Published in final edited form as:

Am J Med Genet B Neuropsychiatr Genet. 2011 March ; 156(2): 168–176. doi:10.1002/ajmg.b.31151.

Fine Mapping of Candidate Regions for Bipolar Disorder Provides Strong Evidence for Susceptibility Loci on Chromosomes 7q

Haiyan Xu^{1,2,8}, Rong Cheng^{2,9}, Suh-Hang Juo^{2,3,4}, Jianjun Liu^{2,5}, Jo Ellen Loth⁶, Jean Endicott⁶, Conrad Gilliam^{1,2,7}, and Miron Baron^{2,6}

¹Department of Genetics and Development, Columbia University, New York, NY

²Columbia Genome Center, Columbia University, New York, NY

³Department of Epidemiology, Columbia University, New York, NY

⁴Graduate Institute of Medical Genetics, Kaohsiung Medical University, Kaohsiung, Taiwan

⁶Department of Psychiatry, Columbia University, New York, NY

Abstract

Genomewide scans of bipolar disorder (BP) have not produced consistent linkage findings. Follow-up studies using enlarged samples and enhanced marker density can bolster or refute claims of linkage and pave the way to gene discovery. We conducted linkage and association analyses, using a ~3-cM density map of 10 candidate regions, in a large BP pedigree sample (865 individuals from 56 pedigrees). The candidate regions were identified in a previous 10-cM genome-wide scan using a subset of this sample (373 individuals from 40 pedigrees). The present sample consists of the expanded original pedigrees ('core' pedigrees) and 16 additional pedigrees. We obtained experiment-wide significant linkage on chromosome 7q34 (LOD score 3.53, $p < 0.001$), substantially stronger than that observed in the genome-wide scan. Support for linkage was sustained on chromosomes 2p13, 4q31, 8q13, 13q32, 14q21 and 17q11, though at a more modest level. Family-based association analysis was consistent with the linkage results at all regions with linkage evidence, except 4q and 8q, but the results fell short of statistical significance. Three of the previously implicated regions – 9q31, 10q21 and 10q24 – showed substantial reduction in evidence of linkage. Our results strongly support 7q34 as a region harboring susceptibility locus for BP. Somewhat lesser, yet notable support was obtained for 2p13, 4q31, 8q13, 13q32, 14q21 and 17q11. These regions could be considered prime candidates for future gene finding efforts.

Correspondence: Rong Cheng, Ph.D., Taub Institute, Columbia University, 630 West 168th Street, New York NY 10032, USA, Tel: 212-342-4192 Fax: 212-342-5144 rc2116@columbia.edu.

²Current: Genome Institute of Singapore, Singapore

⁷Current: Department of Human Genetics, University of Chicago, Chicago, IL

⁸Current: Centre for Molecular Epidemiology, National University of Singapore, Singapore

⁹Current: Taub Institute, Columbia University, New York, NY

Electronic-Database Information

The URLs for data presented herein are as follows:

Center for Medical Genetics, <http://research.marshfieldclinic.org/genetics/> (for the Marshfield genetic map)

Rutger's combined linkage-physical map version 2.0 build 35. http://compugen.rutgers.edu/mapomat/view_query_map.html

Gconvert program, <http://www2.qimr.edu.au/davidD/> (for estimation of allele frequencies)

FBAT, <http://www.biostat.harvard.edu/~fbat/fbat.htm> (for testing of family-based association)

SIB-PAIR, <http://www2.qimr.edu.au/davidD/#sib-pair> (for testing of intermarker LD)

SLINK, <http://linkage.rockefeller.edu/soft/slink.html> (for simulation of families)

Keywords

bipolar; linkage; genome scan; follow-up

Introduction

Bipolar disorder (BP), also known as manic-depression [MIM 125480], is a severe psychiatric disorder characterized by mood disturbances such as mania and depression. BP poses a major public health concern due to a lifetime prevalence of 0.5-1.5% and high rates of hospitalization, psychosocial impairment, substance abuse and suicide. Family, twin and adoption studies point to a strong genetic component [Taylor and others 2002], with heritability estimated at 80%. It is hoped that identification of susceptibility genes will pave the way to effective treatment and prevention measures.

Considerable research effort has been directed at uncovering the underlying genetic mechanism. In particular, numerous linkage studies, including genome-wide scans and meta-analyses, have implicated various candidate regions throughout the genome. Some results are promising, but there is considerable variability among studies (reviewed in [Baron 2002; Craddock and Forty 2006; Segurado and others 2003]). As with other common disorders, this is partly due to genetic and phenotypic complexities. A contributing factor is the variability among studies in research protocols, including procedures for diagnosis and ascertainment of individuals and families, choice of genetic and phenotypic models for linkage analysis, marker density and the population under study.

One strategy that may rectify some of these issues is a follow-up of previously identified candidate regions, using increased marker density, enlarged pedigree sample obtained from the same population, and similar analytical models. If the original linkage findings are sustained or enhanced in the follow-up study, this could be taken as evidence for the veracity of these findings. In contrast, substantial reduction in the evidence of linkage might indicate that the initial finding is in doubt or weaker than originally thought. We demonstrated the utility of this approach in previous follow-up studies of another candidate region for BP, 21q22 [Aita and others 1999; Liu and others 2001].

In a previous 10-cM genome scan [Liu and others 2003], we identified 10 chromosomal regions with significant or suggestive linkage to BP. We used 343 microsatellite markers and genotyped 373 individuals from 40 multiplex pedigrees. In the present study, we saturated these regions with 73 densely mapped microsatellite markers (average spacing ~3 cM) and genotyped 865 individuals (including the original 373 and 492 additional individuals) in 56 pedigrees (the original 40 families and 16 new pedigrees). We report that in two of the 10 regions the evidence of linkage is increased substantially, and in five other regions support for linkage is sustained. These findings are reinforced by significant family-based associations between the disease and marker loci at most of these regions. In contrast, the remaining three regions showed greatly diminished evidence for linkage.

Materials and Methods

Subjects

The present sample consists of 865 individuals (Caucasian, age 16 or over) from 56 extended pedigrees with a high density of BP. All the subjects are Caucasian. Twenty-nine of the 40 'core' pedigrees, as well as the 16 new families, were collected in the US and are of European ancestry; the remaining 11 families ('core' sample) were collected in Israel and are of Middle-Eastern extraction. To determine affection status, we used the research

diagnostic criteria (RDC) [Spitzer and others 1978]. Best-estimate consensus diagnoses were based on direct clinical interviews using the lifetime version of the schedule for affective disorders and schizophrenia (SADS-LA) [Endicott and Spitzer 1978] modified for BP, supplemented by family history and medical records. We used three standard diagnostic models for affection status: BP1 (manic syndromes, mostly bipolar I disorder), BP2 (BP1 plus bipolar II disorder with major depression) and BP3 (BP2 plus recurrent major depressive disorders). Individuals considered unaffected under any of these models (mainly 'never mentally ill') were coded as unknown in the statistical analyses.

The ascertainment, diagnostic procedures, and affection status models are described in detail elsewhere [Baron and others 1994]. The study was approved by the Columbia University-New York State Psychiatric Institute Review Board. Informed consent was obtained from all participants.

As mentioned, our previous study [Liu and others 2003] was based on a subset of the present sample (373 individuals from 40 families), which gleaned the bulk of genetic information, while minimizing genotyping cost. The algorithm used selected only affected individual and connecting lineal pedigree members, while excluding other individuals, including married-in spouses and unaffected siblings. Here we include our previous sample plus an additional 341 genotyped individuals, including married-in spouses and unaffected siblings, for a total of 714 individuals in the extended 40 'core' pedigrees. These individuals were added to minimize ambiguities in marker inheritance patterns and thus decrease the propensity for false-positive linkage findings. In addition, we include 16 new families comprised of 151 individuals. Comparative statistics on the two samples, including affection status, are given in Table 1.

Regions and Markers

The 10 regions investigated in the present follow-up study are those implicated in the previous genome-wide scan [Liu and others 2003]: 2p13, 4q31, 7q34, 8q13, 9q31, 10q21, 10q24, 13q32, 14q21, and 17q11-12.

For eight regions, we genotyped 3 to 5 flanking markers on each side of the locus with the highest LOD score in the genome scan. Two regions – 7q34 and 10q24 – were an exception because the loci with the highest LOD scores (D7S1834 and D10S169) are located towards the end of the chromosomes, and no microsatellite marker on the telomeric side was available. For these two regions, only markers on the centromeric side were selected. In total, 73 microsatellite markers were genotyped in the whole sample, including the 10 markers with the highest LOD scores from the genome scan, and 63 new ones.

The markers genotyped consist of tandem repeated dinucleotides, trinucleotides and tetranucleotides. They were selected for map position and heterozygosity. The average map distance between markers is 3.3cM. However, to ensure the highest resolution around the original linkage finding, the inter-marker distance becomes shorter when approaching the linked marker from the genome-wide scan. Map positions were obtained from Marshfield Center for Medical Genetics (http://research.marshfieldclinic.org/genetics/Map_Markers/maps/IndexMapFrames.html), and verified in Rutgers Combined Linkage-Physical Map version 2.0 build 35 (http://compugen.rutgers.edu/mapomat/view_query_map.html), as well as in deCODE genetic map [Kong and others 2002]. On the basis of optimal map distances, we selected markers with high heterozygosity (≥ 0.70 , mostly ≥ 0.80) for higher information content. Table 2 lists all markers genotyped by region.

Genotyping

The DNA was extracted from blood and lymphoblastoid cell lines by established standard methods. Semi-automated genotyping was performed with fluorescence-labeled microsatellite markers. DNA samples and PCR reagents were pipetted with a TECAN robot, and cycled in MJ Research PTC225 thermocyclers. The electrophoresis of the fluorescent PCR products was performed on ABI 377 DNA sequencers. The genotypes were detected and analyzed using GeneScan v.2.1 (<http://www.perkin-elmer.com/>). The allele sizes were called by Genotyper v. 2 (<http://www.perkin-elmer.com/>) and visually checked. Additional error checking was performed with LABMAN [Adams 1994] and all genotype information was then stored in the LABMAN database.

Linkage and Association Analyses

Genotypes were checked for Mendelian inconsistencies with the Pedcheck program [O'Connell and Weeks 1998]. Allele frequencies were estimated from all genotyped individuals using the Gconvert program (<http://www2.qimr.edu.au/davidD/>).

Two-point and multipoint parametric analyses were performed using the MLINK and LINKMAP programs (FASTLINK package), respectively [Cottingham and others 1993]. In our previous genome scan [Liu and others 2003], the highest LOD scores found in the 10 regions were obtained with various diagnostic and genetic models. For the present study, we used the same models and parameters to determine if the previous results can be verified. For both dominant and recessive genetic models, we assumed a penetrance of 0.8 for carriers based on robustness, and a phenocopy rate of 0.001. The values were originally chosen for their robustness rather than striving with the impossible task of estimating the real penetrance. The disease allele frequency was set to 0.01 and 0.10 for dominant and recessive models, respectively. Due to computational constraints, our multipoint results were based on three-point analysis (including the marker with peak two-point LOD score and two adjacent markers).

Multipoint affected sib-pair (ASP) analysis was conducted using Genehunter version 2.0 [Kruglyak and others 1996]. Multipoint LOD scores were calculated by comparing the likelihood of estimated identity-by-descent (IBD) sharing and the null distribution assuming no linkage. For ASP analysis, the large pedigrees were broken into nuclear families. Only the first sib pair from each nuclear family was analyzed for linkage.

To test for genetic heterogeneity, we computed HLOD scores and corresponding α values (fraction of linked families) using Genehunter 2.0. This was done for loci with pronounced evidence of linkage.

The genome-wide or experiment-wide significance of our linkage findings were assessed in two ways: 1) employing Lander and Kruglyak's guidelines [Lander and Kruglyak 1995]; 2) computing the experimental empirical significance of our results, using genotypic and marker allele frequencies from our data, and simulating datasets (1000 replicates) under the assumption of no linkage, using the program SLINK [Ott 1989; Weeks and others 1990].

FBAT (<http://www.biostat.harvard.edu/~fbat/fbat.htm>) was used to test for family-based association between BP and individual markers. Because the presence of inter-marker linkage disequilibrium (LD) could lead to biased results when the family genotype data are not complete, we tested for inter-marker LD using SIB-PAIR (<http://www2.qimr.edu.au/davidD/#sib-pair>) prior to the FBAT analysis. The association was evaluated assuming the presence of linkage, and under the diagnostic model that produced the linkage finding. Association was considered significant when the number of informative families was 10 or larger, and when the permuted two-sided p-value was

statistically significant ($p \leq 0.05$). For each marker with significant association, Z-scores and p-values for test of linkage disequilibrium, as well as permuted two-sided p-values, are given.

Results

Table 3 shows the maximum LOD scores for both our previous sample [Liu and others 2003] and the present, expanded data set (40 ‘core’ pedigrees, 16 new pedigrees, and all pedigrees). Two-point LOD scores are shown for parametric linkage analysis. ASP results are presented as multipoint LOD scores. Results are displayed for the original marker loci [Liu and others 2003]. Also shown are results for nearby markers (present analysis), if the associated LOD score in any of the three pedigree groups is larger than that obtained for the original marker. The maximum LOD scores in the present analysis were obtained with the same diagnostic and genetic models that produced the strongest results in the previous study, with the exception of D8S1136 at 8q13 (BP1 versus BP2 in the original analysis).

One region, 7q34 showed substantial increase of linkage evidence: the maximum LOD score was 3.53 at D7S1824 (BP2, ASP, all pedigrees, experiment-wide empirical p-value < 0.001) compared to the original score of 2.78. An even higher LOD score was observed in the “core” pedigrees: 4.52 at D7S1824 (BP2, ASP, experiment-wide empirical p-value < 0.001). The SLINK simulated p-value of 0.001 corresponds to a LOD score of 3.07 for all pedigrees and 3.16 for the “core” pedigrees, respectively. Interestingly, the parametric two-point LOD score at the same marker (1.47) is also supportive of linkage and larger than that obtained previously (1.17).

The LOD score plot for 7q34 is shown in Figure 1. Six other regions – 2p13, 4q31, 8q13, 13q32, 14q21 and 17q11 – produced LOD scores consistent with the original linkage findings, though at a lesser magnitude (Table 3). Nearby markers at 2p13, 8q13 and 13q32, 17q11 (D2S1394, D8S1136, D13S781 and D17S1303, respectively) produced larger LOD scores than the previously studied markers (all pedigrees). At 2p13, 8q13, 14q21 and 17q11, the LOD scores for the overall sample (all pedigrees) were larger than those observed for the ‘core’ pedigrees, suggesting that the new pedigrees contributed to the linkage signal. On the other hand, the LOD score at 4q31 and 13q32 (all pedigrees) was smaller than the LOD score for the ‘core’ pedigrees, indicating that the ‘core’ pedigrees made the main contribution to the evidence of linkage at this locus. It should be noted that the LOD score for core pedigrees achieved experimental-wide significance in 13q32: 3.53 at D13S779 (BP2, ASP, experiment-wide empirical p-value < 0.001). The LOD score for all pedigrees slightly decreased at the originally linked marker D13S779 (2.09 versus the original 2.20), although D13S781, an adjacent marker, yielded a larger LOD score of 2.69. The remaining three regions – 9q31, 10q21 and 10q24 – showed substantial reduction in LOD scores compared to the previous results.

The 16 new pedigrees gave LOD scores consistent with linkage at 2p13, 8q13, 14q21 and 17q11, regions with linkage evidence in the ‘core’ pedigrees. The LOD scores at other regions, including all regions with no linkage evidence in the ‘core’ pedigrees (9q31, 10q21 and 10q24), were negligible. For the most part, the new pedigrees produced smaller LOD scores compared to the ‘core’ pedigrees.

The parametric multipoint LOD scores (all pedigrees; not shown in Table 3) for 2p13, 4q31, 7q34, 8q13, 13q32 and 17q11 are 1.04, 1.84, 1.27, 2.40, 0.94 and 2.68, respectively, consistent with the hypothesis of linkage. In contrast, the multipoint parametric LOD scores for 9q31, 10q21 and 10q24 are far smaller (≤ 0.5). Linkage evidence for the latter three

regions is considerably reduced compared to the previous results [Liu and others 2003], consistent with the two-point LOD scores.

The HLOD scores for 7q34 and 13q32, the two regions with the strongest linkage signals in the present analysis, are 2.22 and 1.73, respectively. The corresponding alpha values (the fraction of linked families) are 0.52 and 0.53. Although the HLOD scores do not reach statistical significance, the fact that they exceed the parametric LOD scores for these regions suggests the presence of genetic heterogeneity, with a substantial portion of linked families (~50%).

Using FBAT test for association in presence of linkage, association was tested between BP and 55 markers in the regions with evidence for linkage (2p13, 4q31, 7q34, 8q13, 13q32, 14q21 and 17q11). Evidence was found for association in most regions except 4q31 and 8q13. Table 4 lists all individual markers showing significant association with BP. Also included are the z-scores and p-values for the linkage disequilibrium test, as well as p-values generated by permutation. The sign of the z-scores indicates the direction of the association: the alleles with positive z-scores are considered putative while those with negative z-scores are viewed as protective.

Discussion

This follow-up study expands on our first-stage genome scan [Liu and others 2003] in two ways. First, the data set is much larger, with more extensive meiotic information. The enlarged sample includes the original 40 pedigrees extended to include additional family members, and 16 new pedigrees. Second, marker density is improved for previously identified candidate regions, generating fine mapping for linkage and association analyses.

Our strongest linkage signal is at 7q34. The LOD score is substantially enhanced compared to the previous results: 3.53 (all pedigrees) and 4.52 (“core” pedigrees) versus 2.78 (previous study). Based on Lander and Kruglyak’s guidelines, the LOD score for “core” pedigrees achieves genome-wide significance, while that for all pedigrees approaches significance. The empirical p-values for these LOD scores (present analysis) are both smaller than 0.001, thus indicative of experiment-wide significance.

Linkage evidence on 7q34 seems to increase towards the telomere, and the peak signal likely lies outside of the marker range. It is possible that the telomeric effect caused inflation in the linkage scores for D7S1824. No well-established microsatellite markers are available on the telomeric side of D7S1824. Therefore, genotyping of single-nucleotide polymorphisms (SNPs) would be necessary for further evaluation of the linkage signal.

Our second strongest finding is at 13q32. As with 7q34, the LOD score at the originally linked marker, D13S779, is increased appreciably in the “core” pedigrees: 3.53 versus 2.20 in the original sample. The enhanced LOD score approaches experiment-wide significance by Lander and Kruglyak’s guidelines, and is smaller than the empirical experiment-wide p-value of 0.001. However, the maximum LOD score for the entire sample (all pedigrees) is reduced: 2.09 versus 2.20. Thus, the evidence of linkage is sustained, though somewhat weakened (empirical experiment-wide p-value <0.005). It should be acknowledged that the 16 new pedigrees, analyzed separately, show no linkage at 7q34 and 13q32. This finding, as well as the generally weaker linkage evidence in these pedigrees compared to that produced by the 40 ‘core’ pedigrees, is likely attributable to the smaller sample, i.e., inadequate statistical power. Variable marker informativeness and genetic heterogeneity, which may be present in spite of the identical ascertainment of the two sets of pedigrees, may be contributing factors..

As mentioned, the inclusion of additional individuals in the “core” pedigrees (in particular, married-in spouses and other sibs) reduces ambiguities in marker inheritance patterns, so that propensity for false-positive linkage results is reduced as well. In complex disorders such as BP, it is not uncommon that the initial linkage evidence (often false-positive) is reduced or even disappears when more individuals from the same pedigrees are genotyped. This is mainly because of the increased false-positive rates when founders are not genotyped [Goring and Terwilliger 2000]. Hence, our finding at 7q34 is all the more noteworthy not only because the LOD scores are considerably higher, but also because these results are less likely to be false-positive. Similarly, the fact that the linkage evidence at 2p13, 4q31, 8q13, 13q32, 14q21 and 17q11 is sustained (though not clearly enhanced) in the enlarged sample suggests that the original linkage findings for these regions are not false-positive. On the other hand, the substantial reduction in LOD scores at 9q31, 10q21 and 10q24 indicates that the evidence of linkage in these regions is weaker than previously thought. Not unexpectedly, linkage evidence obtained with the 16 new pedigrees is generally weaker than that produced by the 40 ‘core’ pedigrees, probably due to the smaller sample. Genetic heterogeneity and variable marker informativeness among pedigrees may be contributing factors.

The association (tested in presence of linkage) between BP and marker loci at most regions with linkage evidence provides further support for the linkage results, although none of these p-values achieved global significance using the Bonferroni correction. Although association analysis using microsatellite markers is feasible, it is limited in that the distance between these markers is usually too long to reach significant linkage disequilibrium. Thus, if a disease locus resides in the region of interest but not close enough to a typed marker, no association will be detected, even when a linkage signal is present. This could be one explanation for the lack of association at 4q31 and 8q13, in the present analysis. Additionally, the power of association tests also depends on the number of allele transmissions observed. Microsatellite markers are highly polymorphic, with many alleles of low frequencies, which could lead to diminished signal as well.

Delineation of candidate regions is not straightforward in complex disorders. Our strongest linkage signal is at D7S1824. However, non-trivial LOD scores are in evidence across a large interval of about 10 cM at 7q34. The wide boundaries make it difficult to pinpoint the exact location of putative BP genes in this region.

Our main finding on 7q is supported by other linkage studies (reviewed in [Baron 2002; Liu and others 2003]). Meta-analyses of genome scans are also supportive under some models [Segurado and others 2003]. Support for the other ‘positive’ regions has also been reported. In particular, 13q [Badner and Gershon 2002], 2p and 8q [Dick and others 2003], though the peak signal at 8q is at a more distal location, and 4q, 14q and 17q (reviewed in [Baron 2002; Liu and others 2003]). G72/G30, a promising candidate gene locus for BP, is mapped to 13q32, about 7 cM distal to our peak linkage signal. SNPs and haplotypes at this locus have been shown to be associated with BP in several independent samples [Hattori and others 2003; Schumacher and others 2004]. Also of interest are similar data on schizophrenia [Chumakov and others 2002; Schumacher and others 2004], though it is far from clear if the same locus on 13q contributes to the risk for both BP and schizophrenia. Given the wide boundaries of this candidate region, the presence of other susceptibility genes in this region is not implausible.

The fact that our prior linkage evidence for 9q and 10q [Liu and others 2003] has been attenuated in the present analysis does not necessarily mean that susceptibility loci in these regions are ruled out. Genetic heterogeneity among samples may account for linkage evidence reported elsewhere (reviewed in [Liu and others 2003; Venken and others 2005]).

Also, in a previous reanalysis of our genome scan data [Park and others 2004], we showed that linkage to 9q was enhanced when the analysis centered on a phenotypic subtype (BP in conjunction with psychosis). Fine mapping of these regions in other samples, as well as the use of alternate phenotype definitions, could cast more light on these issues. Meta-analyses of genome scans are also supportive of some of our findings. For example, Segurado and others (2003) reported modest linkage evidence in two of our promising candidate regions, 7q and 14q, although the regions flagged in the meta-analysis were much broader and the diagnostic models did not always coincide with ours. Another region, 10q, with evidence of linkage in our previous genome scan [Liu and others 2003], also showed modest linkage evidence in the meta-analysis, although the LOD scores for 10q were diminished in our current, follow-up study. It should be noted that this meta-analysis included our 39 “core” pedigrees. However, these pedigrees were but a fraction of the total sample, which consisted of 347 pedigrees from 18 genome scans. Another meta-analysis mentioned earlier [Badner and Gershon 2002], which employed a different statistical approach and a smaller number of genome scans, was supportive of linkage at 13q, our second strongest candidate region. We employed several analytical models, a common practice in the analysis of complex disorders. This may have led to inflated linkage evidence due to multiple testing. However, the prior evidence of linkage to these regions and the fact that some of these tests are not truly independent – for example, the diagnostic models are nested in each other (BP1 in BP2, and BP2 in BP3; see Material and Methods) – mitigate against a false-positive finding. Also, the highest LOD scores reported in this study were obtained with the same diagnostic and genetic models used in our genome-wide scan [Liu and others 2003].

Our strongest finding at 7q34 is not supported by genome-wide association studies (GWAS) [The Wellcome Trust Case Control Consortium 2007; Baum and others 2008; Sklar and others 2008], although one of these studies [Baum and others 2008] shows significant association at 13q33, in the general region of our second strongest signal on 13q. As discussed by Gershon and others [2008], multiple linkage scans, including the major meta-analyses of linkage results in bipolar disorder, show discrepancies from GWAS findings. Moreover, there is a lack of clear correspondence in results among the published GWAS [Gershon and others 2008]. Further study addressing design issues, statistical power and other methodological limitations is needed to resolve this issue and to weed out spurious results (Psychiatric GAWs Consortium Coordinating Committee 2009). Until such studies are done, lack of persuasive association finding does not necessarily invalidate linkage evidence, and vice versa.

In summary, using fine mapping of previously identified candidate regions for BP, we provide strong evidence for susceptibility loci on 7q34 and suggestive evidence for such loci on 2p13, 4q31, 8q13, 13q32, 14q21 and 17q11. Further study, including dense mapping with SNPs and mutation screening of candidate genes, is needed to clarify the role of these loci in BP.

Acknowledgments

We thank the families for their cooperation and support. Contributors to data collection and management include Richard Blumenthal, Lawrence Sharpe, Robert Simon, Barbara Lilliston, Melissa Smith and John Nee (Department of Psychiatry, Columbia University Medical Center), and Rachel Alexander, Kyra Kanyas and Bernard Lerer (Department of Psychiatry, Hadassah-Hebrew University Medical Center). Iresha Abeyhayake and Miguel Brito (Columbia University Genome Center) contributed to the genotyping effort. This work was supported by NIMH grants MH42535, MH43979 and MH00176, and by funds from the Columbia Genome Center and the New York State Office of Mental Health. The lead author was also supported by NIH grant MH44292.

References

- Adams P. LABMAN and LINKMAN: a data management system specifically designed for genome searches of complex diseases. *Genet Epidemiol*. 1994; 11(1):87–98. [PubMed: 8013891]
- Aita VM, Liu J, Knowles JA, Terwilliger JD, Baltazar R, Grunn A, Loth JE, Kanyas K, Lerer B, Endicott J. A comprehensive linkage analysis of chromosome 21q22 supports prior evidence for a putative bipolar affective disorder locus. *Am J Hum Genet*. 1999; 64(1):210–7. others. [PubMed: 9915960]
- Badner JA, Gershon ES. Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. *Mol Psychiatry*. 2002; 7(4):405–11. [PubMed: 11986984]
- Baron M. Manic-depression genes and the new millennium: poised for discovery. *Mol Psychiatry*. 2002; 7(4):342–58. [PubMed: 11986978]
- Baron M, Endicott J, Lerer B, Loth JE, Alexander JR, Simon R, Sharpe L, Gibbon M, Hasin D, Lilliston B. A pedigree series for mapping disease genes in bipolar affective disorder: sampling, assessment, and analytic considerations. *Psychiatr Genet*. 1994; 4(1):43–55. [PubMed: 8049902]
- Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B, Schulze TG, Cichon S, Rietschel M, Nothen MM. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Mol Psychiatry*. 2008; 13:197–207. others. [PubMed: 17486107]
- Chumakov I, Blumenfeld M, Guerassimenko O, Cavarec L, Palicio M, Abderrahim H, Bougueleret L, Barry C, Tanaka H, La Rosa P. Genetic and physiological data implicating the new human gene G72 and the gene for D-amino acid oxidase in schizophrenia. *Proc Natl Acad Sci U S A*. 2002; 99(21):13675–80. others. [PubMed: 12364586]
- Cottingham RW Jr, Idury RM, Schaffer AA. Faster sequential genetic linkage computations. *Am J Hum Genet*. 1993; 53(1):252–63. [PubMed: 8317490]
- Craddock N, Forty L. Genetics of affective (mood) disorders. *Eur J Hum Genet*. 2006; 14(6):660–8. [PubMed: 16721402]
- Dick DM, Foroud T, Flury L, Bowman ES, Miller MJ, Rau NL, Moe PR, Samavedy N, El-Mallakh R, Manji H. Genomewide linkage analyses of bipolar disorder: a new sample of 250 pedigrees from the National Institute of Mental Health Genetics Initiative. *Am J Hum Genet*. 2003; 73(1):107–14. others. [PubMed: 12772088]
- Endicott J, Spitzer RL. A diagnostic interview: the schedule for affective disorders and schizophrenia. *Arch Gen Psychiatry*. 1978; 35(7):837–44. [PubMed: 678037]
- Gershon ES, Liu C, Badner JA. Genome-wide association in bipolar. *Mol Psychiatry*. 2008; 13(1):1–2. [PubMed: 18084308]
- Goring HH, Terwilliger JD. Linkage analysis in the presence of errors I: complex-valued recombination fractions and complex phenotypes. *Am J Hum Genet*. 2000; 66(3):1095–106. [PubMed: 10712220]
- Hattori E, Liu C, Badner JA, Bonner TI, Christian SL, Maheshwari M, Detera-Wadleigh SD, Gibbs RA, Gershon ES. Polymorphisms at the G72/G30 gene locus, on 13q33, are associated with bipolar disorder in two independent pedigree series. *Am J Hum Genet*. 2003; 72(5):1131–40. [PubMed: 12647258]
- Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G. A high-resolution recombination map of the human genome. *Nat Genet*. 2002; 31(3):241–7. others. [PubMed: 12053178]
- Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet*. 1996; 58(6):1347–63. [PubMed: 8651312]
- Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet*. 1995; 11(3):241–7. [PubMed: 7581446]
- Liu J, Juo SH, Dewan A, Grunn A, Tong X, Brito M, Park N, Loth JE, Kanyas K, Lerer B. Evidence for a putative bipolar disorder locus on 2p13-16 and other potential loci on 4q31, 7q34, 8q13, 9q31, 10q21-24, 13q32, 14q21 and 17q11-12. *Mol Psychiatry*. 2003; 8(3):333–42. others. [PubMed: 12660806]

- Liu J, Juo SH, Terwilliger JD, Grunn A, Tong X, Brito M, Loth JE, Kanyas K, Lerer B, Endicott J. A follow-up linkage study supports evidence for a bipolar affective disorder locus on chromosome 21q22. *Am J Med Genet.* 2001; 105(2):189–94. others. [PubMed: 11304836]
- O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet.* 1998; 63(1):259–66. [PubMed: 9634505]
- Ott J. Statistical properties of the haplotype relative risk. *Genet Epidemiol.* 1989; 6(1):127–30. [PubMed: 2731704]
- Park N, Juo SH, Cheng R, Liu J, Loth JE, Lilliston B, Nee J, Grunn A, Kanyas K, Lerer B. Linkage analysis of psychosis in bipolar pedigrees suggests novel putative loci for bipolar disorder and shared susceptibility with schizophrenia. *Mol Psychiatry.* 2004; 9(12):1091–9. others. [PubMed: 15241432]
- Psychiatric GWAS Consortium Coordinating Committee. Genomewide association studies: history, rationale, and prospects for psychiatric disorders. *Am J Psychiatry.* 2009; 166(5):540–556. [PubMed: 19339359]
- Schumacher J, Jamra RA, Freudenberg J, Becker T, Ohlraun S, Otte AC, Tullius M, Kovalenko S, Bogaert AV, Maier W. Examination of G72 and D-amino-acid oxidase as genetic risk factors for schizophrenia and bipolar affective disorder. *Mol Psychiatry.* 2004; 9(2):203–7. others. [PubMed: 14966479]
- Segurado R, Detera-Wadleigh SD, Levinson DF, Lewis CM, Gill M, Nurnberger JI Jr. Craddock N, DePaulo JR, Baron M, Gershon ES. Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder. *Am J Hum Genet.* 2003; 73(1):49–62. others. [PubMed: 12802785]
- Sklar P, Smoller JW, Fan J, Ferraira MA, Perlis RH, Chambert K, Nimgaonkar VL, McQueen MB, Faraone SV, Kirby A. Whole-genome association study of bipolar disorder. *Mol Psychiatry.* 2008; 13:558–569. others. [PubMed: 18317468]
- Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: rationale and reliability. *Arch Gen Psychiatry.* 1978; 35(6):773–82. [PubMed: 655775]
- Taylor L, Faraone SV, Tsuang MT. Family, twin, and adoption studies of bipolar disease. *Curr Psychiatry Rep.* 2002; 4(2):130–3. [PubMed: 11914174]
- The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007; 447:661–683. [PubMed: 17554300]
- Venken T, Claes S, Sluijs S, Paterson AD, van Duijn C, Adolfsson R, Del-Favero J, Van Broeckhoven C. Genomewide scan for affective disorder susceptibility Loci in families of a northern Swedish isolated population. *Am J Hum Genet.* 2005; 76(2):237–48. [PubMed: 15614721]
- Weeks DE, Ott J, Lathrop GM. SLINK: a general simulation program for linkage analysis. 1990:A204. abstract.

Figure 1.
Multiple-point MLS scores across region 7q34. BP2 diagnostic model, first affected sib-pair in each nuclear family.

Figure 2. Multiple-point MLS scores across region 13q32. BP2 diagnostic model, first affected sib-pair in each nuclear family.

Table 1

Comparison of the samples in present study and the original genome scan

Study	Pedigrees	Genotyped individuals	Affected individuals		
			BP1	BP2	BP3
Liu et al., 2003	40	373	117	204	304
Present Study					
Core families	40	714	121	217	329
New families	16	151	32	44	59
Total	56	865	153	261	388

Table 2

Markers genotyped in present study

2p11	D2S1337	D2S337	D2S380	D2S285	DS441	D2S2113	D2S2112	D2S1394
Map position (cM)	80.4	80.69	83.88	86.02	86.79	88.15	90.29	90.82
4q31	D4S3250	D4S1579	D4S1644	D4S424	D4S1625	D4S1586	D4S3049	D4S1629
Map position (cM)	126.15	140.64	143.31	144.56	145.98	147.06	155.19	157.99
7q34	D7S1804	D7S500	D7S495	D7S684	D7S1824			
Map position (cM)	136.95	140.63	144.72	147.22	149.9			
8q13	D8S1110	D8S1763	D8S1136	D8S1117	D8S2323	D8S530	D8S279	D8S1119
Map position (cM)	67.27	78.78	82.26	85.23	87.54	90.33	91.46	96.21
9q31	D9S1690	D9S277	D9S938	D9S1832	D9S1677	D9S930		
Map position (cM)	106.63	109.85	110.93	112.6	117.37	120.04		
10q21	D10S1220	D10S539	D10S1788	D10S464	D10S1225	D10S581	D10S1646	D10S1647
Map position (cM)	70.23	72.9	75.57	79.03	80.77	82.5	86.2	88.41
10q24	D10S505	D10S1700	D10S169					
Map position (cM)	165.27	172.26	173.13					
13q32	D13S317	D13S886	D13S1241	D13S1240	D13S779	D13S158	D13S781	D13S796
Map position (cM)	63.9	70.13	76.26	81.64	82.93	84.87	87.03	93.52
14q21	D14S297	D14S741	D14S1014	D14S306	D14S1048	D14S1068	D14S587	
Map position (cM)	31.75	36.76	42.81	44.06	45.12	50.5	55.82	
17q11	D17S1303	D17S1875	D17S1803	D17S1856	D17S900	D17S921	D17S839	D17S2196
Map position (cM)	23.56	25.14	31.96	35.55	36.14	36.14	37.8	41.12
								44.62
								50.74

* The linked loci in the genome scan (Liu et al., 2003) are bold-faced.

Table 3

Summary of LOD scores

Region	Marker	Position (cM)	Diagnostic model	Genetic model	LOD scores			
					Previous Study		Present Study	
					Liu et al. 2003 (N=40)	'Core' pedigrees (N=40)	New pedigrees (N=16)	All pedigrees (N=56)
2p13	D2S441	87	BP2	Dominant	3.20*	1.54	0.73	2.09
	D2S1394	91	BP2	ASP	Not done	1.45	1.78	2.77
4q31	D4S1625	146	BP3	Dominant	3.16*	2.71*	0.25*	1.94*
7q34	D7S1824	150	BP2	ASP	2.78	4.52	0.14	3.53
8q13	D8S2323	88	BP2	Dominant	1.86	1.01	0.00	0.99
	D8S1136	82	BP1	Dominant	Not done	1.05	1.89	1.62
9q31	D9S938	111	BP1	Dominant	1.88*	1.06*	-0.02*	0.42*
10q21	D10S1225	81	BP2	Dominant	1.99	0.27	0.00	0.24
10q24	D10S169	173	BP1	Dominant	2.79*	1.00*	0.11*	0.98*
13q32	D13S779	83	BP2	ASP	2.20	3.53	0.00	2.09
	D13S781	87	BP2	ASP	Not done	2.93	0.13	2.69
14q21	D14S306	44	BP2	ASP	2.36	1.28	0.65	1.90
17q11	D17S921	36	BP1	Dominant	2.68	0.48	1.22	0.85
	D17S1303	24	BP1	Dominant	Not done	1.57	0.83	2.07

* nuclear family analysis

Table 4

Summary of p-values for family-based association

Region	Diagnostic Model	Marker	Position (cM)	Allele	Allele Freq	N of Families	Z-score	p-value	perm. p-value
2p13	BP2	D2S337	80.69	7	0.199	73	2.311	0.02085	0.021923
		D2S441	86.79	2	0.16	51	2.508	0.012143	0.013457
		D2S2113	88.15	12	0.171	60	2.325	0.020064	0.017097
7q34	BP2	D7S1804	136.95	6	0.222	79	-2.056	0.039765	0.042291
				11	0.051	22	2.294	0.021768	0.020784
		D7S684	147.22	2	0.223	83	2.358	0.01837	0.017745
		D7S1824	149.9	1	0.05	31	1.929	0.053731	0.052573
				8	0.07	33	-2.1	0.035692	0.037023
13q32	BP2	D13S779	82.93	5	0.34	65	2.3	0.021458	0.018622
14q21	BP2	D14S306	44.06	6	0.117	53	2.05	0.040409	0.040958
				10	0.075	27	-2.044	0.040951	0.043478
		D14S1048	45.12	13	0.055	18	2.646	0.008151	0.008258
				4	0.049	20	-2.553	0.010674	0.01256
		D14S1068	50.5	6	0.226	77	2.077	0.03782	0.037744
				7	0.027	17	2.443	0.014556	0.014445
				global	N/A	N/A	N/A	0.012825	0.025058
17q11	BP1	D17S900	36.14	1	0.221	46	-2.077	0.037806	0.04025
		D17S921	36.14	7	0.032	15	2.385	0.017094	0.018388
				3	0.183	49	-2.827	0.004692	0.004911
		D17S2196	44.62	5	0.098	37	2.151	0.031474	0.031252
				global	N/A	N/A	N/A	0.021706	0.031496

Family #: number of informative families (353 nuclear families in total)

perm. p-value: two-sided p-value generated by permutation (hbat -p marker_name)

global: global analysis for the marker