Psychiatric Genetics 1999; 9:75–84

Detecting QTLs for uni- and bipolar disorder using a variance component method

P.M. Visscher[a], C.S. Haley[b], S.C. Heath[c], W.J. Muir[d] and D.H.R. Blackwood[d]

[a]*University of Edinburgh, Division of Biological Sciences, Edinburgh, UK;* [b]*Roslin Institute, Roslin, Midlothian, UK;* [c]*Rockefeller University, New York, USA;* [d]*University of Edinburgh, Department of Psychiatry, Royal Edinburgh Hospital, Edinburgh, UK*

Correspondence to P.M. Visscher, University of Edinburgh, Division of Biological Sciences, West Mains Road, Edinburgh, EH9 3JG, UK. E-mail: peter.visscher@ed.ac.uk

Received 16 October 1998; accepted 17 February 1999

The objective of this study was to use a robust variance component method to analyse unipolar and bipolar disorder in a large Scottish extended family ($n = 168$) in which linkage between markers and disease has been previously reported on the short arm of chromosome 4. Data consisted of diagnosed clinical uni- or bipolar disorder on 143 individuals, with microsatellite marker information on 109 of these individuals. The incidence of unipolar and bipolar disorder in the family was 17/143, and 11/143, respectively. Eleven linked markers on chromosome 4, spanning a region of approximately 26 cM, were used in the analysis. The statistical analysis was performed in two steps. First, pairwise identify-by-descent (IBD) coefficients for all individuals in the pedigree were calculated at 1 cM intervals, using all marker data simultaneously, with a Monte Carlo Markov Chain algorithm. Second, the variance in the trait of interest was partitioned using residual maximum likelihood (REML). Three components of variance were estimated: (i) a genetic component associated with the average relationship between individuals using the numerator relationship matrix, (ii) a genetic component associated with a chromosome location using the estimated IBD coefficients, and (iii) a residual component. The test statistic (LOD score) was calculated from the maximum likelihood of the full model, fitting all three variance components, and the maximum likelihood value from the reduced model, fitting a polygenic and residual component. The largest LOD scores (maximum LOD = 5.9), were found in a region spanning about 10 cM, when the trait was defined as the occurrence of either uni- or bipolar disorder. The putative QTL explained about 25% of the total variation in the trait. © **1999 Lippincott Williams & Wilkins.**

Keywords: unipolar disorder, bipolar disorder, genetics, maximum likelihood, quantitative trait locus

INTRODUCTION

Most psychiatric disorders fall into the category of 'complex traits', i.e. traits which are influenced by multiple genetic loci and by non-genetic environmental factors, when we seek to understand the incidence and variation of disorders between and within populations. Understanding the underlying genetics of complex traits is difficult because there is no direct correspondence between genotype and phenotype, and so individuals cannot be divided into genotypic classes. However, with a large enough number of individuals with phenotypic records and pedigree information, the genetic contribution to individual variation can be estimated. With the addition of genetic marker information, linkage analyses can be employed to search for genome regions which are co-segregating with disorders in families. However, published results from genome scans have been somewhat controversial, because of the lack of confirmation of putative genome regions affecting disorders in other populations (Risch and Botstein, 1996). It was suggested by these authors that the lack of success of finding trait loci for psychiatric disorders was due to the publication of results with large false positive rates, because of the methodology used to detect linkage, and the inherently complex nature of psychiatric disorders.

For complex traits, results from linkage analysis are not robust to violations of assumptions regarding the putative genetic model (e.g. dominant, recessive, or additive loci) and the specification of the penetrance function (Almasy and Blangero, 1998). Recently, there has been a renewed interest in variance-based methods to detect quantitative trait loci (QTLs) for complex (quantitative) traits in outbred populations (Amos *et al.*, 1990; Xu and Atchley, 1995; Comuzzie *et al.*, 1997; Duggilari *et al.*,

1996; Grignola *et al.*, 1996; Almasy and Blangero, 1998). These methods do not specify any particular genetic model (in terms of gene action or number of loci influencing trait variation) for the trait under consideration, but look for regions of the genome which explain a significant amount of the phenotypic variation in the trait. The locations on a chromosome (or genome) which explain a significant amount of variation are then our best estimates for the locations of QTL. Variance-based methods, like regression-based sib-pair methods, are generally robust to violations of assumptions (e.g. Almasy and Blangero, 1998). The objective of this study was to analyse data on unipolar and bipolar disorder in a large Scottish extended family ($n = 168$) in which linkage between markers and disease has been previously reported on the short arm of chromosome 4 (Blackwood *et al.*, 1996), using a variance component method. In the previous analysis, which assumed that any genetic component was due to a single locus with incomplete penetrance, a two-point LOD score of 4.1 was reported for marker *D4S394* (Blackwood *et al.*, 1996).

TABLE 1. Linkage map used to calculate IBD coefficients

Marker	Distance (cM)	Location (cM)
d4s431		0
	1	
d4s2366		1
	1	
d4s3007		2
	2	
d4s394		4
	8	
d4s1582		12
	1	
d4s1599		13
	1	
d4s1605		14
	3	
d4s1602		17
	1	
d4s403		18
	4	
d4s1567		22
	4	
d4s419		26

METHODS

Data

Phenotypic data (absence or presence of uni- or bipolar disorder) on a large family was used, consisting of 143 records and 168 individuals in the pedigree. The incidence of the disorder was 17/143 (unipolar), 11/143 (bipolar), and 28/143 (uni- or bipolar).

A linkage group of 11 markers, spanning 26 cM, was used to calculate multipoint pairwise identity-by-descent (IBD) coefficients. These markers were a subset of those used by Blackwood *et al.* (1996), and were chosen because of strong support for their order using the linkage package CRI-MAP (Lander and Green, 1987). In addition to the marker genotype validation carried out by Blackwood *et al.* (1996), marker genotypes were checked for unlikely double crossovers using CRI-MAP, and set to 'unknown' if these occurred. The set of ordered markers and their relative distances are shown in Table 1.

Model

A simple linear model was assumed for all traits investigated. Traits were scored as 0 or 1 (absence or presence, respectively), and defined as (i) absence/presence of unipolar disorder, (ii) absence/presence of bipolar disorder, and (iii) absence/presence of uni- or bipolar disorder. All subjects diagnosed with bipolar disorder also met criteria for major depressive disorder and this combined trait allows for the possibility that bipolar and unipolar illness in this family are presentations of the same disorder.

The aim of variance component analyses is to partition the observed variation in the trait into causes of variation. These causes can be of an environmental or genetic nature. The variance estimation method used was restricted maximum likelihood (REML: Patterson and Thompson, 1971; Lynch and Walsh, 1998). This method is very similar to maximum likelihood (ML), but takes account of the loss in degrees of freedom due to fitted fixed effects when estimating variance components. (In essence, it divides a sum of squares by $[N-1]$ if only a mean is fitted, and not by $[N]$ as ML would do.) The difference in variance estimates between ML and REML is small if few fixed effects are fitted in the model, relative to the total number of observations. All analyses were done on the 0/1 data, because there is little evidence that fitting a generalized linear model (rather than a simple linear model) improves precision or power (e.g. Visscher *et al.*, 1996). For all analyses, the maximum likelihood value was iteratively obtained using a direct search method.

Three different analyses were carried out for each of the three traits. In each model, the only fixed effect fitted was that of sex.

Polygenic model

This model allows us to determine the strength of evidence for genetic variation for the trait within the pedigree. Marker information is not required to fit

this model. Formally, we are assuming that the trait is affected by a large number of independent loci of small effect. However, this model is relatively robust to the assumption concerning the number of loci and, with sufficient data, should provide a good estimate of the genetic contribution to variation even if only a single locus is responsible. It is important to fit this particular model, because the strength of evidence for a QTL should be relative to it, i.e. does a QTL explain a significant amount of variance over and above the amount of variance explained by fitting a polygenic model?

In matrix notation, the first model is:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e} \quad (1)$$

where $\mathbf{y}$ is the vector of observations, $\mathbf{b}$ is the vector of fixed effects, $\mathbf{X}$ is the incidence matrix linking observations to the fixed effects, $\mathbf{Z}$ is the matrix linking observations to polygenic effects, $\mathbf{u}$ is the vector of random polygenic effect and $\mathbf{e}$ is the vector of random environmental effect. The assumed mean and variance structure of the observation is:

$$E(\mathbf{y}) = \mathbf{Xb}$$
$$\text{var}(\mathbf{y}) = \mathbf{V} = \mathbf{ZAZ}'\sigma_u^2 + \mathbf{I}\sigma_e^2$$

with $\mathbf{A}$ the numerator relationship matrix (e.g. Lynch and Walsh, 1998). In a non-inbred population, the elements of matrix $\mathbf{A}$ are the expected proportion of alleles that two individuals share IBD. Typical entries in this matrix are 0.5 (parents and progeny, fullsibs), 0.25 (halfsibs), and 0.125 (first cousins). This model fits the average genomic relationship between individuals, ignoring marker information. All relationships between individuals in the pedigree are used in the analysis, by fitting the numerator relationship matrix.

The results from this model are an estimate of the heritability of the trait ($h^2 = \sigma_u^2/(\sigma_u^2 + \sigma_e^2)$), and the corresponding value of the maximum likelihood. Note that this is an estimate of the heritability within the pedigree or pedigrees analysed, not in the population as a whole.

QTL model

In this model, we use marker information to determine the strength of evidence that particular genetic locations influence trait variation. The model assumes that any genetic variation is solely due to the location being examined (i.e. no loci elsewhere influence the trait). However, the model makes no assumption about the relative contribution of this genetic effect and environmental influences.

For genomic location i, the following model was fitted:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zq} + \mathbf{e} \quad (2)$$

where $\mathbf{q}$ is the vector of random QTL effects, with variance and $\text{var}(q) = \mathbf{Q}_i\sigma_q^2$. The matrix $\mathbf{Q}_i$ was estimated separately from the marker data. In this model, the 'realized' relationship matrix at a particular location in the genome is fitted by using the proportion of alleles which are identical-by-descent (IBD) between any pair of individuals. For fully informative markers and no missing information, IBD coefficients are 0, 0.5, or 1. For example, consider a small nuclear family, in which the father has marker genotype 1/2 and the mother 3/4, with two fullsibs also genotyped. If the sibs are 1/3 and 1/3, then they share both alleles IBD, and their coefficient in the matrix is 1. However, if one sib is 1/3 and the other 2/4, their coefficient is 0. With missing data and markers which are not fully informative, the elements in the IBD matrix are estimates of the proportion of alleles shared IBD between two individuals.

To estimate the IBD matrix using data from multiple markers, both SIMWALK (Sobel and Lange, 1996) and software based on Heath (1998) and Heath and Thompson (1997) were used. These programmes use Monte Carlo Markov Chain methods to estimate the IBD matrix conditional on all marker information, and take by far the most computing time for these analyses. Results from using either package to calculate IBD coefficients were similar (results not shown), and only those pertaining to the software by Heath (1998) are presented. Marker allele frequencies are necessary to distinguish between markers that are IBD and identical-by-state (IBS), and that were determined by considering only founders with marker genotypes. For this estimation, founders were defined as individuals having one or two parents with missing marker genotypes. This approach is likely to overestimate the population frequencies of the marker alleles found in affected individuals. Estimating marker allele frequencies in this way could reduce the statistical power slightly, by increasing the probability that genome segments shared by affected individuals are IBS, rather than IBD. IBD coefficients were also calculated assuming that all alleles at any marker were at equal frequencies.

The results from this model are a 'QTL heritability' ($h^2 = \sigma_q^2/(\sigma_q^2 + \sigma_e^2)$), i.e. the amount of variance within the pedigree explained by a location on chromosome 4, and the corresponding maximum likelihood value. Results from this analysis should be closest to those from the published linkage analysis by Blackwood *et*

al. (1996). Estimation of variance components was performed at 1 cM intervals.

Polygenic and QTL model

This model assumes that the genetic contribution to the trait is due to the genomic region under examination plus an unlinked polygenic effect. All random components are fitted in the model, hence:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Zq} + \mathbf{e} \qquad (3)$$

For this analysis, both an average relationship matrix and an IBD matrix were fitted. Estimates of the polygenic heritability and QTL heritability, and the overall maximum likelihood value, are results from this analysis.

Statistical tests

Since we obtain the maximum likelihood value for each model, we can calculate a LOD score to test for the presence of polygenic and QTL variation. Corresponding to the three different models, we used three different statistical tests: (1) test for polygenic variation, by comparing the maximum likelihood for the best estimate of the polygenic heritability with the likelihood for $h^2 = 0$; (2) test for QTL variation (h_q^2 vs $h_q^2 = 0$), and (3) test for QTL variation over and above polygenic variation. Hence we compared the likelihoods for a model with h^2 and h_q^2 (i.e. model 3) with a model fitting only h^2 (i.e. model 2). This is the most appropriate test for QTL presence. For a single test at a particular genomic location, the likelihood ratio test statistic (LRT) for testing variance components is approximately distributed as a 50:50 mixture of a χ^2 variable with 1 degree of freedom and a point mass at zero, under the null hypothesis of a zero component (e.g. Almasy and Blangero, 1998). The LOD score is calculated from the LRT as LOD = LRT/4.6.

RESULTS

Figures 1–3 and Table 2 summarize the findings. There is strong evidence for genetic variation for unipolar disorder and uni- or bipolar disorder, with (polygenic) heritabilities of 0.28 and 0.37, respectively. These values are significantly different from zero. There appears no evidence for polygenic variance for bipolar disorder ($h^2 = 0.06$).

There is very strong evidence for a QTL on chromosome 4 affecting uni- or bipolar disorder. The LOD of 6.6 is larger than the values published previously (Blackwood *et al.*, 1996) for testing a QTL variance vs no genetic variance. Even when testing the QTL variance relative to polygenic variance (test 3), the LOD is still 3.7, pointing to significant QTL variation, even after taking out polygenic variation. The QTL variance explains ~30% of the total variance, whether fitting a polygenic component or not.

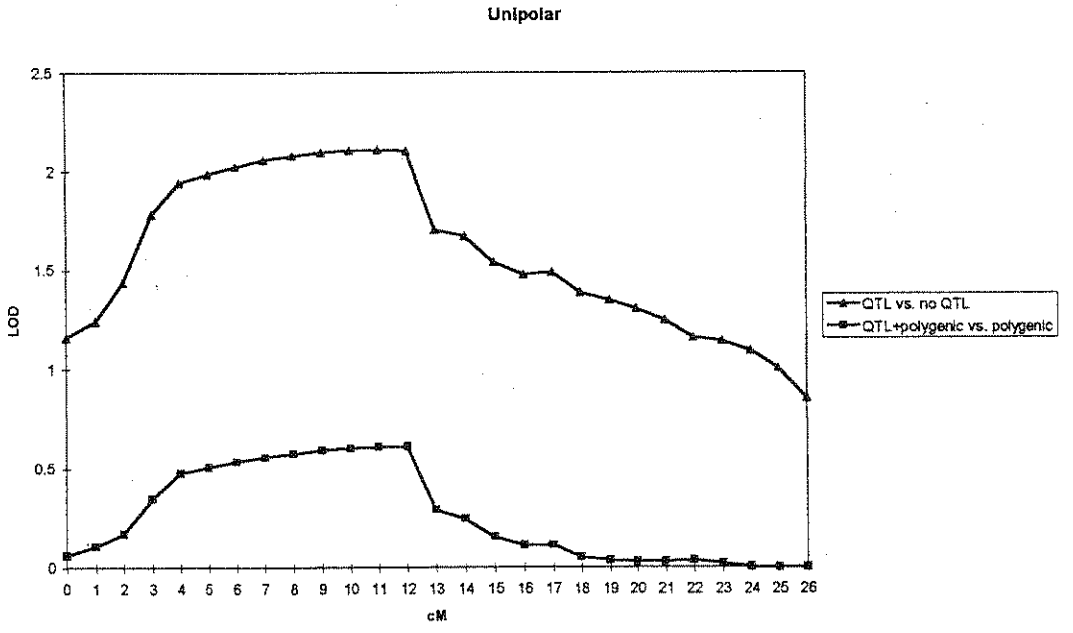

FIGURE 1. LOD scores for testing QTL variation for unipolar disorder on chromosome 4p.

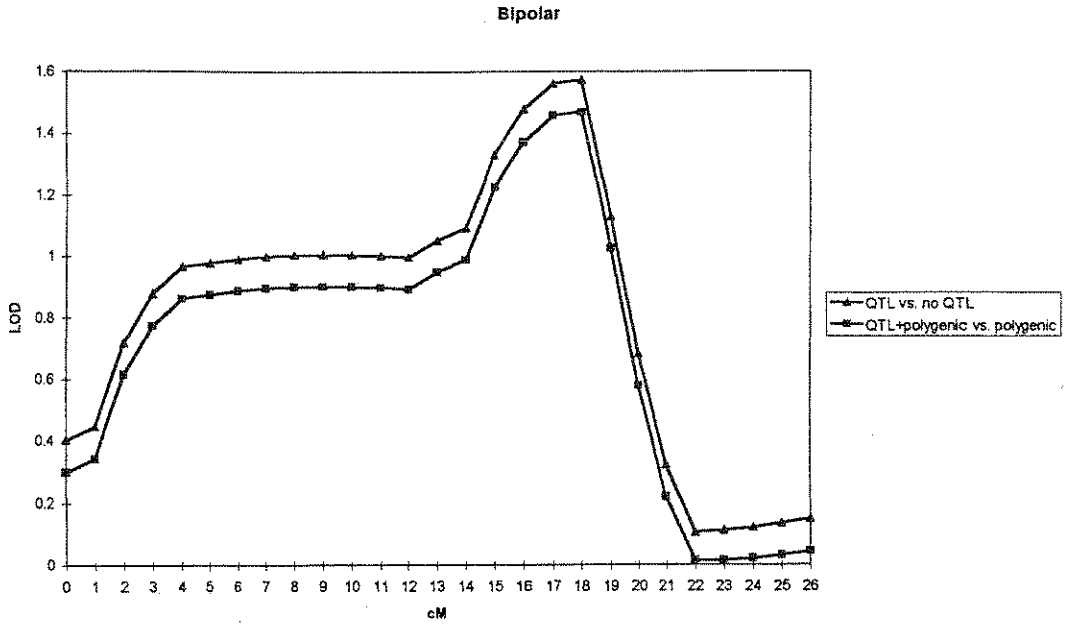

FIGURE 2. LOD scores for testing QTL variation for bipolar disorder on chromosome 4p.

There is less evidence for QTL variation for the individual (unipolar, bipolar) traits. It appears that the LOD of 2.8 for QTL variance (test 2) for unipolar disorder may be partially due to the polygenic component, since there is no significant evidence for QTL variance when a polygenic variance is fitted simultaneously. The situation for bipolar disorder is the reverse: the LOD score when fitting a QTL is hardly affected by fitting a polygenic component.

The evidence for QTL activity for uni- or bipolar disorder on chromosome 4 was stronger when it was assumed that marker alleles were at equal frequencies (Table 3, and Figures 4–6). There is little evidence for a QTL affecting unipolar disorder (test

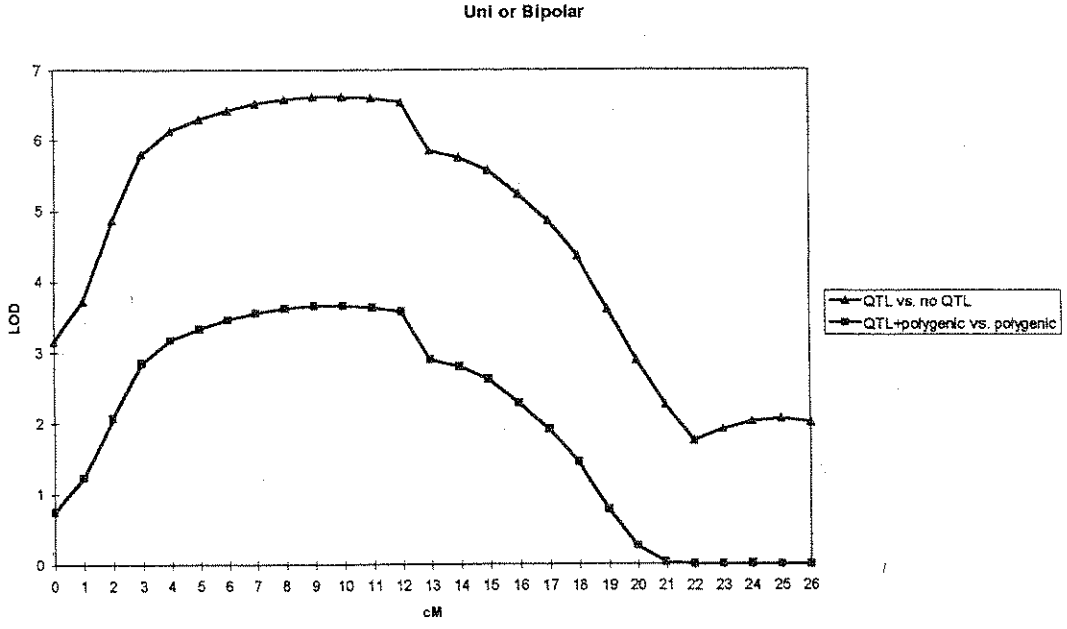

FIGURE 3. LOD scores for testing QTL variation for unipolar or bipolar disorder on chromosome 4p.

TABLE 2. Parameter estimates from the variance component analyses

Trait	Analysis[a]	$h^{2\,b}$	$h_q^{2\,c}$	LOD	Test
Unipolar disorder	I	0.28		1.6	$h^2 = 0$
	II		0.19	2.1	$h_q^2 = 0$
	III	0.12	0.16	0.6	$h_q^2 = 0$
Bipolar disorder	I	0.06		0	$h^2 = 0$
	II		0.14	1.6	$h_q^2 = 0$
	III	0.00	0.14	1.5	$h_q^2 = 0$
Uni- or bipolar disorder	I	0.37		3.0	$h^2 = 0$
	II		0.30	6.6	$h_q^2 = 0$
	III	0.00	0.29	3.7	$h_q^2 = 0$

[a]I: Fitting polygenic component only; II: fitting QTL component only; III: fitting both polygenic and QTL component.
[b]Polygenic heritability.
[c]Heritability due to QTL.

TABLE 3. Parameter estimates from the variance component analyses, when equal frequencies for alleles for a marker are used to estimate IBD coefficients

Trait	Analysis[a]	$h^{2\,b}$	h_q^{2c}	LOD	Test
Unipolar disorder	I	0.28		1.6	$h^2 = 0$
	II		0.17	2.8	$h_q^2 = 0$
	III	0.08	0.16	1.3	$h_q^2 = 0$
Bipolar disorder	I	0.06		0	$h^2 = 0$
	II		0.12	2.4	$h_q^2 = 0$
	III	0.00	0.11	2.3	$h_q^2 = 0$
Uni- or bipolar disorder	I	0.37		3.0	$h^2 = 0$
	II		0.26	8.7	$h_q^2 = 0$
	III	0.00	0.26	5.9	$h_q^2 = 0$

[a]I: Fitting polygenic component only; II: fitting QTL component only; III: fitting both polygenic and QTL component.
[b]Polygenic heritability.
[c]Heritability due to QTL.

3, LOD = 1.3), but stronger evidence for a QTL affecting bipolar disorder (LOD = 2.3) and very strong evidence for a QTL affecting either uni- or bipolar disorder (LOD = 5.9).

DISCUSSION

Significant evidence for QTL activity was detected, using a very different approach from the linkage analysis used previously (Blackwood *et al.*, 1996). This is encouraging, since the method used is more robust and makes fewer assumptions about the genetic aetiology of the trait. The advantage of a variance component method is first that no genetic model is assumed prior to the statistical analysis, and second that the correct test for the presence of a QTL

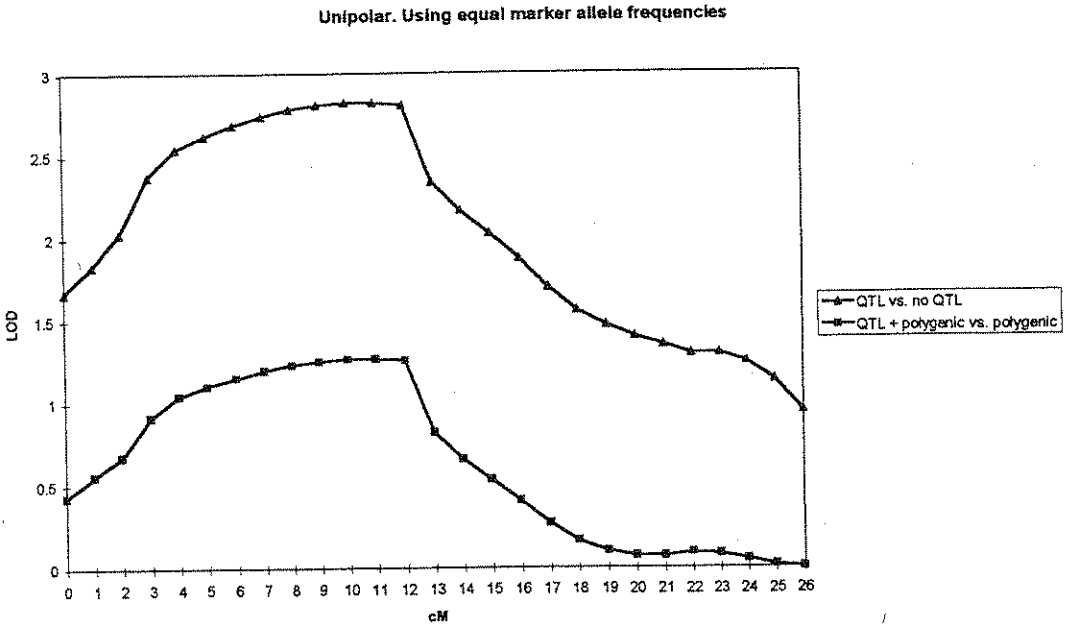

FIGURE 4. LOD scores for testing QTL variation for unipolar disorder, when the IBD coefficients are estimated, assuming that all alleles for a marker have the same frequency on chromosome 4p.

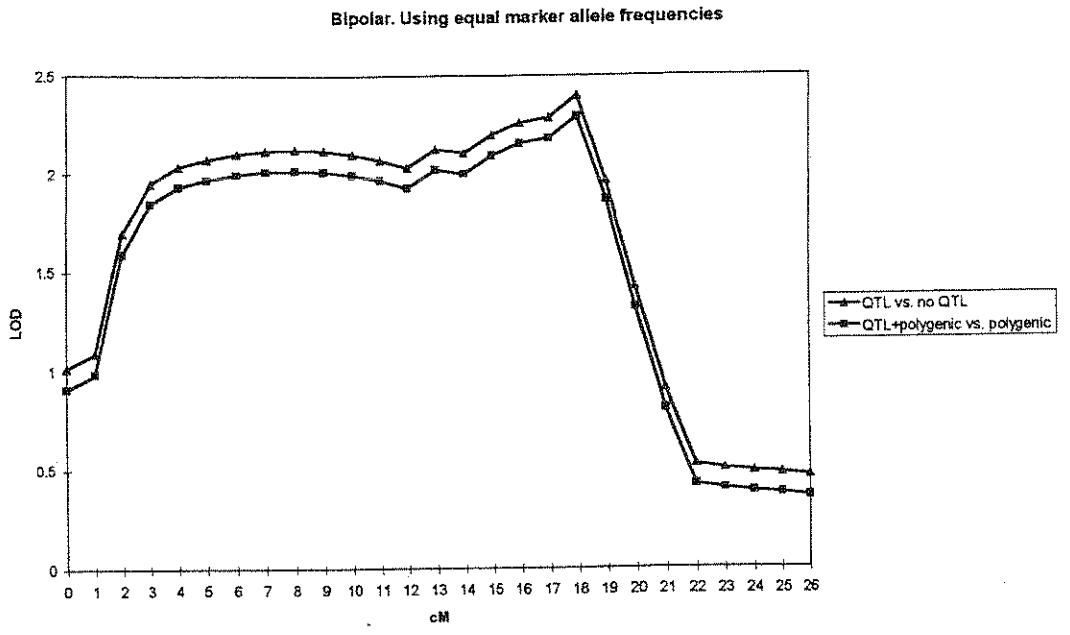

FIGURE 5. LOD scores for testing QTL variation for bipolar disorder, when the IBD coefficients are estimated, assuming that all alleles for a marker have the same frequency on chromosome 4p.

can be performed, i.e. testing whether a chromosome region explains variation in addition to the variation which is associated with the average relationships between individuals in the pedigree. Hence, the variance component method can easily distinguish between single gene effects and polygenic variation, whereas this is not possible with a standard linkage approach. Variance component methods have been successfully used in detecting QTLs for obesity-related traits in man (Duggirala *et al.*, 1996; Comuzzie *et al.*, 1997), and milk production traits in dairy cattle (Zhang *et al.*, 1998). Variance component

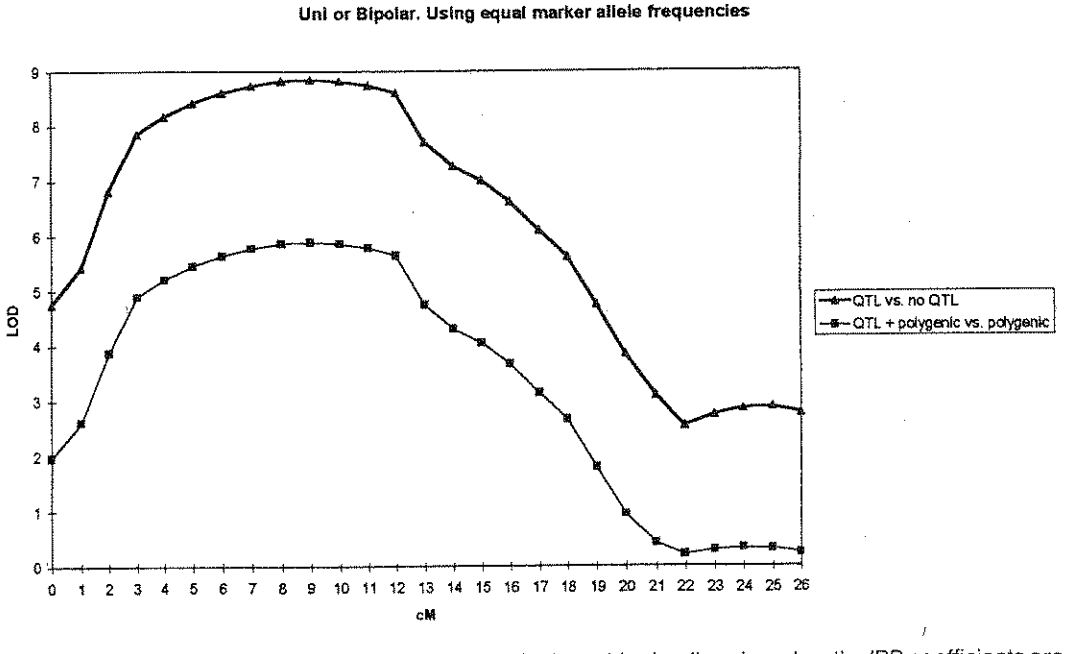

FIGURE 6. LOD scores for testing QTL variation for unipolar or bipolar disorder, when the IBD coefficients are estimated, assuming that all alleles for a marker have the same frequency on chromosome 4p.

models can easily include covariates and fixed effects, such as age at onset and a cohort effect. Extension of the methodology to multiple QTLs is relatively straightforward (Almasy and Blangero, 1998).

Other random effects can also easily be incorporated into the linear model when using the variance component method. For example, we fitted the random effect of a common environment of fullsibs in addition to fitting the polygenic effect. However, the variance component for the common fullsib effect was zero for all three traits. We also fitted a polygenic effect after fitting the common environmental effect, and that resulted in significant LOD scores for the polygenic heritability (h^2) for unipolar disorder (LOD = 3.9) and for uni- or bipolar disorder (LOD = 9.7). These results indicate that there is no evidence to suggest that the variation between nuclear families within this large pedigree is due to common environmental or within-family effects.

We have not produced any *P*-values to denote significance of the statistical tests, because the data (phenotypic and genotypic) used in our analyses are very similar to that previously published (Blackwood *et al.*, 1996), so that we had prior information that the region of chromosome 4 contained QTL activity. We performed many multiple tests in our analyses: three traits were tested for three models at 26 locations. However the three traits are correlated, one of the tests (that for polygenic variation) is independent of location, and tests on different chromosome locations are highly correlated, so that the total number of 'independent' tests from our analyses is likely to be less than 10 (say, two uncorrelated traits × two statistical tests × two independent locations). In practice, multiple tests need to be taken into account when declaring significant QTLs, because otherwise too many false positives would be reported (e.g. Lander and Kruglyak, 1995). However, even for complete genome scans the genome-wide threshold is unlikely to be larger than a LOD of 4.0, so that our reported results would be significant had they been selected from a genome scan. When performing a single test for the significance of a variance component, a LOD of 3.0 is approximately significant at the 0.0001 level, and a LOD of 5.9 approximately at the 10^{-7} level. If bipolar disorder is a more severe form of unipolar disorder, the power to detect QTLs should be larger if individuals are coded 'affected' if they suffer from either uni- or bipolar disorder. In our analyses, there was much stronger evidence for QTL activity for this combined trait than for unipolar or bipolar disorder separately. Blackwood *et al.* (1996) reported the highest LOD score (4.1) for linkage with bipolar disorder (using their narrow definition), although the broader definition used in their analysis (unipolar or bipolar disorder) produced a similar LOD score of 4.0 at the same marker. It is not clear why Blackwood *et al.* (1996) detected a chromosome location affecting bipolar disorder, when in the present analysis we find little evidence of genetic variation for the narrow definition of bipolar disorder. The difference may be that Blackwood *et al.* (1996) coded unipolar disorder (and other psychiatric diagnoses) as 'unknown' in their linkage analysis of bipolar disorder, whereas we coded them as 'zero' (i.e. absence of disorder) in our analysis. For example, if two siblings, one affected with bipolar disorder and one affected with unipolar disorder, share a common haplotype, the evidence for linkage would be stronger if unipolar was coded as 'unknown' than if it was coded as 'absence of bipolar'. In effect, coding unipolar as absence of bipolar disorder would be interpreted as evidence against linkage. Other possibilities for the differences in the strength of evidence for a region affecting bipolar disorder between the two analyses are the assumption regarding phenocopy rate and marker allele frequencies. To test the hypothesis that bipolar disorder is a more severe form of unipolar disorder, we also looked at an analysis in which no disorder was coded as '0', unipolar was coded as '1', and bipolar as '2'. The results from this study were very similar to the analysis of the trait 'presence of uni- or bipolar disorder': a LOD score of 5.2 for test 2 (QTL vs no QTL) and 3.5 for test 3 (QTL plus polygenic vs polygenic).

In the method, we use recombinations between the markers, and a putative linked QTL or major gene is accounted for in estimating the IBD matrix. Nonetheless, any such recombination would result in part of the variation caused by the major gene or QTL being absorbed by the polygenic component. However, the data is from a single extended family and we concentrated on a small region of the genome with a relatively high marker density. Hence, we would not expect much recombination to occur in this pedigree between a marker on chromosome 4p and a QTL for bipolar disorder in that region. By looking for a single QTL (or major gene), the variation caused by another (unlinked) gene with large effect would be partitioned into any of the three variance components (polygenic, QTL, residual), depending on the IBD pattern of that gene versus the IBD pattern of the chromosome 4 region versus the genome-wide average IBD pattern. On average, we would expect the variation of an unlinked QTL to be absorbed by the polygenic variance.

The region with the high LOD scores is relatively large (> 10 cM), and on the basis of these analysis more information is needed before fine mapping or candidate gene approaches are undertaken. More meioses are needed to fine-map the QTL on chromosome 4. Additional markers would allow a better estimation of IBD coefficients (by following haplotypes through the pedigree), but are unlikely to reduce the confidence interval around the QTL.

When analyses were performed assuming that alleles were at equal frequencies, the pairwise IBD coefficients between all affected individuals were estimated as 0.5 (results not shown), i.e. they share a common haplotype (as in Blackwood *et al.*, 1996). For bipolar disorder, the LOD scores increase by nearly one unit when fitting a QTL component (Figure 5), to a maximum value of about 2.4. The reason that the LOD score does not increase further is due to incomplete penetrance or recombination, because the IBD coefficients are calculated between both affected and unaffected individuals for haplotype sharing. For the trait uni- or bipolar disorder, the assumption of equal marker allele frequencies means that the LOD scores are increased even further (Figure 6). The LOD increases to a maximum value of nearly 9.0 for the test of genetic variation in the linkage group or elsewhere on the genome (test 2), and to a value of nearly 6.0 for the evidence of the presence of a QTL in the chromosome 4 region (Table 3).

The power of the analysis could be increased substantially if more families and more marker data were available. More marker information should provide more precise estimates of IBD coefficients between members of the pedigree, and more families should increase the power to distinguish between polygenic variance and QTL variance. If there is a strong polygenic component, it could easily mask the detection of a QTL in linkage studies, and the variance-based methods may therefore yield additional, previously undetected, QTL. Between-family heterogeneity in genetic aetiology of a trait is less of a concern in variance component analyses than it is in linkage analyses. For a particular genomic region, the addition of families which do not have causative polymorphism in the region to one or more that does may dilute the evidence for a gene, but does not add evidence against the presence of a gene in that region, as it can in linkage analyses. In addition, two QTL models can easily be accommodated in the analysis, by including a third random effect in the model.

The calculation of maximum likelihood values for several models allow the testing of several hypotheses. Although we have shown results from only three statistical tests, we could perform others. For example the difference between the sum of the LOD scores for tests 1 and 3 and the LOD for test 2 is a test for genetic variation in addition to the variation explained by the QTL. However, for the present pedigree the values of this test are close to zero (results not shown) and not significant at the nominal 5% level, suggesting little remaining genetic variation after fitting the QTL. This is perhaps not surprising, since the estimate of a polygenic heritability based upon a single extended family of size 168 is very imprecise, so that a likelihood ratio test is unlikely to reveal a significant (small) variance component. Information from multiple families should improve the detection of a polygenic component of variance.

In conclusion, we have demonstrated that a robust variance component method, which makes few assumptions about the genetic model of uni- and bipolar disorder, is appropriate to analyse data on complex psychiatric traits in extended pedigrees. We now plan to use this method to investigate QTL activity throughout the genome, using information from multiple families.

Acknowledgements

We thank the Medical Research Council and the Biotechnology and Biological Sciences Research Council for support. Thanks to Eric Sobel for help with the SIMWALK software.

REFERENCES

Almasy L, Blangero J (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* **62**:1198–1211.

Amos CI, Zhu DK, Boerwinkle E (1990). Assessing genetic linkage and association with robust components of variance approaches. *Ann Hum Genet* **60**:143–160.

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

Comuzzie AG, Hixson JE, Almasy L, Mitchell BD, Mahaney MC, Dyer TD *et al.* (1997). A major quantitative trait locus determining serum leptin levels and fat mass is located on human chromosome 2. *Nat Genet* **15**:273–275.

Duggirala R, Stern MP, Mitchell BD, Reinhart LJ, Shipman PA, Uresandi OC *et al.* (1996) Quantitative variation in obesity-related traits and insulin precursors linked to the *OB* gene region on human chromosome 7. *Am J Hum Genet* **59**:694–703.

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.

Grignola FE, Hoeschele I, Tier B (1996). Mapping quantitative trait loci in outcross populations via residual maximum likelihood. I. Methodology. *Genet Sel Evol* **28**:479–490.

Heath SC (1998). Generating consistent genotypic configurations for multi-allelic loci and large complex pedigrees. *Hum Hered* **48**:1–11.

Heath CS, Thompson EA (1997). MCMC samplers for multilocus analyses on complex pedigrees. *Am J Hum Genet* **61**:1623.

Lynch M., Walsh B (1998). *Genetics and analysis of quantitative traits*. Sinauer Sunderland, USA: Sinauer.

Patterson HD, Thompson R (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**:545–555.

Risch N, Botstein D (1996). A manic depressive history. *Nat Genet* **12**:351–353.

Sobel E, Lange K (1996). Descent graphs in pedigree analysis: application to haplotyping, location scores, and marker sharing statistics. *Am J Hum Genet* **58**:1323–1337.

Visscher PM, Haley CS, Knott SA (1996). Mapping QTLs for binary traits in backcross and F_2 populations. *Genet Res* **68**:55–63.

Xu S, Atchley WR (1995). A random model approach to interval mapping of quantitative trait loci. *Genetics* **141**:1189–1197.

Zhang Q, Boichard D, Hoeschele I, Ernst C, Eggen A, Murkve B *et al.* (1998). Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. *Genetics* **149**:1959–1973.