Am. J. Hum. Genet. 66:205–215, 2000

Full-Genome Scan for Linkage in 50 Families Segregating the Bipolar Affective Disease Phenotype

Carl Friddle,[1,*] Rebecca Koskela,[1,2,†] Koustubh Ranade,[1] Joan Hebert,[1] Michele Cargill,[1,‡] Chris D. Clark,[1] Melvin McInnis,[3] Sylvia Simpson,[3] Francis McMahon,[3,§] O. Colin Stine,[3,‖] Deborah Meyers,[3,] Jianfeng Xu,[3,‖] Dean MacKinnon,[3] Theresa Swift-Scanlan,[3] Kay Jamison,[3] Susan Folstein,[4] Mark Daly,[5] Leonid Kruglyak,[5] Thomas Marr,[2,#] J. Raymond DePaulo,[3] and David Botstein[1]

[1]Department of Genetics, Stanford University, Stanford, CA; [2]Department of Computational Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; [3]Department of Psychiatry, Johns Hopkins University School of Medicine, Baltimore; [4]Tufts University School of Medicine, Boston; and [5]Whitehead Institute for Biomedical Research, Cambridge, MA

Summary

A genome scan of ~12-cM initial resolution was done on 50 of a set of 51 carefully ascertained unilineal multiplex families segregating the bipolar affective disorder phenotype. In addition to standard multipoint linkage analysis methods, a simultaneous-search algorithm was applied in an attempt to surmount the problem of genetic heterogeneity. The results revealed no linkage across the genome. The results exclude monogenic models and make it unlikely that two genes account for the disease in this sample. These results support the conclusion that at least several hundred kindreds will be required in order to establish linkage of susceptibility loci to bipolar disorder in heterogeneous populations.

Introduction

Bipolar (BP) disorder, also known as "manic-depressive illness" (MIM 125480), is a common disease affecting ~1% of the world's population (Weissman et al. 1996). Since the earliest descriptions of BP disorder, a striking familial incidence has been noted (Kraepelin 1921). Further twin, adoption, and family studies have supported a genetic etiology, and, with the advent of modern genetic maps, this disease has become an attractive target for linkage analysis (see MacKinnon et al. 1997). Since 1968, >50 studies have attempted to identify the loci responsible for BP disorder. Although >20 loci have been implicated, few of the data in these reports could be confidently distinguished from chance findings (Dupuis et al. 1995; Kruglyak and Lander 1995).

One of the most widely studied regions is located on chromosome 18p, where Berrettini et al. (1994) first reported evidence for linkage to BP disorder. Several groups, both ours (Stine et al. 1995; McMahon et al. 1997) and others (McInnes et al. 1996; Nöthen et al. 1999), have reported evidence supporting linkage to chromosome 18, including this region. Each of these data sets provided statistical evidence reaching the suggestive level but not the level recommended (see Lander and Kruglyak 1995) as being statistically significant. A metanalysis was performed on the three data sets noted above and on two others (Knowles et al. 1998; Rice 1997). The genotypes and phenotypes of the 382 affected sib pairs in the 185 kindreds analyzed (Lin and Bale 1997) supported linkage, with statistical significance ($P = 2 \times 10^{-8}$).

Evidence for linkage to loci in 18q21-23 also has been observed by several groups (Stine et al. 1995; McInnes et al. 1996; McMahon et al. 1997; Nöthen et al. 1999). However, the number of loci, the location(s) of the loci, and the role of a parent-of-origin effect (McMahon et al. 1995) on the linkage evidence are not well defined. Evidence for linkage to 21q (Straub et al. 1994), 4p (Blackwood et al. 1996), and 12q (Barden and Morissette 1999), implicated in single kindreds selected from larger family sets, has also been supported by the results

Received February 12, 1999; accepted for publication October 26, 1999; electronically published December 28, 1999.

Address for correspondence and reprints: Dr. J. Raymond DePaulo, Department of Psychiatry, Johns Hopkins University School of Medicine, Meyer 3-181, 600 N. Wolfe Street, Baltimore, MD 21287-7381. E-mail: jrd@welchlink.welch.jhu.edu

* Present affiliation: Lawrence Berkeley National Laboratory, Berkely, CA.

† Present affiliation: Mayo Clinic, Rochester, MN.

‡ Present affiliation: Whitehead Institute for Biomedical Research, Cambridge, MA.

§ Present affiliation: University of Chicago, Chicago.

‖ Present affiliation: University of Maryland School of Medicine, Baltimore.

Present affiliation: Genomica Corporation, Boulder, CO.

0002-9297/2000/6601-0024$02.00

of some studies (Detera-Wadleigh et al. 1996; Smyth et al. 1997; Gurling et al. 1998; Aita et al. 1999; Craddock and Lendon 1999) but not by findings from the majority of independent follow-up studies. The results of a large number of studies of chromosomes 4, 12, 18, and 21 have recently been published (Curtis 1999; Detera-Wadleigh 1999; Kennedy et al. 1999; Van Broeckhoven and Verheyen 1999).

Potential reasons for difficulty in linkage mapping of BP disorder are not hard to find. Because the pathology and pathophysiology of the disorder are unknown, the diagnosis and consequent phenotypic classification of BP disorder are made solely on the basis of clinical grounds. The underlying genetic mechanisms are unknown: several genes could act together in a given individual to cause disease, or several genes, each of which causes the disease in different families (genetic heterogeneity), could be present in the population. In either circumstance, we agree with others who suggest that, without evaluation of several hundred or more kindreds with BP disorder, genomewide linkage studies will usually fail to detect or confirm linked loci. In addition, maps that are much more informative than those typically in use and analytic strategies that are explicitly designed to detect loci in complex disorders (Risch and Merikangas 1996; Visscher et al. 1998; Cox et al. 1999) might be required.

At the time when we began the ascertainment of families with multiple cases of BP disorder, we assumed that genetic heterogeneity would constitute the main complexity. This seemed surmountable with the use of linkage methods just then being reduced to practice (Ott 1991), combined with a strategy that we called "simultaneous search" (Lander and Botstein 1986*b*). In principle, with the use of data covering the entire genome, simultaneous search would evaluate loci pairwise, thereby minimizing the chance that data from a family whose disease is caused by one locus would decrease the significance of the data from families whose disease is caused by another locus. The simultaneous-search algorithm might therefore be expected to increase the power of linkage analysis to resolve two or more loci acting independently in a population. We believed that this would be particularly advantageous in an instance in which a small number of dominant loci were each segregating in a different subset of our pedigrees.

Families were ascertained through a treated proband with BP type 1 (BPI) who had two or more affected siblings or at least one affected sibling and one affected parent. In all families, one parent had to be unaffected—that is, free of recurrent unipolar depression (RUP) or BP affective disorder after being directly examined in the study. All available family members were interviewed by fully trained psychiatrists using the Schedule for Affective Disorders and Schizophrenia–Lifetime Version (SADS-L) interview. Two noninterviewing psychiatrists made best-estimate diagnoses on the basis of all available family-history medical record and interview data and by use of Research Diagnostic Criteria (Spitzer et al. 1978). They also assigned phenotypes by means of a conservative method (i.e., both psychiatrists had to agree that the case met criteria either for one of the "affected" conditions or for one of the "unaffected" diagnoses, or the case was called "phenotype unknown"). Despite the observation (Simpson et al. 1992) that most ascertained families were bilineal (i.e., affected individuals were found among the first-degree relatives of both parents), 51 apparently unilineal families were chosen for this study; one of the families was dropped as a result of allele-segregation inconsistencies. We estimated that we needed 50 families with three fully informative meioses per family, on the basis of a simulation of a simultaneous-search method, given a perfect 20-cM marker map of the genome (Lander and Botstein 1986*b*). To bias the sample toward families with the simplest modes of inheritance, we included only those families that appeared to be unilineal. Indeed, we found it extremely difficult to find sufficient numbers of families with three affected siblings (proband and two other siblings) and one clearly unaffected parent (Simpson et al. 1992), and therefore we included families with two or more affected siblings and one affected parent as well as one unaffected parent. When the ascertainment bias was ignored, the segregation of the disease phenotype in most of the 50 families resembled the expectation for simple dominant inheritance of a highly predisposing gene in each family (fig. 1).

The results of previous studies (McMahon et al. 1995) in this sample demonstrated clinical evidence for a parent-of-origin effect, with excess maternal transmission. When this effect was included in an early linkage analysis of chromosome 18 (Stine et al. 1995), the evidence for linkage was increased in paternal transmission and in paternal families (i.e., families that show predominant transmission of the disease through the father). Several studies show evidence of a parent-of-origin effect in familial BP disorder but raise questions about the nature of the effect (Gershon et al. 1996) and about the best way to include it in linkage analysis (McMahon et al. 1997; Nöthen et al. 1999). In addition, at least a few studies report no parent-of-origin effect (Kato et al. 1998). Although inclusion of this effect in this genomewide analysis would be reasonable, we chose not to do so, since it was not part of our original hypothesis, since it is not clear how to include it as a variable in the analysis, and since its inclusion would increase the df of the analysis.

Here we present the results of the experiment as originally planned: genomewide analysis, by means of both standard analysis and the simultaneous-search strategy, of 50 families with BP disorder (in this case, 50 of the

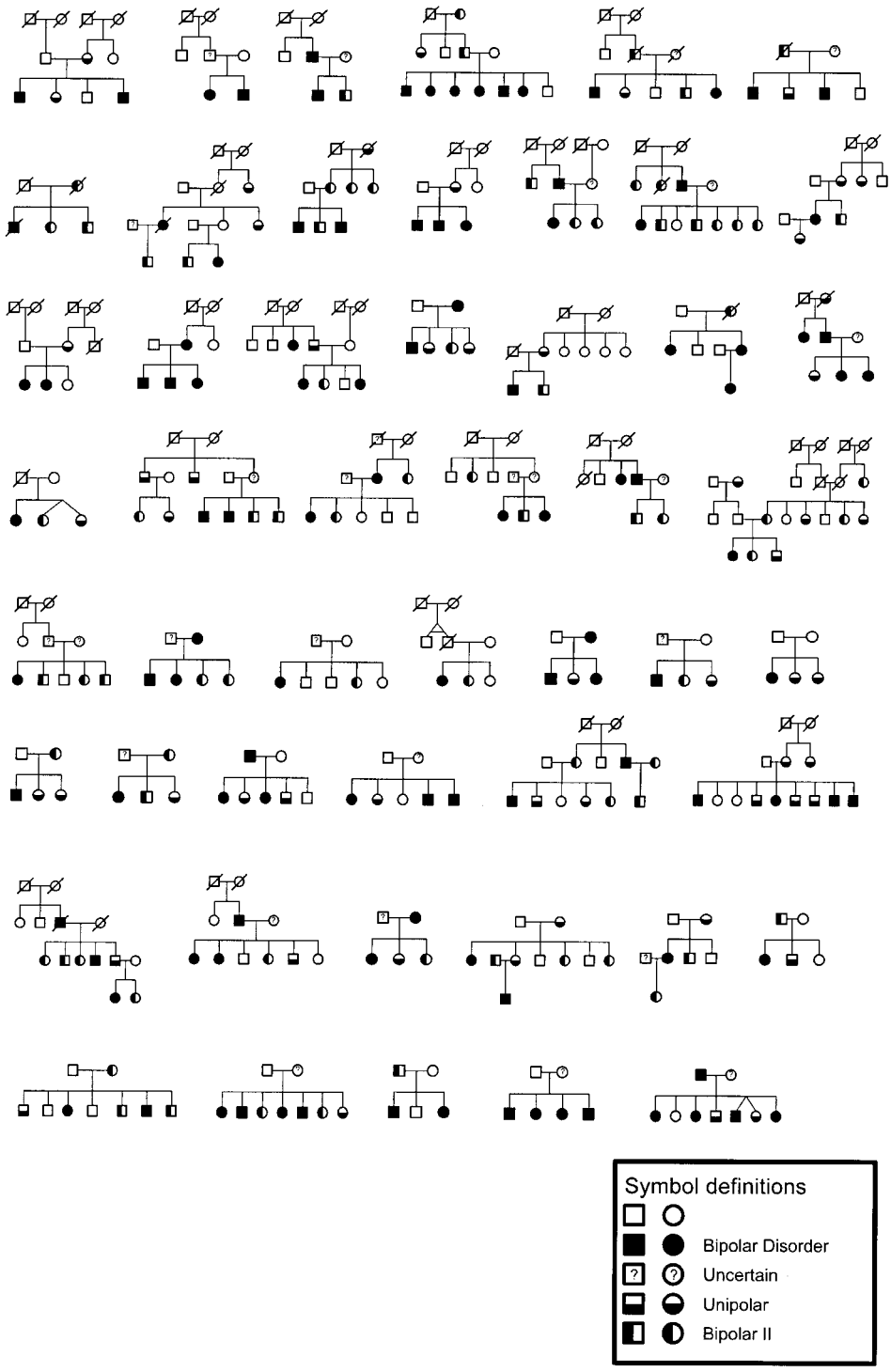


Figure 1 Pedigrees of the 50 families studied. Only the generations used for genotyping are shown, and unaffected individuals as well as those with BPI, BPII, and RUP diagnoses are shown.

58 families had been previously studied by McMahon et al. 1997). The results are sufficient to exclude—in most of the genome, including the loci previously implicated in the studies cited above—a single gene accounting for disease in half the families in this sample. A similar genome scan of 96 kindreds has been published (Edenberg et al. 1997; Rice et al. 1997; Detera-Wadleigh et al. 1997; Stine et al. 1997), with similar results. We conclude, in agreement with findings from the earlier reports, that one gene or a combination of two genes acting independently are unlikely to account for the observed familial clustering of BP disorder generally.

Subjects and Methods

Family Ascertainment

Families were ascertained through a treated proband with BPI who had either two or more affected siblings or at least one affected sibling and one affected parent (Simpson et al. 1992). In all families, one parent had to be unaffected—that is, free of RUP or BP affective disorder after being directly examined in the study. All available family members were interviewed by fully trained psychiatrists using the SADS-L interview. Two noninterviewing psychiatrists made best-estimate diagnoses on the basis of all available family-history medical record and interview data and by the use of Research Diagnostic Criteria (Spitzer et al. 1978), with the additional requirement that, for the diagnosis of BP type 2 (BPII), there be recurrent episodes of major depression. Phenotypes were assigned by means of a conservative method. Both psychiatrists had to agree that the case met criteria for one of the "affected" diagnoses (BPI, BPII, RUP, or schizoaffective manic-type [SAM]) or for one of the "unaffected" diagnoses; otherwise, the case was called "phenotype unknown." Families in which individuals with BPI, BPII, or RUP were found among both the parents or among parental first-degree relatives were excluded. Although 51 families were ascertained, the collection analyzed comprises 50 families containing 470 individuals, among whom there were 108 BPI phenotypes, 75 BPII phenotypes, and 53 RUP phenotypes. Figure 1 summarizes the family pedigrees.

Genotyping

All the genotypes reported here were determined by the multiplex method, as described elsewhere (Clark and Gschwend 1994). The first 28 families ascertained had been analyzed by linkage analysis with simple-tandem-repeat polymorphism (STRP) markers, as reported elsewhere (Stine et al. 1995). Of an additional 23 families that were later ascertained by the same methods, 22 were genotyped again, with use of an improved set of markers. On the basis of a preliminary analysis of the first 135,000 genotypes, 36 additional markers were added to improve the information in regions (with multipoint LOD score under the assumption of heterogeneity [HLOD] >1.5). This threshold was chosen on the basis of simulations that indicated that any locus capable of generating an HLOD score >3 in this sample would yield an HLOD >1.5 in our first pass of genotyping. A total of three to six markers were added to these regions to maximize the HLOD and/or information content. In all, 147,185 genotypes were determined, resulting in an average resolution of 12 cM; not all the same markers were used throughout, but all of the Cooperative Human Linkage Center–Set 6 markers were applied to all the individuals. Marker data were analyzed with the use of CRIMAP (Lander and Green 1987; Goldgar et al. 1989).

Simulation Methods

Simulations were conducted with the use of our pedigree structure and marker map, with a modification of the GENEHUNTER algorithm. We ran our simulations by placing the disease locus within a simulated map generated to match the markers used in this linkage analysis. We then varied the heterogeneity parameter α, which denotes the percentage of families in which the disease allele is causative, during the simulation. This was followed by GENEHUNTER analysis of the simulated data (with α maximized in the analysis), to determine our ability to detect genes of the percentage α used in the simulation. Several independent simulations were run for each simulated disease locus and simulated α.

Linkage Methods

Multipoint linkage analysis was done with the use of both the GENEHUNTER software package (Kruglyak et al. 1996) and an unpublished implementation (by L. Kruglyak) of simultaneous search (Lander and Botstein 1986*b*). Linkage was evaluated under four models of inheritance. The first was a dominant model that considered only individuals with BPI, BPII, or SAM as being affected. These diagnoses were distinct from RUP, on the basis of the presence of mania. The second model was a dominant model, in which individuals with BPI, BPII, SAM, or RUP were considered to be affected. The third model was a recessive model, in which individuals with BPI or BPII were considered to be affected. Similarly, the fourth model was a recessive model, in which individuals with BPI, BPII, or RUP were considered to be affected. We did not have enough individuals with BPI to analyze them separately. In each case, an age-dependent penetrance was assumed (.63–.85), as was an age-dependent phenocopy rate (10%–19%). The disease-allele frequency was set at .02 for the dominant models and at .10 for the recessive models. Allele frequencies were calculated from the data, and the recombination fractions generated by CRIMAP were used as input map distances.

GENEHUNTER calculated the multipoint HLOD and estimated the fraction of the families segregating each locus scored, as well as the information content at the position of each marker and at the midpoint between markers. The information content is a measure (range .0–1.0) of the fraction of the total inheritance information extracted by the available marker data. GENEHUNTER runs were made with α allowed to vary, to maximize the HLOD score, and also with α fixed, to estimate the power to exclude models in which one or more loci could account for the disease in 50% and 33% of the families, respectively. The HLOD scores and their positions, as generated by GENEHUNTER, were used as input for a simultaneous-search program written by Leonid Kruglyak and Eric Lander, as described elsewhere (Gschwend et al. 1996), with computer code optimized by Mark Schroeder. The simultaneous-search program evaluated linkage at all marker positions and at the midpoint between markers, for all chromosomes and all pedigrees. The parameters α_1 and α_2 (in this case, the fraction of families attributed to each potential locus) were set to .44 for the simultaneous search. This corresponds to the average α that yielded the highest HLOD scores in our genome scan.

Table 1

HLOD Scores for Simulations of Low Information Content on Chromosome 1

HLOD Score at α = .50		HLOD Score at α = .33	
At Peak	At Disease Locus	At Peak	At Disease Locus
3.9	3.9	2.6	2.0
3.1	3.1	2.4	1.6
2.5	2.5	1.9	1.9
2.4	2.4	1.6	1.6
2.1	1.4	1.5	1.3
2.1	1.1	1.1	1.1
1.8	1.2	.8	8.4
1.7	1.7	.6	.0
1.3	1.3	.3	.1
1.0	1.0	.1	.1
1.0	.5	.1	.0
.1	.0	.0	.0

NOTE.—HLOD scores are given for each of 12 simulations of a disease locus (α = .5 or .33) within a large gap between markers. Information content was .3. Peak HLOD score is given, as well as the HLOD score for the exact position of the disease locus. These scores are sorted from highest to lowest.

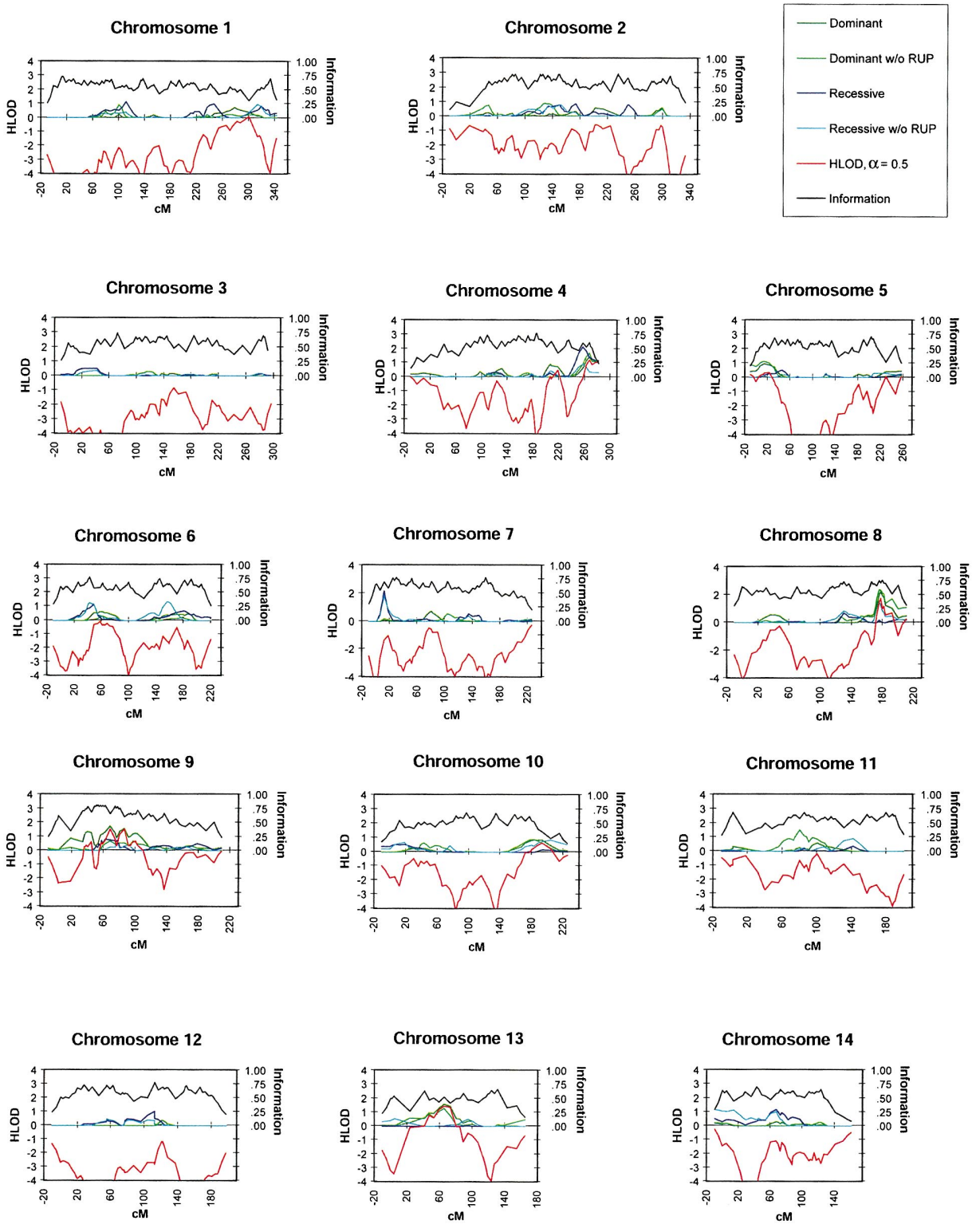
Results

Simulations

Using the actual family set and our actual markers, we performed two sets of simulations designed to test the power to detect loci for BP disorder in this sample. To determine our ability to detect genes of low α in the weakest region of our genome scan, we placed the disease locus in the middle of a 22-cM gap on chromosome 1 (information content .3) and then allowed the computer to generate the genotypes at the surrounding markers, for several values of α. Similarly, we placed a simulated disease locus at a position of good information content (.8) on chromosome 8, in an effort to determine what LOD score would be expected, in our family set, from a true linkage in a region of high information, such as that which we obtained when adding additional nearby markers.

Simulations placing the disease locus in the middle of the gap on chromosome 1, with an α of 1.0 (i.e., no heterogeneity), produced a LOD-score range of 9–10. Thus, there is little question that, with this family and marker set, we would have detected linkage if a single gene anywhere in the genome were causing most of the disease in this population. With α at .5, the peak-LOD-score range was 0.1–3.9 (table 1). Of 12 simulations, 4 (33%) yielded peak LOD scores >1.5, which is the threshold that we used for application of additional markers in the neighborhood; this finding suggests that, in a future scan, one might want to use a slightly lower threshold. Only 1 (8%) of 12 simulations had a LOD score that fell below the score of 0.4 that was actually observed in this interval (fig. 2). To determine our ability to detect a less-common locus, we simulated a disease locus causing disease in only one-third (i.e., α = .33) of the families. Seven (58%) of 12 simulations failed to yield a LOD score >1.5, which is the threshold used for secondary genotyping (table 1). Thus, where our information is weakest, we would generally have detected a locus accounting for disease in half the families, but we would have succeeded only half the time if the locus accounted for disease in only one-third of the families.

The simulations in the gap on chromosome 1 estimate the likelihood that a true disease locus at a given α in our weakest map position would be detected in the first round of genotyping. Once a disease locus was detected (i.e., provided a LOD score >1.5) in the first analysis, we added additional markers. This provided an information density similar to that in the region that we simulated on chromosome 8, where the disease locus was placed in a region where the marker spacing is 4 cM. In the case of no heterogeneity (α = 1), the LOD scores were 16–20. At α = .5, the LOD-score range was 3.4–8.7 (table 2). This indicates that we would have observed a significant LOD score (11 successes in 11 tries) for any true locus that passed the original screening at LOD >1.5. We then addressed the issue of detection of a less common locus, by setting α at .33. Five (42%) of 12 runs gave us a LOD score >3.0, indicating that identification of a locus responsible for disease in one-third of the pedigrees would have been difficult (table 2). The power would have been lower in samples with higher


Chromosome 1
Chromosome 2
Dominant
Dominant w/o RUP
Recessive
Recessive w/o RUP
HLOD, α = 0.5
Information
Chromosome 3
Chromosome 4
Chromosome 5
Chromosome 6
Chromosome 7
Chromosome 8
Chromosome 9
Chromosome 10
Chromosome 11
Chromosome 12
Chromosome 13
Chromosome 14
HLOD
Information
cM

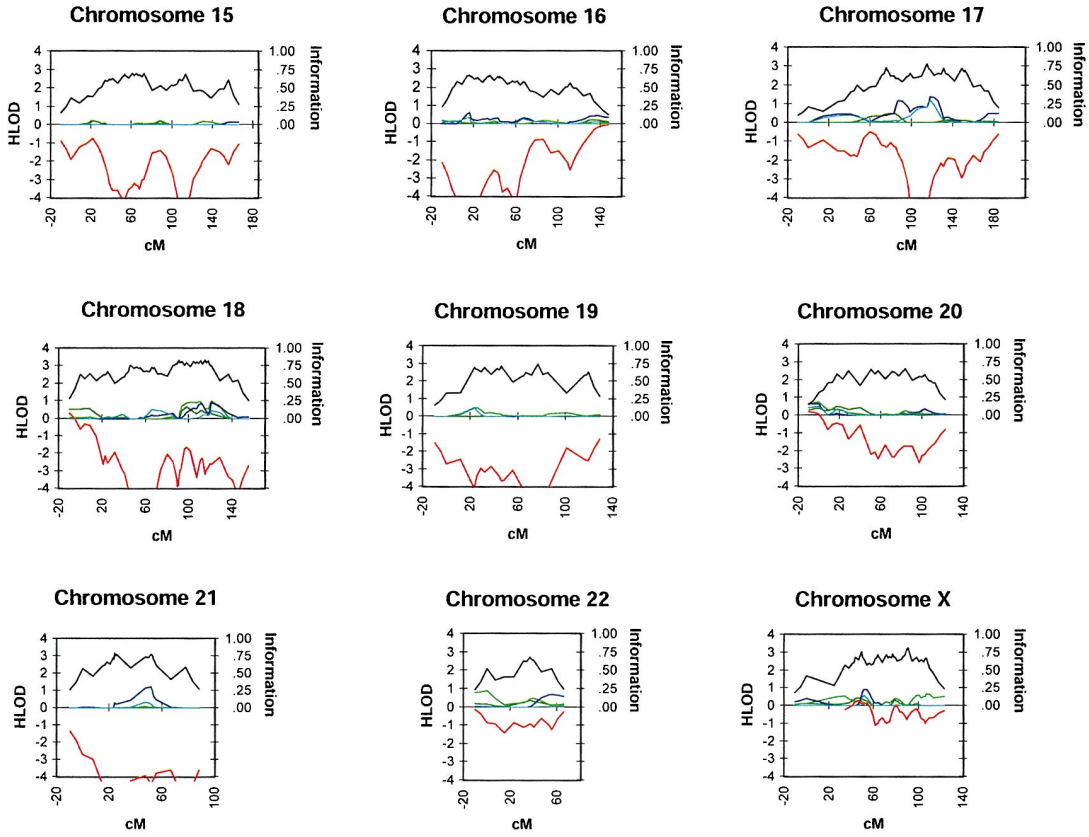


Figure 2 Results from GENEHUNTER analysis. Each chromosome is separately plotted. The information content (scale in black on right-side ordinate) at each marker and at the midpoint between markers is shown. Also shown are the HLOD scores (scale on the leftside ordinate) for four models (see text), as well as for $\alpha = .5$ (to estimate exclusion), again evaluated at markers and at the midpoint between markers.

heterogeneity. For example, in the linkage analysis of 58 families with BP (including the 50 families reported here), the maximal HLODs obtained at loci in 18q21-22 were associated with an α value of 25% (McMahon et al 1997). Thus, the estimates for the likelihood of detection and the power to exclude agree. They show that this family set should have been sufficient for detection of a gene that accounts for disease in $\geq$50% of the families. We would very likely have missed a gene responsible for disease in one-third of the families.

Linkage-Analysis Results

Several statistics were obtained from the linkage analysis, by use of GENEHUNTER (Kruglyak et al. 1996); a summary of these statistics, over the entire genetic length of the human genome, is given infigure 2. First, the position of each marker is given, along with the information content (see the Subjects and Methods section). Second, the multipoint HLODs from GENEHUNTER are given. The results are shown for four models (for details, see the Subjects and Methods section): a dominant model (with heterogeneity), in which only individuals with BPI, BPII, or SAM were considered to be affected; the same dominant model, in which only individuals with BPI, BPII, SAM, or RUP were considered to be affected; a recessive model (with heterogeneity), in which only individuals with only BPI, BPII, or SAM were considered to be affected; and the same recessive model, in which only individuals with BPI, BPI I, SAM, or RUP were considered to be affected. In these small, nuclear families, the nonparametric-linkage statistics closely followed the HLOD statistics.

Examination offigure 2 shows three peaks with HLOD scores >2. It should be recalled, in this context, that the threshold for significance for genome scans is $\geq$3.3, although a consensus is emerging in favor of 3.6 (Kruglyak et al. 1996). The detailed data for the peaks are given in table 3. The peaks may have arisen by chance alone, as might be expected at this level of significance (Dupuis et al. 1995; Lander and Kruglyak 1995). They almost certainly reflect that our assessment of lineality, albeit painstaking, was done on the basis of an insufficiently complex model of the genetics of BP disorder.

Table 2

LOD Scores for Simulations of High Information Content on Chromosome 8

LOD Score at α = .50		LOD Score at α = .33	
At Peak	At Disease Locus	At Peak	At Disease Locus
8.7	7.6	7.1	6.0
6.6	6.5	5.7	5.3
6.3	5.4	5.6	5.5
5.3	5.3	5.4	4.8
5.3	4.5	5.2	5.2
4.3	4.3	3.9	3.3
4.3	4.3	3.4	2.8
4.2	3.7	2.3	2.3
4.1	4.1	1.0	.9
4.0	4.0	1.0	.8
3.4	3.4	.9	.7
		.7	.6

NOTE.—HLOD scores are given for each simulation of a disease locus (α = .50 or .33) within a 4-cM gap between markers (information content .8). Both the peak LOD score and the LOD score for the exact position of the disease locus are given. These scores are sorted from highest to lowest.

Simultaneous-Search Results

A major part of the motivation for this study was the possibility that simultaneous search (Lander and Botstein 1986*a*, 1986*b*) might, under restricted circumstances, allow resolution of two or more loci in the case of heterogeneity, by separation of families in which disease is caused by one locus from those in which disease is caused by the other locus. Simple calculations under heterogeneity, such as those done forfigure 2, might make use of less information than is inherent in the data. We calculated the LOD scores over the genome by simultaneous search using two dominant loci or two recessive loci. The highest points in the LOD-plot range are 3.4–3.8 at the locus pair D8S272 and D9S264 and at the locus pair D9S264 and D13S151. This is near the level of significance for the second locus in a conditional search and is therefore well below the threshold for an unconditional search (Dupuis et al. 1995; Gschwend et al. 1996).

Exclusion Analysis

Since we succeeded, with use of our markers, in extraction of 50%–80% of the genetic information over virtually all of the genome, we attempted to estimate how strongly one might be able to exclude models in which one or more loci could account for a substantial fraction of the phenotype observed in the 50 families. We did this in two ways. One of these ways, which is a feature of GENEHUNTER (Kruglyak et al. 1996), is to set the parameter α (i.e., the fraction of families attributed to a potential locus) so that the resulting HLOD score estimates the likelihood that a gene at a given locus can account for that fraction of the families. The results when α was set at .5 are plotted infigure 2. It is clear that most of the genome (except for the X chromosome and parts of chromosomes 9 and 13) is unlikely to contain a gene that is causative for half of the families (i.e., HLOD is negative, generally >-1). The conventional criterion for exclusion (i.e., LOD >-2) is met for approximately half the genome). We also did the calculations for other values of α, with the result that, with

Table 3

Multipoint HLOD Peaks >2.0

Flanking Marker(s)[a]	Position[b] (cM)	Information Content[c]	HLOD	α[d]	Model[e]
D4S408, D4S426	258.01	.60	2.11	.41	Recessive
	267.93		1.70	.46	Dominant w/o RUP
	267.93		1.45	.34	Dominant
	258.01		.91	.30	Recessive w/o RUP
D7S513	12.27	.69	2.17	.32	Recessive
			1.77	.31	Recessive w/o RUP
			.16	.11	Dominant
			.05	.06	Dominant w/o RUP
D8S256, D8S272	177.88	.71	2.39	.42	Dominant w/o RUP
			2.23	.32	Dominant
			1.24	.29	Recessive w/o RUP
			.08	.07	Recessive

[a] Located at or to either side of the peak HLOD.
[b] From the pter of the chromosome.
[c] Measure of how well the inheritance of marker loci is known in the families.
[d] Fraction of linked families that gives the maximum HLOD.
[e] Mode of disease inheritance under which the data were analyzed. "w/o RUP" indicates that individuals with RUP were classified as being unaffected.

α at .3, the fraction of the genome that is excluded is substantially less.

Discussion

Our results show that no single locus or pair of loci can account for the disease in a substantial fraction of the 50 families that we studied. The level of information provided by the density and heterozygosity of the markers was sufficient to exclude from most of the genome—including both the loci implicated in previous studies and 18q21 (where our data are consistent with those of McMahon et al. [1997], when we analyze the markers used in both studies)—such a causative gene accounting for disease in half the families.

How then can we account for the generally negative results of this analysis and for the strongly familial occurrence of BP disorder? One possibility is extreme heterogeneity—namely, that the disease is frequently caused by a single dominant gene but that there are many alternative such genes segregating in the population. The observation, by Blackwood et al. (1996), of a strong indication of linkage (LOD score 4.1) in only 1 of 12 comparable families, is consistent with this view. This would not, however, account for the frequent bilineality of BP disorder, as defined in our ascertainment (Simpson et al. 1992).

An alternative genetic model, which may produce pseudodominant segregation in some cases and bilineality in others, is the "complex inheritance" or as a "quantitative-trait model," in which a number of different, relatively common disease alleles, when combined in a single individual, predispose to BP disorder. Such disease alleles could contribute relative risks of $\leqslant 5$ or much less than 2. Such a circumstance would generally produce bilineality, could show a 1:1 ratio in affected sibships, and might even occasionally result in a rare family (Straub et al. 1994; Blackwood et al. 1996; Barden and Morissette 1999; Craddock and Lendon 1999) in which a particular allele could be followed by linkage methods, by use of a simple dominant model of BP disorder. In general, however, both the number of genes involved in individuals with a common condition such as BP disorder and the consequent high frequency of them all would defeat linkage analysis of the kind that we have performed.

We conclude that any underlying genetic etiology of BP disorder is too complex to be resolved in genome scans with single-major-locus assumptions and in linkage analyses involving fewer than many hundreds of multiplex families. Since the data on the subset of 28 families reported by Stine et al. (1995) were found to be useful to analysts and generated more than a dozen published papers, with this publication we are making all genetic data, including all 147,000 genotypes collected, available to the scientific community. The data can be found at the Dana Bipolar Project Web site (for password access, contact J.R.D.).

Acknowledgments

This work was supported by grants from the Charles A. Dana Foundation Consortium on the Genetic Basis of Manic Depressive Illness, the National Institutes of Mental Health, the National Alliance for Research on Schizophrenia and Depression, and the Theodore and Vada Stanley Foundation and by contributors to the Affective Disorders Fund and the George Browne Laboratory Fund at Johns Hopkins University. We thank Paul McHugh and James Watson for their contributions to the planning of this study and for helpful criticism. We thank the residents and faculty of the Johns Hopkins University Department of Psychiatry, who referred families for this study. Finally, we thank the many families without whose collaboration this study would not have been possible.

Electronic-Database Information

The accession number and URLs for data in this article are as follows:

Cooperative Human Linkage Center, The, http://lpg.nci.nih.gov/CHLC/

Dana Bipolar Project, http://www.cshl.org/dana (for password access, contact J.R.D.)

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for BP disorder [MIM 125480])

References

Aita VM, Liu J, Knowles JA, Terwilliger JD, Baltazar R, Grunn A, Loth JE, et al (1999) A comprehensive linkage analysis of chromosome 21q22 supports prior evidence for a putative bipolar affective disorder locus. Am J Hum Genet 64:210–217

Barden N, Morissette J (1999) Chromosome 13 workshop report. Am J Med Genet 88:260–262

Berrettini WH, Ferraro TN, Goldin LR, Weeks DE, Detera-Wadleigh S, Nurnberger JI Jr, Gershon ES (1994) Chromosome 18 DNA markers and manic-depressive illness: evidence for a susceptibility gene. Proc Natl Acad Sci USA 91: 5918–5921

Blackwood DH, He L, Morris SW, McLean A, Whitton C, Thomson M, Walker MT, et al (1996) A locus for bipolar affective disorder on chromosome 4p. Nat Genet 12:427–430

Clark CD, Gschwend M (1994) Nonradioactive multiplex analysis of SSLPs. In: Dracopoli NC (ed) Current protocols in human genetics: PCR methods of genotyping: unit 2.5. John Wiley & Sons, New York, pp 2.5.7–2.5.16

Cox NJ, Frigge M, Nicolae DL, Concannon P, Hanis CL, Bell

GI, Kong A (1999) Loci on chromosomes 2 (NIDDM1) and 15 interact to increase susceptibility to diabetes in Mexican Americans. Nat Genet 21:213–215

Craddock N, Lendon C (1999) Chromosome workshop: chromosomes 11, 14, and 15. Am J Med Genet 88:244–254

Curtis D (1999) Chromosome 21 workshop. Am J Med Genet 88:272–275

Detera-Wadleigh SD (1999) Chromosomes 12 and 16 workshop. Am J Med Genet 88:255–259

Detera-Wadleigh SD, Badner JA, Goldin LR, Berrettini WH, Sanders AR, Rollins DY, Turner G, et al (1996) Affected-sib-pair analyses reveal support of prior evidence for a susceptibility locus for bipolar disorder, on 21q. Am J Hum Genet 58:1279–1285

Detera-Wadleigh SD, Badner JA, Yoshikawa T, Sanders AR, Goldin LR, Turner G, Rollins DY, et al (1997) Initial genome scan of the NIMH genetics initiative bipolar pedigrees: chromosomes 4, 7, 9, 18, 19, 20, and 21q. Am J Med Genet 74:254–262

Dupuis J, Brown PO, Siegmund D (1995) Statistical methods for linkage analysis of complex traits from high-resolution maps of identity by descent. Genetics 140:843–856

Edenberg HJ, Foroud T, Conneally PM, Sorbel JJ, Carr K, Crose C, Willig C, et al (1997) Initial genomic scan of the NIMH genetics initiative bipolar pedigrees: chromosomes 3, 5, 15, 16, 17, and 22. Am J Med Genet 74:238–246

Gershon ES, Badner JA, Detera-Wadleigh SD, Ferraro TN, Berrettini WH (1996) Maternal inheritance and chromosome 18 allele sharing in unilineal bipolar illness pedigrees. Am J Med Genet 67:202–207

Goldgar DE, Green P, Parry DM, Mulvihill JJ (1989) Multipoint linkage analysis in neurofibromatosis type I: an international collaboration. Am J Hum Genet 44:6–12

Gschwend M, Levran O, Kruglyak L, Ranade K, Verlander PC, Shen S, Faure S, et al (1996) A locus for Fanconi anemia on 16q determined by homozygosity mapping. Am J Hum Genet 59:377–384

Gurling H (1998) Chromosome 21 workshop. Psychiatr Genet 8:109–113

Kato T, Winokur G, Coryell W, Rice J, Endicott J, Keller MB, Akiskal HS (1998) Failure to demonstrate parent-of-origin effect in transmission of bipolar II disorder. J Affect Disord 50:135–141

Kennedy JL, Basile VS, Macciardi FM (1999) Chromosome 4 Workshop summary: Sixth World Congress on Psychiatric Genetics, Bonn, Germany, October 6–10, 1998. Am J Med Genet 88:224–228

Knowles JA, Rao PA, Cox-Matise T, Loth JE, de Jesus GM, Levine L, Das K, et al (1998) No evidence for significant linkage between bipolar affective disorder and chromosome 18 pericentromeric markers in a large series of multiplex extended pedigrees. Am J Hum Genet 62:916–924

Kraepelin E (1921) Manic-depressive insanity and paranoia, translated by RM Barclay. Ayer Publishing, Salem, NH

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58:1347–1363

Kruglyak L, Lander ES (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet 57:439–454

Lander ES, Botstein D (1986*a*) Mapping complex genetic traits in humans: new methods using a complete RFLP linkage map. Cold Spring Harb Symp Quant Biol 51:49–62

——— (1986*b*) Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms. Proc Natl Acad Sci USA 83: 7353–7357

Lander ES, Green P (1987) Construction of multilocus genetic linkage maps in humans. Proc Natl Acad Sci USA 84:2363–2367

Lander ES, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247

Lin JP, Bale SJ (1997) Parental transmission and D18S37 allele sharing in bipolar affective disorder. Genet Epidemiol 14: 665–668

MacKinnon DF, Jamison KR, DePaulo JR (1997) Genetics of manic depressive illness. Annu Rev Neurosci 20:355–373

McInnes LA, Escamilla MA, Service SK, Reus VI, Leon P, Silva S, Rojas E, et al (1996) A complete genome screen for genes predisposing to severe bipolar disorder in two Costa Rican pedigrees. Proc Natl Acad Sci USA 93:13060–13065

McMahon FJ, Hopkins PJ, Xu J, McInnis MG, Shaw S, Cardon L, Simpson SG, et al (1997) Linkage of bipolar affective disorder to chromosome 18 markers in a new pedigree series. Am J Hum Genet 61:1397–1404

McMahon FJ, Stine OC, Meyers DA, Simpson SG, DePaulo JR (1995) Patterns of maternal transmission in bipolar affective disorder. Am J Hum Genet 56:1277–1286

Nöthen MM, Cichon S, Rohleder H, Hemmer S, Franzek E, Fritze J, Albus M, et al (1999) Evaluation of linkage of bipolar affective disorder to chromosome 18 in a sample of 57 German families. Mol Psychiatry 4:76–84

Ott J (1991) Analysis of human genetic linkage, rev ed, trans. Johns Hopkins University Press, Baltimore

Rice J (1997) Genetic analysis of bipolar disorder: summary of GAW10. Genet Epidemiol 14:549–561

Rice JP, Goate A, Williams JT, Bierut L, Dorr D, Wu W, Shears S, et al (1997) Initial genome scan of the NIMH genetics initiative bipolar pedigrees: chromosomes 1, 6, 8, 10, and 12. Am J Med Genet 74:247–253

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. Science 273:1516–1517

Simpson SG, Folstein SE, Meyers DA, DePaulo JR (1992) Assessment of lineality in bipolar I linkage studies. Am J Psychiatry 149:1660–1665

Smyth CS, Kalsi G, Curtis D, Brynjolfsson J, O'Neill J, Rifkin L, Moloney E, et al (1997) Two-locus admixture linkage analysis of bipolar and unipolar disorder supports the presence of susceptibility loci on chromosomes 11p15 and 21q22. Genomics 39:271–278

Spitzer RL, Endicott J, Robins E (1978) Research diagnostic criteria: rationale and reliability. Arch Gen Psychiatry 35: 773–782

Stine OC, McMahon FJ, Chen L, Xu J, Meyers DA, MacKinnon DF, Simpson S, et al (1997) Initial genome screen for bipolar disorder in the NIMH genetics initiative pedigrees: chromosomes 2, 11, 13, 14, and X. Am J Med Genet 74:263–269

Stine OC, Xu J, Koskela R, Mcmahon FJ, Gschwend M, Frid-

dle C, Clark CD, et al (1995) Evidence for linkage of bipolar disorder to chromosome 18 with a parent-of-origin effect. Am J Hum Genet 57:1384–1394

Straub RE, Lehner T, Luo Y, Loth JE, Shao W, Sharpe L, Alexander JR, et al (1994) A possible vulnerability locus for bipolar affective disorder on chromosome 21q22.3. Nat Genet 8:291–296

Van Broeckhoven C, Verheyen G (1999) Report of the chromosome 18 workshop. Am J Med Genet 88:263–270

Visscher PM (1998) On the sampling variance of intraclass correlations and genetic correlations. Genetics 149:1605–1614

Weissman MM, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG, Joyce PR, et al (1996) Cross-national epidemiology of major depression and bipolar disorder. JAMA 276:293–299