	0.013	13:5	4.7	0.03		1:3
Haplotype							
# 1 # 2 # 5 # 6							
+			6:2			0:1	2:0
+ +			4:0				
+ + +			10:4	3.7	0.055		1:3
+ +			1:0				
+			2:1				
+			6:7				

[a]The UCSD/NIMH family set was separated into the indicated number of northern European Caucasian, non-northern European Caucasian and non-Caucasian families. For each group the number of transmissions (*T*) from heterozygotic parents to affected offspring determined by Transmit is listed. Since Transmit output lists the number of expected transmissions (*E*) rather than non-transmissions (*N*), *N* is calculated here from the relationship $N=T-2(T-E)$. For northern Europeans corresponding χ^2 values determined by the 'robust' option and one-sided *P* values are indicated. Allele frequency in bipolars is the allele frequency of the variant in affected offspring of northern European Caucasians. Allele frequency on non-transmitting control chromosomes is estimated by dividing the allele frequency on transmitted chromosomes by the *T:N* ratio. Unambiguous haplotypes could be determined for all but one family. (+) indicates the haplotype has the variant base at that position.

Table 4 TDT data on the ''flanking SNPs'' in the UCSD/NIMH families[a]

Haplotype	*Northern European Caucasian triads*		
	Haplotype frequency	*T*	*N*
A486a/A486b			
1.1	0.65	260	253
2.1	0.079	29	31
1.2	0.14	56	61
2.2	0.12	49	49
A630/A665			
1.1	0.54	208	218
2.1	0.027	12	10
1.2	0.10	40	43
2.2	0.33	130	119

[a]For the indicated haplotypes allele 1 is the common allele, allele 2 is the variant. Haplotype frequencies are calculated for the UCSD/NIMH northern European Caucasian triads: (*T*) transmissions; (*N*) non-transmissions. None of the differences are significant.

variant. Therefore, because of this limited amount of data, it is not possible to make definitive conclusions about the portion of linkage evidence at D22S419 that is explained by P-5. One of the two families with P-5 from the UCSD 20 set yielded a lod score of 0.91. This compares to a summed lod score of 0.71 in the P-5 negative families in this family set. This suggests that P-5 makes a disproportionate contribution to the total lod score at this locus in this family set. The entire UCSD 32 family set provides additional support for linkage to the GRK3 region (to be reported separately). However, little evidence for linkage at D22S419 is present in either the one family with P-5 or the 14 without. Therefore, this family set contributes little information to this question. Similarly, in the NIMH set, three families with P-5 contribute little to the total lod score of 1.49 observed in the 38 northern European Caucasian families. In addition, two native American families, one with P-5 and one without, had individual family lods of 0.60 and 0.11. Hence, while families with haplotypes marked by the variant P-5 contributed to the total linkage among the subset of families of northern European Caucasian ancestry, families with other haplotypes (haplotypes without P-5) contributed more overall to the evidence for linkage at 22q12.1 among these northern European Caucasian families and may contribute to linkage among non-northern European Caucasian families. This argues that P-5 may explain a portion, but not the majority of the linkage signal seen at D22S419.

Table 5 TDT data on 2 SNPs in the Toronto bipolar triads (275)[a]

	Caucasian						*Non-Caucasian*
	Northern European					*Non-northern European*	
	Allele frequency		*237*			*33*	*5*
SNP	*BP*	*ctrl*	*T:N*	*χ2*	*P value*	*T:N*	*T:N*
P-1	0.048	0.039	23 : 18.6	0.6	0.45	2 : 4	0 : 2
P-5	0.023	0.007	11 : 3.5	3.8	0.055	2 : 3	1 : 2
Haplotype							
# 1 # 5							
+ −	0.026	0.032	12.5 : 15.5			1 : 1	0 : 1
+ +	0.023	0.007	11 : 3.5	3.8	0.055	1 : 3	0 : 1
− +			0 : 0			1 : 0	1 : 1

[a]Separation into ethnic groups and analyses are as in Table 3.

Table 6 TDT data on all triads combined (428)[a]

	Caucasian					*Non-Caucasian*
	Northern European				*Non-northern European*	
	Allele frequency		*367*		*48*	*13*
SNP	*BP*	*ctrl*	*T:N*	*χ^2* *P value*	*T:N*	*T:N*
P-1	0.049	0.028	43 : 24.5	5.0 0.025	2 : 5	3 : 5
P-5	0.030	0.009	26 : 7.7	9.6 0.0019	2 : 3	2 : 5
Haplotype						
# 1 # 5						
+ −	0.021	0.020	18.6 : 17.4	0 0.8		
+ +	0.029	0.009	25 : 7.6	8.8 0.003		
− +	0.001	0.000	1 : 0			
Global chisquared test, 3 df				10 0.019		

[a]Separation into ethnic groups and analyses are as in Table 3.

Clinical phenotypes in subjects with the P-5 variant

Since variants conferring risk might be associated with specific clinical phenotypes, we examined the frequency of BP I *vs* BP II, anxiety diagnoses, and the presence of psychotic features during episodes of mania or depression among subjects inheriting P-5 and compared this to subjects without P-5 for the families of northern European Caucasian ancestry. A nonsignificant excess of BP II diagnoses was present in subjects with P-5. Among the 47 UCSD and 237 Toronto families were 17 subjects with P-5 of whom eight were BP I (47%) and nine were BP II (53%), whereas the subjects without P-5 were 66% BP I and 34% BP II. (Because the NIMH families were ascertained through BP I probands, they were not included in this analysis.) In the three family sets combined, a history of psychotic features was present in 11 of 17 P-5 positive BP I subjects (65%) compared with 60% among BP I subjects without P-5. The number of P-5 positive subjects with a secondary diagnosis of alcoholism or other drug dependence, panic disorder, panic attacks, or any anxiety diagnosis, was small and not clearly different from the frequency in subjects without P-5.

Discussion

We have screened the GRK3 gene for functional variants and identified six SNPs in the putative promoter region. We tested four of these SNPs for association with BPD. A statistically significant evidence was obtained indicating that the SNP-designated P-5 is associated with disease in families of northern European ancestry. A trend toward support of this result was then obtained in an

independent sample. Among the 367 northern European Caucasian families in the combined sample, the allele frequency of SNP P-5 in affected offspring was 3% and an excess of transmission from heterozygotic parents to affected offspring was observed at a ratio of 26 transmissions to 7.7 non-transmissions, giving a χ^2 of 9.6 and a P value for association of 0.0019. No evidence for association was found for SNPs 28 kb upstream or 114 and 150 kb downstream of exon 1, thereby providing some data to bracket the region of association. Together, these data argue for the association of BPD to the GRK3 gene.

These results must be qualified based on the statistical consideration that the overall estimate for the strength of the association in the combined data set should take into consideration a correction for multiple testing. Since the two promoter SNPs examined in the combined triad set are in strong LD and are found on shared haplotypes, the number of independent tests performed is less than two. Making an overly conservative Bonferroni correction for six tests (ie for two SNPs and for separation into three ethnic groups) gives a corrected P value of 0.011 which is still significant.

Further support for these results comes from a similar relative risk suggested by the two triad sets. Although the number of families in which P-5 is observed is small in both sets, the second set is a qualitatively similar replication of the first since the transmission-to-non-transmission ratio is approximately 3.4 in both triad sets.

The transmission-to-non-transmission ratio of 3.4 provides us with a first-order estimate for the relative risk conferred by the variant P-5. This is a relatively high-risk ratio for a polygenic complex disorder and suggests that a relatively strong biological effect is conferred by inheritance of the P-5 variant in individuals of northern European Caucasian ancestry. However, the overall population frequency of the P-5 variant is low; the observed allele frequency is 3% among offspring affected with BPD in these two family sets. Therefore, the P-5 variant conveys a moderate increase in risk, but in only a small portion of patients, and therefore likely has a relatively low population attributable risk. The combined set of 367 northern European Caucasian families (with 439 affected offspring) provides adequate power to detect association with disease as long as the marker is the disease variant or is in strong LD with the disease variant. For a marker with a population allele frequency of 1%, a heterozygotic genotype relative risk (Aa) of 3, a homozygotic genotype relative risk (AA) of 9, and an assumed disease frequency of 1%, this family set provides 93% power to detect association at a P value of 0.05.[29,30]

We chose to pursue a candidate gene approach based on the biological function, anatomical distribution, and chromosomal location of GRK3. We have judged that these considerations make it a more compelling candidate than other genes in the region. However, it is possible that either the linkage or LD results really reflect the functional effect of a mutation in another nearby gene. A 10 cM interval centered on the peak in evidence for linkage in the 22q12.1 region (that is, centered at the markers D22S925, 533, and 419, which, as discussed above, are located immediately centromeric to GRK3) spans just 4.3 Mb, according to the Marshfield sex-averaged gene map. A distance of 5 cM proximal to GRK3 is approximately 3.1 Mb, while 5 cM distal to GRK3 is only 1.2 Mb. The Sanger Centre lists 80 known and proposed genes in this region, discounting a large cluster of immunoglobulin light chains. GRK3 is located at pos. 22.7 Mb. In considering genes in this region, we judged 11 of the known or predicted genes as good candidates: three zinc finger proteins (pos. 19.6, 19.6, and 20.8), G protein alpha *z* (20.2), RAB36 (20.2), RALGDS form A (20.8), metalloproteinase 11 (20.9), calcineurin-binding protein 1 (21.1), adenosine A2a receptor (21.6), lipoprotein receptor-related protein (22.5), and seizure-related gene 6 (mouse)-like (23.4). The Celera gene list parallels the Sanger list except for four additional predicted/named genes of which we sorted one as a 'good candidate': WD-40 domain protein (20.3). However, the absence of association with the two SNPs 28 kb 5′ and SNPs at 114 and 150 kb 3′ to the promoter provides some limited support that the observed association is indeed to the GRK3 gene.

If the P-5 variant is associated to BPD, then it could result from two possibilities. The P-5 variant might affect promoter function and be directly responsible for the susceptibility; alternatively, it may be close to and in linkage disequilibrium with the functional mutation which has yet to be identified. Several observations argue that the P-5 variant itself may not be the functional variant. One of these stems from the differences seen when stratifying by ethnicity. The necessity for separating the test families into major ethnic groups is justified by the long-standing observation that the prevalence for many disease mutations is highly distinct for different populations. Further, linkage disequilibrium relationships may vary between different populations. Although the numbers of non-Caucasian families in this study are very small, we see no trend toward excess transmission among non-Caucasians. Precedence is also well established for the prevalence of disease alleles varying markedly among subgroups of Caucasians. A large number of such examples have been extensively characterized in Ashkenazi Jews. While in our study the number of non-European Caucasians is too small to make definitive statements, there appears to be no trend toward excess transmission of any of the promoter SNPs tested. The simplest conclusion is that the variant is not itself a functional mutation, but in linkage disequilibrium with an adjacent functional variant. In this case, in northern European Caucasians, but not in non-northern Europeans, the P-5 variant would lie on a haplotype which carries a disease mutation outside of the region which we have sequenced. To investigate this possibility, we will

need to determine the physical extent of linkage disequilibrium and of association with disease along these haplotypes, as well as to screen for mutations over a larger region. Alternatively, if the P-5 variant is the disease mutation, it is possible that an independently inherited epistatic risk co-factor operates together with P-5 to produce disease. Such a factor might be more prevalent in northern European Caucasians than in other populations. Conversely, a protective factor might be highly prevalent in non-northern Europeans.

Sequence analysis of the region containing the P-5 variant suggests some possible functions. The P-5 variant lies in the sequence GGGAGGG which has been shown to be a conserved DNA sequence motif in enhancer regions in a number of genes.[31,32] Strikingly, when we compared the human to mouse genomic sequence of the upstream region of GRK3, we found not only that this GGGAGGG element is not present in mouse, but also that the P-5 variant is located in a short DNA stretch which is unique to human sequence (see Figure 3). The P-6 variant is a single G deletion out of a run of 8 G's lying at the base of a potential hairpin structure formed by 12 G-C pairs. Such hairpin structures in UTRs have been shown to have significant regulatory effects on translation.[33] However, as is true for P-5, this hairpin and the P-6 variant do not have homology in the murine genomic sequence. Interestingly then, both these variants might pertain to functions of GRK3 that are unique to only certain mammals or to primates. Likewise, the other promoter variants we have identified lie in sequence for which there is only moderate-to-minimal similarity between human and mouse, and only one variant, P-2, is the identical nucleotide in both species.

In summary, we have identified variants in GRK3 which may affect the regulation of gene expression, and have shown that they are associated with increased susceptibility for disease. These data suggest the hypothesis that a dysregulation in the time or place of GRK3 expression results in aberrant neurotransmitter signaling and an increased susceptibility to develop BPD. The narrowest formulation of this hypothesis is that a failure of dopamine receptor homologous desensitization results in an effective hypersensitivity to dopamine and can cause a susceptibility to disease. A broader formulation takes into account the fact that GRK3 has been demonstrated to desensitize many GPCRs, including adrenergic, muscarinic, histaminic, dopaminergic, CRF-1, and opioid receptors.[12–16,21] GRK3 might, therefore, play an important role in the regulation of any one of many GPCRs, and the regulation might be either presynaptic or postsynaptic and might involve either stimulatory or inhibitory circuits. The complexity of these functions raises the possibility that different mutations each responsible for distinct changes in the time and place of GRK3 expression could have strikingly different effects on GPCR regulation and thus be associated with distinct subphenotypes of BPD.

A final possibility is that variants in the time or place of GRK3 expression might result in a more complicated perturbation (dysregulation) of signaling than a mere failure in homologous desensitization. Following phosphorylation by GRKs, β-adrenergic receptors are bound by β-Arrestin which terminates G protein-induced signaling and initiates clatherin-mediated endocytosis.[14,15] β-arrestin has been shown capable of initiating subsequent intracellular signaling through MAP kinase pathways.[34] Therefore, changed expression of GRK3 in specific neurons could have complicated effects on multiple aspects of neurotransmitter signaling. Together, the multiple intriguing roles of GRK3 in neuronal signaling in conjunction with the data for association that we here report argue for the role of GRK3 in the susceptibility to BPD.

Acknowledgements

We thank the family members who participated in this study, without whom it would not have been possible. This work was supported by Novartis Pharma AG, and by grants to JRK from the Department of Veterans Affairs and the National Institute of Mental Health (NIMH) (MH47612, MH59567), the UCSD Mental Health Clinical Research Center (MH30914), and the UCSD General Clinical Research Center (M01 RR00827). RLH was supported by a VA Merit Review grant, the VA Mental Illness Research, Education and Clinical Center (MIRECC) of VISN22, and the NIMH Mental Health Clinical Research Center (PHS MH20914-14). TBB was supported by a Veterans Administration Fellowship in Psychiatric Research/Neurosciences and by a NARSAD Young Investigator Award. We also acknowledge the technical assistance of Tatyana Shekhtman.

References

1 Kendler KS, Pedersen N, Johnson L, Neale MC, Mathe AA. A pilot Swedish twin study of affective illness, including hospital- and population-ascertained subsamples. *Arch Gen Psychiatry* 1993; **50**: 699–700.
2 Gershon ES, Hamovit J, Guroff JJ, Dibble E, Leckman JF, Sceery W *et al*. A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. *Arch Gen Psychiatry* 1982; **39**: 1157–1167.
3 Rice J, Reich T, Andreasen NC, Endicott J, van Eerdewegh M, Fishman R *et al*. The familial transmission of bipolar illness. *Arch Gen Psychiatry* 1987; **44**: 441–447.
4 Sadovnick AD, Remick RA, Lam R, Zis AP, Yee IM, Huggins MJ *et al*. A mood disorder service genetic database: morbidity risks for mood disorders in 3942 first-degree relatives of 671 index cases with single depression, recurrent depression, bipolar I or bipolar II. *Am J Med Genet* 1994; **54**: 132–140.
5 Spence MA, Flodman PL, Sadovnick AD, Bailey-Wilson JE, Ameli H, Remick RA. Bipolar disorder: evidence for a major locus. *Am J Med Genet* 1995; **60**: 370–376.
6 Prathikanti S, McMahon FJ. Genome scans for susceptibility genes in bipolar affective disorder. *Ann Med* 2001; **33**: 257–262.
7 Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; **11**: 241–247.

8 Badner JA, Gershon ES. Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. *Mol Psychiatry* 2002; **7**: 405–411.
9 Lachman HM, Kelsoe JR, Remick RA, Sadovnick AD, Rapaport MH, Lin M *et al*. Linkage studies suggest a possible locus for bipolar disorder near the velo-cardio-facial syndrome region on chromosome 22. *Am J Med Genet* 1997; **74**: 121–128.
10 Kelsoe JR, Spence MA, Loetscher E, Foguet M, Sadovnick AD, Remick RA *et al*. A genome survey indicates a possible susceptibility locus for bipolar disorder on chromosome 22. *Proc Natl Acad Sci USA* 2001; **98**: 585–590.
11 Edenberg HJ, Foroud T, Conneally PM, Sorbel JJ, Carr K, Crose C *et al*. Initial genomic scan of the NIMH genetics initiative bipolar pedigrees: chromosomes 3, 5, 15, 16, 17, and 22. *Am J Med Genet* 1997; **74**: 238–246.
12 Carman CV, Benovic JL. G-protein-coupled receptors: turn-ons and turn-offs. *Curr Opin Neurobiol* 1998; **8**: 335–344.
13 Macrae AD, Lefkowitz RJ. The function and regulation of G-protein-coupled receptor kinases. In: O'Malley BW (ed). *Hormones and Signaling*. Academic Press: San Diego, CA, 1998, pp 65–88.
14 Penn RB, Benovic JL. Mechanisms of receptor regulation. In: Conn PM (ed), *Endocrinology Vol 1: Cellular Endocrinology*. Oxford University Press: New York, 1998, pp 125–164.
15 Pitcher JA, Freedman NJ, Lefkowitz RJ. G protein-coupled receptor kinases. *Annu Rev Biochem* 1998; **67**: 653–692.
16 Dautzenberg FM, Hauger RL. The CRF peptide family and their receptors: yet more partners discovered. *Trends Pharmacol Sci* 2002; **23**: 71–77.
17 Arriza JL, Dawson TM, Simerly RB, Martin LJ, Caron MG, Snyder SH *et al*. The G-protein-coupled receptor kinases beta ARK1 and beta ARK2 are widely distributed at synapses in rat brain. *J Neurosci* 1992; **12**: 4045–4055.
18 Erdtmann-Vourliotis M, Mayer P, Ammon S, Riechert U, Hollt V. Distribution of G-protein-coupled receptor kinase (GRK) isoforms 2, 3, 5 and 6 mRNA in the rat brain. *Brain Res Mol Brain Res* 2001; **95**: 129–137.
19 Tiberi M, Nash SR, Bertrand L, Lefkowitz RJ, Caron MG. Differential regulation of dopamine D1A receptor responsiveness by various G protein-coupled receptor kinases. *J Biol Chem* 1996; **271**: 3771–3778.
20 Kim KM, Valenzano KJ, Robinson SR, Yao WD, Barak LS, Caron MG. Differential regulation of the dopamine D2 and D3 receptors by G protein-coupled receptor kinases and beta-arrestins. *J Biol Chem* 2001; **276**: 37409–37414.
21 Dautzenberg FM, Braun S, Hauger RL. GRK3mediates homologous desensitization of the CRF1 receptor: a novel mechanism contributing to stress adaptation. *Am J Physiol Regul Integr Comp Physiol* 2001; **280**: R935-R946.
22 Niculescu III AB, Segal DS, Kuczenski R, Barrett T, Hauger RL, Kelsoe JR. Identifying a series of candidate genes for mania and psychosis: a convergent functional genomics approach. *Physiol Genomics* 2000; **4**: 83–91.
23 Mundo E, Zai G, Lee L, Parikh SV, Kennedy JL. The 5HT1Dbeta receptor gene in bipolar disorder: a family based association study. *Neuropsychopharmacology* 2001; **25**: 608–613.
24 Spitzer RL, Williams JBW, Gibbon M. Structured Clinical Interview for DSM-3-R (SCID). New York State Psychiatric Institute, Biometrics Research: New York, 1987.
25 Nickerson DA, Tobe VO, Taylor SL. Polyphred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res* 1997; **25**: 2745–2751.
26 Clayton D. A generalization of the transmission/disequilibrium test for uncertain-haplotype transmission. *Am J Hum Genet* 1999; **65**: 1170–1177.
27 Ramos-Ruiz R, Penela P, Penn RB, Mayor Jr F. Analysis of the human G protein-coupled receptor kinase 2 (GRK2) gene promoter: regulation by signal transduction systems in aortic smooth muscle cells. *Circulation* 2000; **101**: 2083–2089.
28 Schug J, Overton GC. Computational Biology and Informatics Laboratory, editor. TESS: Transcription Element Search Software on the WWW. School of Medicine, University of Pennsylvania: PA, 1997, CBIL-TR-1997-1001-v0.0.
29 Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996; **273**: 1516–1517.
30 Chen WM, Deng HW. A general and accurate approach for computing the statistical power of the transmission disequilibrium test for complex disease genes. *Genet Epidemiol* 2001; **21**: 53–67.
31 Song J, Ugai H, Kanazawa I, Sun K, Yokoyama KK. Independent repression of a GC-rich housekeeping gene by Sp1 and MAZ involves the same cis-elements. *J Biol Chem* 2001; **276**: 19897–19904.
32 Nordhoff V, Hubner K, Bauer A, Orlova I, Malapetsa A, Scholer HR. Comparative analysis of human, bovine, and murine Oct-4 upstream promoter sequences. *Mamm Genome* 2001; **12**: 309–317.
33 Klausner RD, Rouault TA, Harford JB. Regulating the fate of mRNA: the control of cellular iron metabolism. *Cell* 1993; **72**: 19–28.
34 McDonald PH, Chow CW, Miller WE, Laporte SA, Field ME, Lin FT *et al*. Beta-arrestin 2: a receptor-regulated MAPK scaffold for the activation of JNK3. *Science* 2000; **290**: 1574–1577.