Hum Genet (2005) 117: 160–167
DOI 10.1007/s00439-005-1282-3

ORIGINAL INVESTIGATION

Elizabeth Holliday · Bryan Mowry · David Chant
Dale Nyholt

The importance of modelling heterogeneity in complex disease: application to NIMH Schizophrenia Genetics Initiative data

Received: 22 November 2004 / Accepted: 18 January 2005 / Published online: 21 April 2005

Abstract As for other complex diseases, linkage analyses of schizophrenia (SZ) have produced evidence for numerous chromosomal regions, with inconsistent results reported across studies. The presence of locus heterogeneity appears likely and may reduce the power of linkage analyses if homogeneity is assumed. In addition, when multiple heterogeneous datasets are pooled, inter-sample variation in the proportion of linked families (α) may diminish the power of the pooled sample to detect susceptibility loci, in spite of the larger sample size obtained. We compare the significance of linkage findings obtained using allele-sharing LOD scores (LOD_{exp})—which assume homogeneity—and heterogeneity LOD scores (HLOD) in European American and African American NIMH SZ families. We also pool these two samples and evaluate the relative power of the LOD_{exp} and two different heterogeneity statistics. One of these (HLOD-P) estimates the heterogeneity parameter α only in aggregate data, while the second (HLOD-S) determines α separately for each sample. In separate and combined data, we show consistently improved performance of HLOD scores over LOD_{exp}. Notably, genome-wide significant evidence for linkage is obtained at chromosome 10p in the European American sample using a recessive HLOD score. When the two samples are combined, linkage at the 10p locus also achieves genome-wide significance under HLOD-S, but not HLOD-P. Using HLOD-S, improved evidence for linkage was also obtained for a previously reported region on chromosome 15q. In linkage analyses of complex disease, power may be maximised by routinely modelling locus heterogeneity within individual datasets, even when multiple datasets are combined to form larger samples.

E. Holliday · B. Mowry · D. Chant
Queensland Centre for Mental Health Research,
Level 3, Dawson House, The Park, Centre for Mental Health,
Wacol, QLD, 4076, Australia

E. Holliday (✉) · B. Mowry · D. Chant
Department of Psychiatry,
University of Queensland,
Brisbane, Queensland, Australia
E-mail: lizh@qcmhr.uq.edu.au
Tel.: +61-7-32718697
Fax: +61-7-32718682

E. Holliday · D. Nyholt
Queensland Institute of Medical Research,
P.O. Royal Brisbane Hospital, Brisbane, Queensland, Australia

Introduction

The power of linkage analyses for complex disease may be severely confounded by the presence of locus heterogeneity. Locus heterogeneity exists when mutations at different genomic loci produce indistinguishable phenotypes. The presence of heterogeneity can severely reduce the power of standard statistics to detect linkage, particularly if penetrance at the disease locus is low (Abreu et al. 1999). A number of methods have been devised to reduce locus heterogeneity in complex disease, including the collection of population isolates and stratification of samples on the basis of refined phenotype definitions. However, population isolates are rare, and data stratification can result in power losses if more homogeneous analytic groups are not obtained (Leal and Ott 2000).

If the presence of locus heterogeneity cannot be avoided, its impact may be reduced by implementing heterogeneity LOD (HLOD) score analysis (Ott 1999). This method represents an extension of standard (homogeneity) likelihood-based methods, but maximises the likelihood with respect to both θ, the recombination frequency and α, the proportion of families linked at the disease locus. In this way, HLOD scores may better reflect the linkage evidence provided by families segregating mutations at the disease locus, and reduce the deflation of linkage signals by unlinked pedigrees. In the presence of locus heterogeneity, simulations suggest superior power of the HLOD over homogeneity LODs and "model-free" methods, both to detect linkage (Greenberg and Abreu 2001; Huang and Vieland 2001) and provide a more precise estimate of trait locus position (Finch et al. 2001; Vieland and Logue 2002).

An additional factor affecting statistical power is clearly sample size. A common approach to overcoming this issue is the combination and joint analysis of multiple datasets, via, for example, the formation of multicentre collaborations. However, if only a proportion of affected individuals harbour mutations at a given locus ($\alpha < 1$), sampling differences may produce variation in α across different datasets. Under these circumstances, the simple HLOD calculated for the entire "pooled" sample (HLOD-P) can suffer dramatic losses in power, compared with the individual dataset/s in which α is highest (Vieland et al. 2001). To maintain power under these circumstances, simulations suggest that α parameters should be estimated separately for constituent datasets (Huang and Vieland 2001; Vieland et al. 2001). The correct statistic for doing so has been denoted the "compound" HLOD (HLOD-C), a log-likelihood ratio maximised over θ and one α_i for each sample. Calculation of the HLOD-C statistic is computationally intensive, but for affected sib-pair (ASP) data HLOD-C is well approximated by an alternative score, the "summed" HLOD (HLOD-S), formed by summing individual HLOD scores over separate datasets. By computing α parameters separately for each sample, HLOD-S implicitly models heterogeneity differences between the samples. Compared with the HLOD-P, which estimates a single value of α for the entire pooled sample, HLOD-S shows consistent improvements in power when the value of α varies markedly between samples (Vieland et al. 2001).

Given these considerations, we have assessed the evidence for locus heterogeneity within and between two extant schizophrenia (SZ) datasets. SZ is a complex psychiatric disorder with a lifetime prevalence of ~1% (Jablensky et al. 1992). A variety of family, twin and adoption studies suggest that genetic factors have a substantial impact on SZ risk, and heritability estimates are in the range of 70–80% (Owen et al. 2002). In spite of such a strong genetic component, >20 genome-wide linkage scans for SZ have produced evidence for numerous genomic regions, with promising evidence accumulating for some, but no single region reported by a majority of studies. Locus heterogeneity in SZ thus appears likely. In spite of this evidence, many SZ linkage studies have adhered to statistics which assume homogeneity, although these are likely to suffer reduced power when only a proportion of families segregate mutations at peak loci.

We performed heterogeneity analyses in European American and African American affected sib pair-family samples ascertained by the National Institute of Mental Health (NIMH) Schizophrenia Genetics Initiative (Cloninger et al. 1998). Originally, the two datasets were analysed separately, due to the presence of significantly different marker allele frequencies. Using the criteria of Lander and Kruglyak (1995), nonparametric analyses found no region with statistically significant evidence for linkage, although genome-wide "suggestive" linkage was found on chromosome 10p (NPL $Z = 3.4$, $P = 0.0004$) in European Americans (Faraone et al. 1998). No region approached the threshold for significant or even suggestive evidence for linkage in the African American pedigrees (Kaufmann et al. 1998). A subsequent analysis of the pooled sample produced a heterogeneity LOD score of 3.97 on chromosome 15q14 under a co-dominant model (Freedman et al. 2001). We have analysed these two datasets separately, allowing for locus heterogeneity within each and also analysed the combined data, estimating heterogeneity parameters either across the entire sample or separately within each dataset.

Materials and methods

Ascertainment and diagnosis

The collection of this sample has been previously described (Cloninger et al. 1998). Briefly, 73 affected sib-pair families were ascertained at Harvard University, Columbia University and Washington University, with all patients giving informed consent for participation in genetic studies. Patients were interviewed using the Diagnostic Instrument for Genetic Studies (Nurnberger et al. 1994), and diagnoses were made using DSM-IIIR criteria (American Psychiatric Association 1987). Subjects satisfying a "core" diagnosis of SZ or schizoaffective disorder, depressed type, were classed as affected. Individuals with non-core diagnoses were considered unaffected. The affection status of remaining subjects for whom no interview data were obtained, was classed as unknown. Subjects' ethnicity was based on personal report of parental descent, yielding 43 European-American and 30 African-American pedigrees. All pedigrees contained at least two affected siblings.

Genotyping and marker characteristics

Subjects were genotyped at 459 microsatellite markers spanning autosomes and the X-chromosome at an average intermarker distance of 10 cM (Cloninger et al. 1998). Non-Mendelian inheritance errors were identified and marker allele frequencies calculated using the PEDMANAGER software (M.P. Reeve and M.J. Daly, personal communication), with allele frequencies based on founder genotypes. Allele frequencies were estimated separately for the two ethnicities, to allow for previously reported differences. Where possible, marker position was obtained from the deCODE genetic map (Kong et al. 2002). The position of markers not present on the decode map was estimated by interpolation from known Marshfield map positions.

Genetic analyses

For each dataset, multipoint linkage scores were computed across the entire genome using the ALLEGRO

(ver. 1.2c) program (Gudbjartsson et al. 2000). Parametric heterogeneity LOD (HLOD) scores were calculated under a simple (i.e. no phenocopies) dominant (HLOD-D) and a simple recessive (HLOD-R) mode of inheritance (MOI) (Greenberg et al. 1998), assuming 50% penetrance (Hodge et al. 1997). As recommended by Pal et al. (2001), we specified disease gene frequencies of 0.01 and 0.1 for the dominant and recessive models, respectively. To facilitate comparison with the previously reported NPL analysis, we computed exponential allele-sharing LOD scores (LOD_{exp}) for the score-pairs statistic. This likelihood ratio statistic provides similar results to the NPL score, but is less conservative than the NPL when genotype data are missing and markers are not perfectly informative (Kong and Cox 1997).

For multipoint analysis of the pooled sample, unique marker allele frequencies were preserved for each ethnic group by giving markers at all loci distinct designations for each group and treating the two sets as unique markers spaced 0.1 cM apart. Pedigree genotypes were coded null at map positions corresponding to the alternate ethnicity. The HLOD analyses of the pooled data estimated the single combination of θ and α which maximised the LOD score at each locus. This bivariate statistic is referred to as the ''pooled'' HLOD (HLOD-P) (Vieland et al. 2001) and was calculated under dominant and recessive models (HLOD-DP and HLOD-RP) as specified for the constituent datasets. LOD_{exp} scores were obtained in pooled data as for individual samples. The multivariate ''summed'' HLOD (HLOD-S) was calculated by summing HLODs obtained for the individual datasets under dominant and recessive models (HLOD-DS and HLOD-RS) (Vieland et al. 2001).

Empirical *P*-value determination

Genome-wide significance levels for all scores were determined empirically, and defined as the frequency with which equivalent or higher scores occurred in 1,000 replicates of data simulated under the null hypothesis. Linkage peaks separated by more than 30 cM were considered independent. Significance levels for African and European American HLOD statistics were corrected for the dominant and recessive models tested, by multiplying *P* values by the effective number of independent tests performed. This value relates to the variance of eigenvalues derived from a correlation matrix of the two sets of genome-wide linkage scores (Cheverud 2001). We did not correct for calculating LOD_{exp} scores. This approach was adopted to facilitate comparison between HLOD and LOD_{exp} scores. For the pooled data, empirical genome-wide *P* values corrected for testing both dominant and recessive models were analogously obtained for the HLOD-P and HLOD-S statistics, to allow their comparison with each other and with LOD_{exp} scores. Thresholds for genome-wide suggestive and significant linkage were established from the simulated null distribution of each statistic and based on the criteria of Lander and Kruglyak (1995). Suggestive linkage thresholds were defined as scores occurring with probability 1 in every genome scan (i.e. 1,000 total independent peaks in 1,000 genome-wide simulations) and significant linkage thresholds as scores occurring with probability 0.05 in every genome scan (50 peaks in 1,000 simulations). To account for the two genetic models specified for HLOD statistics, the number of required peaks for each threshold was divided by the number of independent tests performed.

Results

Analysis of individual samples

Multipoint HLOD and LOD_{exp} scores for the individual datasets are presented in Fig. 1. The correlation between genome-wide HLOD-D and HLOD-R scores was 0.76 for the European American sample and 0.23 for the African American sample, yielding estimates of 1.42 (European American) and 1.95 (African American) effectively independent tests. Based on each statistic's simulated null distribution, and incorporating multiple testing for HLOD statistics, genome-wide suggestive linkage thresholds of 1.82 (HLOD-D), 1.80 (HLOD-R) and 1.64 (LOD_{exp}) and significant linkage thresholds of 3.14 (HLOD-D), 3.11 (HLOD-R) and 2.97 (LOD_{exp}) were determined for the European American sample. For African American data, suggestive linkage thresholds of 1.94 (HLOD-D), 1.97 (HLOD-R) and 1.63 (LOD_{exp}) and significant linkage thresholds of 3.17 (HLOD-D), 3.24 (HLOD-R) and 2.87 (LOD_{exp}) were analogously obtained. Linkage scores and corrected significance levels for peaks exceeding suggestive thresholds for any score are shown in Table 1. Distances are expressed in Kosambi cM from the p-terminus.

The strongest evidence for linkage was detected in European American data, on chromosome 10p12.32 (45.1 cM: D10S1423-D10S582) under the simple recessive heterogeneity model (HLOD-R = 3.92). Only three peaks exceeded this value in 1,000 genome-wide simulations of the HLOD-R statistic. Following correction for testing both HLOD-R and HLOD-D (1.42 independent tests), this result satisfied our criteria for genome-wide significant linkage (corrected genome-wide $P = 0.004$). A peak LOD_{exp} score of 2.85 was observed slightly telomeric to this peak, at 38 cM (D10S2325-D10S1423). This LOD_{exp} result met our criteria for suggestive, but not significant linkage (genome-wide $P = 0.061$) (Table 1). No additional region satisfied our criteria for suggestive linkage in the European American sample using any of the three statistics.

In the African American dataset, the peak linkage score was obtained at chromosome 15q13.3 (30.25 cM, D15S128-D15S118), under the simple dominant heterogeneity model (HLOD-D = 3.07). Following multiple testing correction (1.95 tests), this result easily met our criteria for suggestive linkage, and nearly achieved

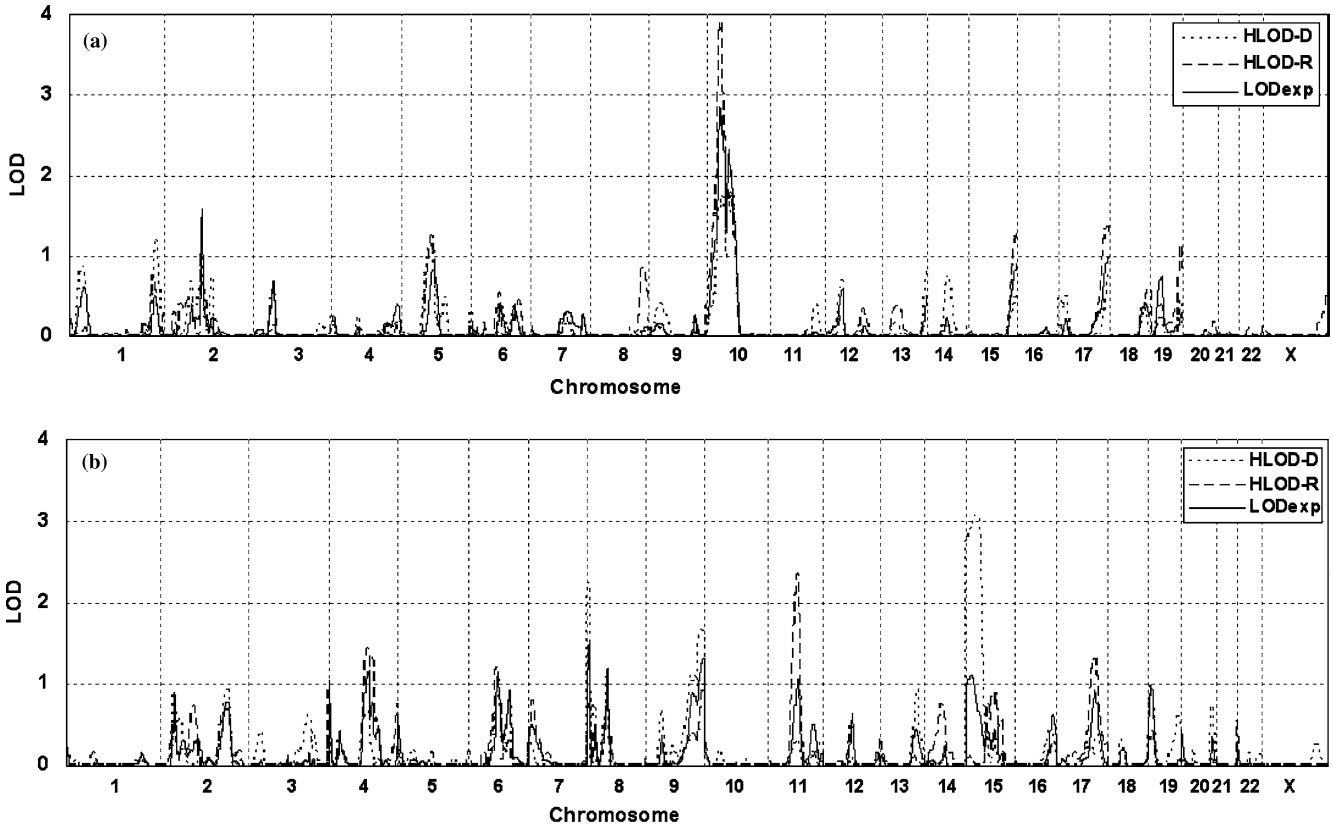


Fig. 1a,b Individual datasets. Multipoint allele-sharing LOD (*LOD exp*) and heterogeneity LOD (*HLOD*) scores computed under simple dominant (*HLOD-D*) and recessive (*HLOD-R*) models for **a** European American and **b** African American pedigrees. Individual chromosomal locations are indicated along the *x*-axis

genome-wide significance (corrected $P = 0.056$). Two additional regions also demonstrated suggestive linkage under heterogeneity. These were located at chromosome 8p23.2 (6.57 cM, HLOD-D = 2.25, corrected $P = 0.391$), between D8S264 and D8S439 and at 11q14.1 (86.78 cM, HLOD-R = 2.35, corrected $P = 0.347$), between D11S2371 and D11S2002 (Table 1). In contrast, no region demonstrated suggestive linkage using LOD_{exp} in the African American sample.

Analysis of combined samples

Multipoint HLOD-P and LOD_{exp} scores for the pooled European and African American data are shown in Fig. 2. The HLOD-S scores are presented in Fig. 3. The correlation between dominant and recessive genome-wide scores was 0.38 for HLOD-P and 0.42 for HLOD-S, yielding estimates of 1.86 (HLOD-P) and 1.82

Table 1 HLOD and LOD_{exp} scores for individual samples. Multipoint scores are shown for regions demonstrating genome-wide suggestive linkage in European American or African American pedigrees. Genome-wide empirical significance levels are shown *in parentheses* for scores exceeding suggestive thresholds. For HLOD scores, *P* values were multiplied by the effective number of independent tests (1.42: European American; 1.95: African American) to correct for testing both dominant and recessive models [*cM* location of peak score in cM (deCODE map), *HLOD* heterogeneity LOD score, LOD_{exp} allele-sharing LOD score]

Sample and location	cM	HLOD		LOD_{exp}
		Dominant	Recessive	
European American				
10p12.32	45.14	1.65	3.92 (0.004)[b]	2.70 (0.087)[a]
10p12.33	38	1.62	3.76 (0.004)[b]	2.85 (0.061)[a]
African American				
8p23.2	6.57	2.25 (0.391)[a]	0.6	1.41
11q14.1	86.78	0.31	2.35 (0.347)[a]	1.06
15q13.3	30.25	3.06 (0.056)[a]	0	0.78

[a] Meets threshold for genome-wide suggestive linkage
[b] Meets threshold for genome-wide significant linkage

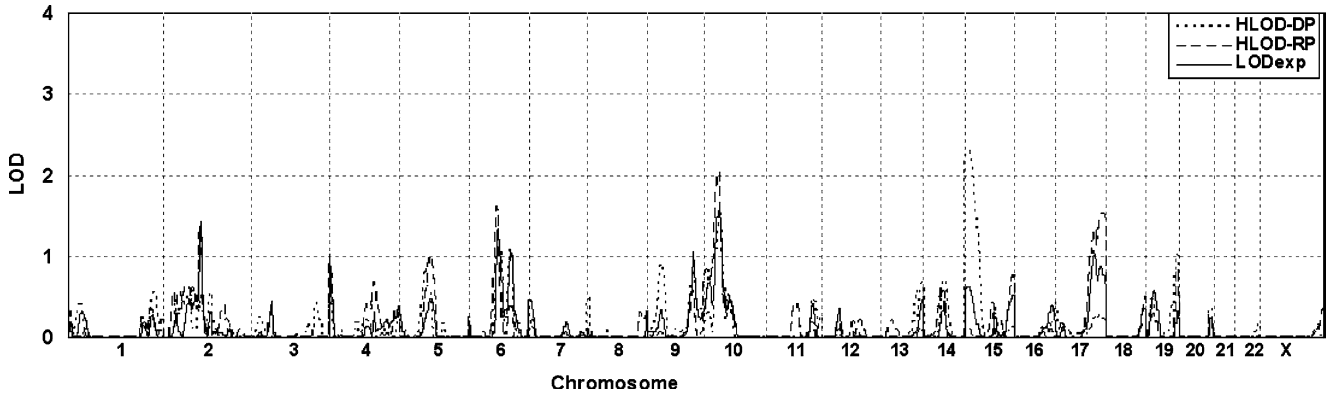
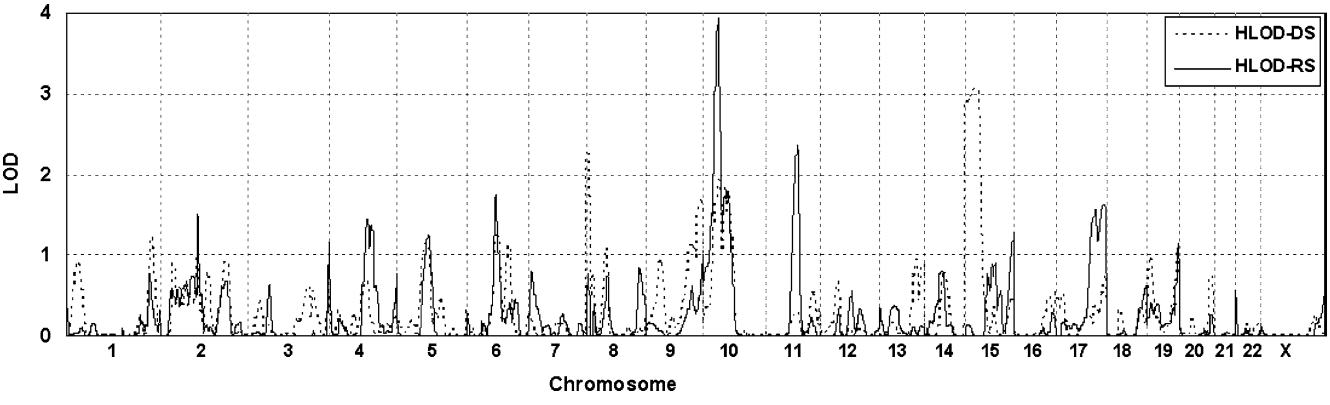


Fig. 2 Combined European and African American data: "pooled" scores. Multipoint allele-sharing LOD (LOD_{exp}) and HLOD-P scores calculated under simple dominant (*HLOD-DP*) and recessive (*HLOD-RP*) models for the pooled data

(HLOD-S) independent tests. Suggestive linkage thresholds of 1.86 (HLOD-DP), 1.90 (HLOD-RP), 1.63 (LOD_{exp}), 2.39 (HLOD-DS) and 2.42 (HLOD-RS) and significant thresholds of 3.22 (HLOD-DP and HLOD-RP), 2.94 (LOD_{exp}), 3.72 (HLOD-DS) and 3.70 (HLOD-RS) were empirically determined. Linkage scores and corrected significance levels for peaks exceeding suggestive thresholds are shown in Table 2.

In the combined sample, no region exceeded our threshold for genome-wide significant or suggestive linkage using LOD_{exp}. In contrast, each of our statistics which incorporated locus heterogeneity produced two genome-wide suggestive linkages. With HLOD-P, the highest score was obtained at chromosome 15q12 (12.9 cM, D15S128-D15S118) under the dominant model (HLOD-DP = 2.29), corresponding with a corrected genome-wide *P* value of 0.316. Suggestive linkage was also obtained at chromosome 10p12.32 (44.65 cM) between D10S1423-D10S582, using HLOD-P under a recessive model (HLOD-RP = 2.04; corrected $P = 0.553$) (Table 2).

Fig. 3 Combined European and African American data: "summed" scores. Multipoint HLOD scores summed across independent datasets for dominant (*HLOD-DS*) and recessive (*HLOD-RS*) models



Using HLOD-S, the recessive score at chromosome 10p12.32 achieved genome-wide significance (HLOD-RS = 3.92; corrected $P = 0.038$), with this result being nearly 15 times more significant than the HLOD-RP obtained in this region. Using HLOD-S under the dominant model, we also obtained suggestive evidence for linkage at chromosome 15q14 (D15S128-D15S118: HLOD-DS = 3.07; corrected $P = 0.195$). The significance of this result was nearly twice that of the HLOD-DP score in this region. Variability in the location of linkage peaks for the different statistics was negligible. On chromosome 10p, the position of the HLOD-RS peak (44.65 cM) corresponded well with both the peak HLOD-RP (45.14 cM), and HLOD-R for the European American data (45.14 cM). On chromosome 15, the position of the peak HLOD-DS (33.4 cM) corresponded well with the peak HLOD-D observed in African American data (30.25 cM), while the HLOD-DP result was more centromeric (12.9 cM).

Discussion

The presence of locus heterogeneity is likely to confound the linkage analysis of many complex diseases. In SZ, locus heterogeneity is suggested by the relatively high first-degree relative risk ($\lambda_s = 10$) (Levinson and Mowry 2000) and lifetime population prevalence of the disease (~1%), combined with conflicting results of many genome scans. The detection of genes influencing liability to SZ and other complex diseases may be enhanced by the

Table 2 HLOD-P and HLOD-S scores for combined data. Multipoint scores are shown for regions exceeding empirical suggestive thresholds in combined European and African American data. Genome-wide empirical significance levels are shown *in parentheses* for scores exceeding suggestive thresholds. For HLOD-P and HLOD-S scores, *P* values were multiplied by the effective number of independent tests (1.86: HLOD-P and 1.82: HLOD-S) to correct for testing both dominant and recessive models (*cM* location of peak score in cM (deCODE map), *HLOD-P* heterogeneity LOD score for pooled sample, *HLOD-S* sum of HLOD scores from separate datasets, LOD_{exp} allele-sharing LOD score)

Location	cM	HLOD-P		HLOD-S		LOD_{exp}
		Dominant	Recessive	Dominant	Recessive	
10p12.32	44.65	1.35	2.04 (0.553)[a]	1.73	3.88 (0.04)[b]	1.56
10p12.32	45.14	1.39	2.03 (0.568)[a]	1.78	3.92 (0.038)[b]	1.58
15q12	12.9	2.29 (0.316)[a]	0.11	2.89 (0.283)[a]	0.13	0.62
15q14	33.4	1.30	0	3.07 (0.195)[a]	0	0.16

[a] Meets threshold for genome-wide suggestive linkage
[b] Meets threshold for genome-wide significant linkage

use of linkage analytic methods which allow for the presence of locus heterogeneity.

Utilising these relevant and well characterised datasets, we provide evidence that heterogeneity LOD (HLOD) scores calculated under simple dominant and recessive models have better power than allele-sharing LOD scores (LOD_{exp}) to detect linkage in ASP data for complex disease. This effect was evident within both separate and combined samples. These individual SZ samples have previously been analysed using only the NPL statistic, with one suggestive linkage detected at chromosome 10p in the European American data. No region approached the threshold required for genome-wide significance in either sample (Faraone et al. 1998; Kaufmann et al. 1998). We obtained identical results using the LOD_{exp} score. The power of both the NPL and LOD_{exp} is likely to have been reduced by their inability to allow for heterogeneity, and the contribution of negative linkage scores by individual pedigrees at peak loci (not shown). Once heterogeneity was incorporated within a parametric model, we detected significant evidence for linkage at chromosome 10p in both European American and the combined data (HLOD-S). These results suggest the presence of locus heterogeneity in these samples, and support those of a body of work (Faraway 1993; Greenberg and Abreu 2001; Vieland and Logue 2002) demonstrating the superior power of HLOD analyses for detecting linkage in complex disease. As reported for homogeneity LOD scores (Greenberg et al. 1998), we show that calculating the HLOD under simple dominant and recessive models with reduced penetrance, and correcting for the two tests, has good power to detect linkage even when the true MOI at the disease locus is unknown. The increased information obtained by using the two models is suggested by the low correlation between dominant and recessive scores for each statistic. Accordingly, a previous analysis of this combined data used only a dominant model, and reported convincing evidence for linkage to 15q, but not the 10p locus detected here (Freedman et al. 2001). This result indicates sensitivity of the HLOD to the penetrance model assumed, as reported for homogeneity LOD scores (Durner et al. 1999). However, when multiple inheritance models are tested, determination of the appropriate multiple testing correction is complicated by their non-independence. Failure to account for test dependence yields conservative significance estimates. To approach this problem, we have utilised the Bonferroni method proposed by Cheverud (2001), which estimates the *effective* number of independent tests based on the genome-wide correlations among different linkage scores. Linear regression techniques have also been proposed to estimate test independence (Camp and Farnham 2001). These methods may be readily applied to analyses involving any number of genetic models.

In addition to confirming the importance of incorporating heterogeneity, our analyses of these combined data suggest that the heterogeneity parameter α should be estimated separately for each sample. Notably, genome-wide significant linkage was detected at 10p using HLOD-S, but not HLOD-P, and linkage evidence at 15q was also stronger using HLOD-S. These results suggest that the European and African American samples contained different proportions of families (α) segregating mutations at these loci (Vieland et al. 2001). Such an effect is also supported by the different size of linkage signals for the two samples in these regions. Even using HLOD-S, however, the evidence for linkage was reduced at all peak loci (10p, 15q, 8p and 11q) compared with the individual dataset demonstrating the initial linkage signal. This highlights the power losses, which can result from combining heterogeneous samples, even when heterogeneity is modelled during analysis.

We believe these results to have important implications for linkage studies of complex disease, especially where larger samples are obtained by pooling multiple, independent datasets (Cox et al. 2001; Mowry et al. 2004). For such data, sampling differences are likely to produce variable levels of heterogeneity across different samples. When common marker maps and diagnostic schemes are used and linkage evidence can readily be combined across studies, the use of statistics which incorporate variation in α may better maintain the power of aggregate data. Failure to do so may reduce the power of the combined sample, in spite of its larger size. Note that we considered only two individual

samples here, and the simulations of Vieland et al. (2001) were based upon three separate datasets. The relative power of the HLOD-S statistic as the number of datasets increases will require further investigation, particularly as this statistic accumulates degrees of freedom with each additional sample. Alternative statistics to the HLOD-S also exist: for an explanation of these, see the report by (Vieland et al. 2001).

In conclusion, we report significant evidence of linkage to chromosome 10p and highly suggestive evidence of linkage to chromosome 15q in NIMH SZ data. Convincing detection of these regions required the acknowledgement of heterogeneity within individual samples, even following the combination of multiple datasets. Sampling differences against a background of considerable locus heterogeneity may contribute to the variability of linkage findings for SZ and other complex diseases. The incorporation of unique heterogeneity parameters for individual samples may be necessary for extracting the quality and consistency of linkage information necessary for motivating candidate region analyses.

Acknowledgements We are grateful to all members of the NIMH Schizophrenia Genetics Initiative for establishing this beneficial resource.

References

Abreu PC, Greenberg DA, Hodge SE (1999) Direct power comparisons between simple LOD scores and NPL scores for linkage analysis in complex diseases. Am J Hum Genet 65:847–857

American Psychiatric Association (1987) Diagnostic and statistical manual of mental disorders, 3rd edn, revised edn. American Psychiatric Association Press, Washington DC

Camp NJ, Farnham JM (2001) Correcting for multiple analyses in genomewide linkage studies. Ann Hum Genet 65:577–582

Cheverud JM (2001) A simple correction for multiple comparisons in interval mapping genome scans. Heredity 87:52–58

Cloninger CR, Kaufmann CA, Faraone SV, Malaspina D, Svrakic DM, Harkavy-Friedman J, Suarez BK, Matise TC, Shore D, Lee H, Hampe CL, Wynne D, Drain C, Markel PD, Zambuto CT, Schmitt K, Tsuang MT (1998) Genome-wide search for schizophrenia susceptibility loci: the NIMH genetics initiative and millennium consortium. Am J Med Genet 81:275–281

Cox NJ, Wapelhorst B, Morrison VA, Johnson L, Pinchuk L, Spielman RS, Todd JA, Concannon P (2001) Seven regions of the genome show evidence of linkage to type 1 diabetes in a consensus analysis of 767 multiplex families. Am J Hum Genet 69:820–830

Durner M, Vieland VJ, Greenberg DA (1999) Further evidence for the increased power of LOD scores compared with nonparametric methods. Am J Hum Genet 64:281–289

Faraone SV, Matise T, Svrakic D, Pepple J, Malaspina D, Suarez B, Hampe C, Zambuto CT, Schmitt K, Meyer J, Markel P, Lee H, Harkavy Friedman J, Kaufmann C, Cloninger CR, Tsuang MT (1998) Genome scan of European-American schizophrenia pedigrees: results of the NIMH genetics initiative and millennium consortium. Am J Med Genet 81:290–295

Faraway JJ (1993) Distribution of the admixture test for the detection of linkage under heterogeneity. Genet Epidemiol 10:75–83

Finch SJ, Chen CH, Gordon D, Mendell NR (2001) A study comparing precision of the maximum multipoint heterogeneity LOD statistic to three model-free multipoint linkage methods. Genet Epidemiol 21:315–325

Freedman R, Leonard S, Olincy A, Kaufmann CA, Malaspina D, Cloninger CR, Svrakic D, Faraone SV, Tsuang MT (2001) Evidence for the multigenic inheritance of schizophrenia. Am J Med Genet 105:794–800

Greenberg DA, Abreu PC (2001) Determining trait locus position from multipoint analysis: accuracy and power of three different statistics. Genet Epidemiol 21:299–314

Greenberg DA, Abreu P, Hodge SE (1998) The power to detect linkage in complex disease by means of simple LOD-score analyses. Am J Hum Genet 63:870–879

Gudbjartsson DF, Jonasson K, Frigge ML, Kong A (2000) Allegro, a new computer program for multipoint linkage analysis. Nat Genet 25:12–13

Hodge SE, Abreu PC, Greenberg DA (1997) Magnitude of type I error when single-locus linkage analysis is maximized over models: a simulation study. Am J Hum Genet 60:217–227

Huang J, Vieland VJ (2001) Comparison of "model-free" and "model-based" linkage statistics in the presence of locus heterogeneity: single data set and multiple data set applications. Hum Hered 51:217–225

Jablensky A, Sartorius N, Ernberg G, Anker M, Korten A, Cooper JE, Day R, Bertelsen A (1992) Schizophrenia: manifestations, incidence and course in different cultures. A World Health Organization ten-country study. Psychol Med Monogr Suppl 20:1–97

Kaufmann CA, Suarez B, Malaspina D, Pepple J, Svrakic D, Markel PD, Meyer J, Zambuto CT, Schmitt K, Matise TC, Harkavy Friedman JM, Hampe C, Lee H, Shore D, Wynne D, Faraone SV, Tsuang MT, Cloninger CR (1998) NIMH genetics initiative millenium schizophrenia consortium: linkage analysis of African-American pedigrees. Am J Med Genet 81:282–289

Kong A, Cox NJ (1997) Allele-sharing models: LOD scores and accurate linkage tests. Am J Hum Genet 61:1179–1188

Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K (2002) A high-resolution recombination map of the human genome. Nat Genet 31:241–247

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

Leal SM, Ott J (2000) Effects of stratification in the analysis of affected-sib-pair data: benefits and costs. Am J Hum Genet 66:567–575

Levinson DF, Mowry BJ (2000) Genetics of schizophrenia. In: Pfaff D, Berrettini W, Maxson S, Joh T (eds) Genetic influences on neural and behavioural functions. CRC Press, New York, pp 47–82

Mowry BJ, Holmans PA, Pulver AE, Gejman PV, Riley B, Williams NM, Laurent C, Schwab SG, Wildenauer DB, Bauche S, Owen MJ, Wormley B, Sanders AR, Nestadt G, Liang KY, Duan J, Ribble R, Norton N, Soubigou S, Maier W, Ewen-White KR, DeMarchi N, Carpenter B, Walsh D, Williams H, Jay M, Albus M, Nertney DA, Papadimitriou G, O'Neill A, O'Donovan MC, Deleuze JF, Lerer FB, Dikeos D, Kendler KS, Mallet J, Silverman JM, Crowe RR, Levinson DF (2004) Multicenter linkage study of schizophrenia loci on chromosome 22q. Mol Psychiatry 9:784–95

Nurnberger JI Jr, Blehar MC, Kaufmann CA, York Cooler C, Simpson SG, Harkavy Friedman J, Severe JB, Malaspina D, Reich T (1994) Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. Special issue: Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. Arch Gen Psychiatry 51:849–859

Ott J (1999) Analysis of human genetic linkage, 3rd edn. The John Hopkins University Press, Baltimore

Owen MJ, O'Donovan MC, Gottesman II (2002) Schizophrenia. In: McGuffin P, Owen MJ, Gottesman II (eds) Psychiatric genetics and genomics. Oxford University Press, Oxford, pp247–266

Pal DK, Durner M, Greenberg DA (2001) Effect of misspecification of gene frequency on the two-point LOD score. Eur J Hum Genet 9:855–859

Vieland VJ, Logue M (2002) HLODs, trait models, and ascertainment: implications of admixture for parameter estimation and linkage detection. Hum Hered 53:23–35

Vieland VJ, Wang K, Huang J (2001) Power to detect linkage based on multiple sets of data in the presence of locus heterogeneity: comparative evaluation of model-based linkage methods for affected sib pair data. Hum Hered 51:199–208