Follow-up Study on a Susceptibility Locus for Schizophrenia on Chromosome 6q

Maria Martinez,[1]* Lynn R. Goldin,[2] Quihe Cao,[2] Jing Zhang,[2] Alan R. Sanders,[2] Derek J. Nancarrow,[3,4] Jennifer M. Taylor,[3] Douglas F. Levinson,[5] Andrew Kirby,[5] Raymond R. Crowe,[6] Nancy C. Andreasen,[6] Donald W. Black,[6] Jeremy M. Silverman,[7] David P. Lennon,[4] Deborah A. Nertney,[4] Donna M. Brown,[8] Bryan J. Mowry,[4] Elliot S. Gershon,[2] and Pablo V. Gejman[2]

[1]Génétique des Maladies Humaines, Institut National de la Recherche et de la Santé Médicale, Paris, France
[2]Clinical Neurogenetics Branch, National Institute of Mental Health, Bethesda, Maryland
[3]Queensland Institute of Medical Research, Brisbane, Australia
[4]Department of Psychiatry, University of Queensland, and Queensland Centre for Schizophrenia Research, Wolston Park Hospital, Brisbane, Australia
[5]Department of Psychiatry, Allegheny University of the Health Sciences, Philadelphia, Pennsylvania
[6]Mental Health Clinical Research Center and Department of Psychiatry, The University of Iowa College of Medicine, Iowa City, Iowa
[7]Department of Psychiatry, Mt. Sinai School of Medicine, New York, New York
[8]Research Genetics, Inc., Huntsville, Alabama

Evidence for suggestive linkage to schizophrenia with chromosome 6q markers was previously reported from a two-stage approach. Using nonparametric affected sib pairs (ASP) methods, nominal *p*-values of 0.00018 and 0.00095 were obtained in the screening (81 ASPs; 63 independent) and the replication (109 ASPs; 87 independent) data sets, respectively. Here, we report a follow-up study of this 50cM 6q region using 12 microsatellite markers to test for linkage to schizophrenia. We increased the replication sample size by adding an independent sample of 43 multiplex pedigrees (66 ASPs; 54 independent). Pairwise and multipoint nonparametric linkage analyses conducted in this third data set showed evidence consistent with excess sharing in this 6q region, though the statistical level is weaker (*p*=0.013). When combining both replication data sets (total of 141 independent ASPs), an overall nominal *p*-value=0.000014 (LOD=3.82) was obtained. The sibling recurrence risk (λs) attributed to this putative 6q susceptibility locus is estimated to be 1.92. The linkage region could not be narrowed down since LOD score values greater than three were observed within a 13cM region. The length of this region was only slightly reduced (12cM) when using the total sample of independent ASPs (204) obtained from all three data sets. This suggests that very large sample sizes may be needed to narrow down this region by ASP linkage methods. Study of the etiological candidate genes in this region is ongoing. Am. J. Med. Genet. (Neuropsychiatr. Genet.) 88:337–343, 1999.
© 1999 Wiley-Liss, Inc.

KEY WORDS: schizophrenia; linkage; genome scan

INTRODUCTION

We have previously reported evidence for suggestive linkage of chromosome 6q markers to schizophrenia in two independent family samples: the sample of 81 affected-sib-pairs (ASPs) in 53 multiplex families from the NIMH IntraMural sample (NIMH-IM) and in the NIMH Schizophrenia Genetics Initiative sample (NIMH-GI) of 109 ASPs in 69 families [Cao et al., 1997]. In that study we screened 41 markers covering all of chromosome 6 in the NIMH-IM data set. Using pairwise and multipoint ASP statistics, we found the strongest evidence for linkage within a 14-cM region of the 6q area (from D6S301 to D6S303, highest LOD = 3.06 at D6S454/423, p = 0.00018). In the NIMH-GI dataset, we genotyped 13 markers from this region, and evidence for suggestive linkage was also obtained (highest LOD = 2.35 at D6S424/434, p = 0.00095).

Although evidence for linkage to markers on chromosome 6q is not definitive, our results are of sufficient interest to merit further investigation. To increase the

replication sample size, we added an independent sample of 66 ASPs in 40 schizophrenia pedigrees (43 nuclear families) collected in the United States and Australia [Levinson et al., 1998]. This design is part of a previously agreed upon collaboration between our NIMH group and the US/Australia group to examine areas according to criteria of follow-up utilized in either center. For NIMH-IM we follow up linkage results with pairwise nominal *p*-values less than 0.01. We applied the same overall methodology for genotyping and linkage analysis as in our previous publication [Cao et al., 1997]. We now report the results of 12 markers spanning a region of 50cM in 6q; from D6S455 (position 83.7cM) to D6S472 (position 135.7cM). We genotyped more markers at the extremes and also increased map density in specific areas: at 110.9cM in both NIMH-IM and NIMH-GI data sets and at 91.2cM in NIMH-GI data set. We found additional evidence for linkage in this region in the US/Australia data set (nominal *p*-value = 0.013).

MATERIALS AND METHODS
Families: Ascertainment and SCZ Definition

Genome scanning data set: NIMH-IM. This data set has been previously described in Cao et al. [1997]. A total of 53 families with two participating siblings affected with either schizophrenia (SCZ) or schizoaffective disorder (SA) were studied. This study was approved by the NIMH Institutional Review Board, and all participants gave written informed consent. Diagnostic resources included a semistructured diagnostic interview [Schedule for Affective Disorders and Schizophrenia -Lifetime version (SADS-L)] [Endicott and Spitzer, 1978], systematic review of medical records, and family history interviews [Gershon et al., 1988]. The predominant ethnic composition per family was: Caucasian 84.1%, African-American 3.5%, others 12.4%. Of all the possible affected sib-pair combinations, the ASP types were: SCZ-SCZ 43.7%, SCZ-SA, bipolar type 31%; SCZ-SA, depressive type 10.3%; SA, depressive type-SA, depressive type 4.6%; SA, bipolar type-SA, depressive type 3.4%; SA, bipolar type-SA, bipolar type 6.9%.

First replication data set: NIMH-GI. Sixty-nine families with an affected sib pair with SCZ or SA were obtained from the NIMH-GI collection. Data collection included a structured interview with the diagnostic interview for genetic studies [Nurnberger et al, 1994], and family history data were collected by using the family instrument for genetic studies and diagnosis by *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III R) [American Psychiatric Association, 1987]. The ethnic composition of the subset of families used by us for chromosome 6 linkage mapping is 51% Caucasian, 35% African-American, and 14% of other ethnic origins. Of all the possible ASP combinations, the ASP types were SCZ-SCZ 73%; SCZ-SA, depressive type 25%; and SA, depressive type-SA, depressive type 2%.

Second replication data set: US/Australia Sample. The 43 pedigrees were ascertained in Australia (18), Philadelphia (14), Iowa (8), and New York (3). Ethnic backgrounds were European-Caucasian (72%), African-American (21%), and other (Caribbean-Hispanic, Aboriginal/Micronesian, and Asian, 9%). Modified SADS [Endicott and Spitzer, 1978] or comprehensive assessment of symptoms and history (CASH) [Andreasen et al., 1992] diagnostic interviews were completed with probands and available relatives. Pedigrees were included if there was a proband with chronic schizophrenia and either a sibling with schizophrenia or two relatives with any nonaffective psychosis (see below) by consensus best estimate diagnosis using DSM-III R criteria, with affected individuals connected by first- or second-degree relationships. Written informed consent was obtained from all subjects.

Diagnostic reliability and differences across samples. For the NIMH-IM and NIMH-GI samples, only subjects with consensus diagnoses of SCZ or SA were included as described above. For the US/Australia sample, previous linkage analyses had been based on a model that included other nonaffective psychoses, and the diagnosis of schizophrenia was assigned conservatively. To determine the best common model for this study, a diagnostic reliability exercise was conducted for the NIMH-IM and US/Australia samples. Each research group used its usual consensus procedure to diagnose 15 cases from the other site, for a total of 30 cases. Cohen's kappa was 0.84 for the category of nonaffective psychosis (including schizophrenia, schizoaffective, delusional, and nonaffective atypical psychosis) and 0.73 for a narrow category including schizophrenia and schizoaffective disorder. Most disagreements were cases given a narrow diagnosis by NIMH diagnosticians and a broader diagnosis by US/Australia diagnosticians. Based on these results it was decided to include all nonaffective psychosis cases in the US/Australia sample because the narrower model appeared to have been applied in a more restrictive fashion in this sample. It was not possible to conduct a similar study with NIMH-GI cases because only the consensus diagnosis is made publicly available. However, reported inter-rater reliability results across multiple academic sites were quite high for schizophrenia [Nurnberger et al., 1994], and while lower for schizoaffective disorder, reliability was quite high for the category of "schizophrenia or schizoaffective disorder" [Faraone et al., 1996], as has been reported by others [Roy et al., 1997].

Genotyping

The genotyping of the NIMH-IM data set was performed by radioactive methods or fluorescence-based genotyping depending on the marker [Cao et al., 1997; Gejman et al., 1993]. Genotyping of the US/Australia sample was performed by semiautomated methods using fluorescinated markers at Research Genetics, Inc., as previously described [Levinson et al., 1998], or by manual methods using radioactive markers at Queensland Institute of Biomedical Research as previously described [Mowry et al., 1995], depending on the marker. For each marker, all subjects were typed at the same laboratory.

Statistical Analysis

Marker allele frequencies were set to their maximum likelihood values estimated in each data set with the

computer program ILINK from the Linkage package [Lathrop and Lalouel, 1984].

Pairwise affected sib-pair analysis was conducted with the SIBPAIR program [Terwilliger, 1996], which provides a likelihood-based test statistic for linkage that is equivalent to the LOD score calculated under the assumption of a simple recessive disease model with phase-unknown matings [Knapp et al., 1994]. To test for linkage, the likelihoods over all families are maximized as a function of the sharing rate (y) and the likelihood ratio test statistic is calculated against the null hypothesis of $y = 0.5$. The statistic follows a chi-square distribution with one degree of freedom (df) and can thus be expressed as a LOD score.

Multipoint nonparametric linkage analysis, which uses information from all markers simultaneously, was performed by using two statistical methods. First, we used the program sib_phase from the ASPEX package [version 1.12, Hinds and Risch, 1996] to estimate rate of identity by descent (IBD) sharing in ASPs. The maximized LOD score, at any chromosomal location, is obtained as a function of y, where y is the maximum likelihood estimate of the rate of IBD sharing. The independent sharing model (or linear model) is defined by: $z_2 = y^2$, $z_1 = 2y(1-y)$ and $z_0 = (1-y)^2$, where z_2, z_1, and z_0 are the probabilities for an affected sib-pair of IBD status 2, 1, and 0, respectively. The maximized LOD scores follow a χ^2 distribution with 1 df. An evaluation of the sibling recurrence risk (λs) attributed to a locus can be calculated as $0.25/z_0$ [Risch, 1990]. To check the sensitivity of our results to marker allele frequency misspecification we also used the program sib_ibd from the ASPEX package, which restricts linkage analysis to ASPs with unambiguous IBD status. For families with more than two (r>2) affected sibs, ASPEX programs can use either all possible pairs ($r[r-1]/2$) or only the $r-1$ independent sib pairs obtained from a sibship by pairing all other sibs with the first sib. Since creating all pairs from larger sibships may underestimate the p values in the far tail of the statistic distribution [Davis and Weeks, 1997; Ebers et al., 1996], we computed multipoint LOD score values using only the independent sib-pairs.

Multipoint linkage analysis was also undertaken with the GENEHUNTER-PLUS program [Kong and Cox, 1997; Kruglyak et al., 1996,] which allows non-parametric linkage (NPL) analysis of moderately sized pedigrees. The Z-all statistic compares the observed IBD sharing among all affected family members with that expected under the null hypothesis of no linkage. We chose also to apply the NPL method because it allows for all 43 US/Australia multiplex pedigrees to be used and for non-sib affected relatives in the other data sets to be included. The Z score statistic is assumed to asymptotically follow a standard normal distribution. However, when data are not fully informative, the reported p-values overestimate true significance levels [Kruglyak et al., 1996]. A less conservative method has been implemented into the GENEHUNTER-PLUS program [Kong and Cox, 1997]. It calculates a likelihood ratio-based statistic as a function of the IBD sharing among all affected relatives (Zlr).

Multipoint analysis in combined data sets was carried out preserving the specific marker allele frequencies, i.e., using the allele frequencies estimated in each data set separately. For instance, when combining both replication data sets, each marker, k, was assumed to have a total number of alleles equal to nk+mk, where nk and mk are the total number of marker alleles in first and second replication data sets, respectively. Then, for each marker, k, observed genotypes in the second data set were modified so that new marker allele numbers were set equal to the observed allele number +nk.

RESULTS

The characteristics of the family data sets are given in Table I. The 53 families of the screening data set (NIMH-IM) have a total sample of 81 ASPs (63 independent), and about 70% of ASPs have both parents genotyped (78% = 63/81; 71% = 45/63). The first replication data set (NIMH-GI) is larger—69 families with 109 ASPs (87 independent)—but the proportion of ASPs with available parental marker data is much lower; both parents are genotyped in about 20% of ASPs (18% = 10/109; 22% = 19/87). In this data set, the majority of ASPs have only one parent genotyped (63% = 69/109; 59% = 51/87). The US/Australia data set is the smallest family sample: 66 ASPs (54 indepen-

TABLE I. Description of Nuclear Families With ASPs Obtained From NIMH-IM, NIMH-GI, and US/Australia SZ Multiplex Pedigrees

Panel	NIMH-IM				NIMH-GI				US/Australia			
Number of
Multiplex pedigrees		55				63				43		
Pedigrees with ⩾ 1 ASP		53				63				40		
Nuclear families with ⩾ 1 ASP		53				69				43		
# Parents with DNA available	2	1	0	Total	2	1	0	Total	2	1	0	Total
Number of affected sibs, r:
r = 2	30	17	1	48	17	26	11	54	20	5	8	33
r = 3	2	—	—	2	1	9	3	13	3	6	—	9
r = 4	2	—	—	2	—	1	—	1	1	—	—	1
r = 5	—	—	—	—	—	1	—	1	—	—	—	—
r = 6	1	—	—	1	—	—	—	—	—	—	—	—
Total nuclear families	35	17	1	53	18	37	14	69	24	11	8	43
Total affected sib-pairs	63	17	1	81	20	69	20	109	35	23	8	66
Total independent ASPs	45	17	1	63	19	51	17	87	29	17	8	54

dent) obtained from 43 nuclear families (40 multiplex families). The proportion of ASPs with both parents genotyped is intermediate (53% = 35/66; 54% = 29/54). The size of the combined replication data set (NIMH-GI plus US/Australia: 112 families; 141 independent ASPs) is about twice that of the NIMH-IM data set, but the rate of ASPs with complete parental marker data is about twofold lower (34%).

The results of the single point analyses are given in Table II. The NIMH-IM and NIMH-GI data sets have already been presented [Cao et al., 1997]. We have added additional markers in these samples: D6S1021 in NIMH-IM ($p > 0.10$) and D6S445 ($p > 0.10$), D6S1021 ($p = 0.0009$), and D6S472 ($p = 0.09$) in the NIMH-GI sample. The genomewide scan of the US/Australia data set has been recently presented [Levinson et al., 1998], including two (D6S1021 and D6S474) markers from our 6q map. Here, we have thus typed 10 new markers, resulting in a denser map with much higher information content. In the US/Australia data set, there is excess allele sharing for several markers but only D6S424 is nominally "significant," (LOD = 0.89, $p = 0.022$).

Multipoint results obtained from the ASPEX program sib_phase are shown in Figure 1, which shows LOD score values at each location point for the screening, the two replication data sets, and the combined replication data set. The NIMH-IM and NIMH-GI data sets show similar significance as previously obtained with MAPMAKER/SIBS [Cao et al., 1997]. In the screening (NIMH-IM) data set, the maximum LOD score is observed close to D6S416 (LOD = 2.17, $y = 64.6\%$). Maximum LOD scores greater than two are also obtained in the first replication data set (NIMH-GI), but the highest LOD score is observed at a more centromeric area, close to D6S301 (LOD = 3.18, $y = 65\%$). In the US/Australia data set the peak location is consistent with the NIMH-GI data set (LOD = 1.07, $y = 62\%$, $p = 0.013$). We combined the NIMH-GI and US/Australia data sets because they are both independent replication samples for evaluating an a priori hypothesis: linkage to schizophrenia with this 6q region. In the combined replication data set significance for
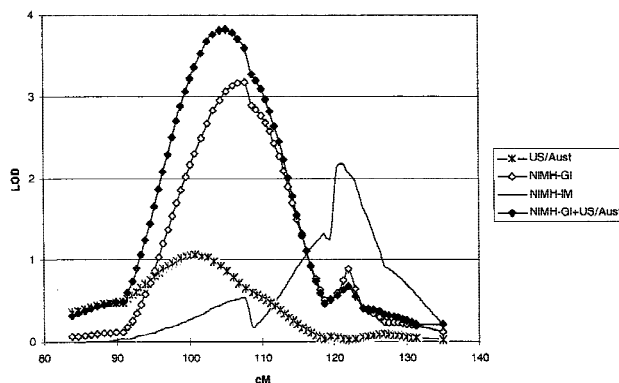

Fig. 1.   LOD score at each point location along the chromosome 6q area in ASPs from the NIMH-IM, NIMH-GI, US/Australia, and the combined NIMH-GI+US/Australia data sets, using the ASPEX sib_phase program (v. 1.12). Markers are ordered from the Genethon genetic map.

linkage is increased, as compared with either data set separately, in the 20-cM area delimited by D6S445 and D6S1021. The highest LOD score is found between D6S424 and D6S301 (LOD = 3.82 $y = 63.8\%$, $p = 0.000014$). The sibling recurrence risk attributed to this putative 6q susceptibility locus is estimated to be 1.92 ($z_0 = 0.13$).

Restricting the analysis to ASPs with unambiguous IBD information using ASPEX sib-ibd program provides similar figures (results not shown). The highest LOD score values are equal to 2.84, 3.5, 0.68, and 3.52 in the NIMH-IM, NIMH-GI, US/Australia and the combined NIMH-GI+US/Australia data sets, respectively. Overall, this analysis indicates that our statistical evidence for linkage of schizophrenia with chromosome 6q markers in the replication data sets are robust to misspecification of marker allele frequencies.

The same trends were also obtained when extending IBD sharing analysis to all affected relatives within the SCZ pedigrees. Nearly identical p-values were observed for the GENEHUNTER Z-all statistic and for the Zlr statistic (as computed by GENEHUNTER-PLUS), indicating that the marker data were sufficiently informative. Figure 2 reports the Z-all statistic

TABLE II.   Results of Pointwise Analysis (SIBPAIR program) of Chromosome 6q Markers in the Screening and the Two Replication Data Sets

Marker	Pos[a]	NIMH-IM			NIMH-GI			US/Australia		
		y[b]	LOD	p-val	y[b]	LOD	p-val	y[b]	LOD	p-val
D6S455	83.7	0.52	0.04	0.34	0.52	0.08	0.27	0.54	0.11	0.24
D6S445	91.2	0.51	0.00	0.44	0.53[c]	0.13	0.22	0.54	0.08	0.27
D6S424	102.4	0.57	0.44	0.078	0.66	2.59	0.00028	0.62	0.89	0.022
D6S301	110.0	0.63	1.39	0.0057	0.62	1.60	0.0033	0.55	0.22	0.16
D6S1021	110.9	0.52[c]	0.02	0.37	0.64[c]	2.10	0.0009	0.51	0.01	0.41
D6S416	122.0	0.65	1.60	0.0033	0.60	1.00	0.016	0.50	0.00	—
D6S474	122.8	0.64	1.81	0.0020	0.56	0.56	0.050	0.50	0.00	—
D6S454	123.8	0.60	0.75	0.031	0.57	0.71	0.036	0.50	0.00	—
D6S423	125.3	0.65	1.49	0.0045	0.64	2.15	0.00083	0.50	0.00	—
D6S267	127.4	0.62	1.22	0.0090	0.58	0.73	0.033	0.52	0.03	0.36
D6S408	130.6	0.53	0.05	0.31	0.55	0.18	0.18	0.56	0.15	0.20
D6S472	135.7	0.50	0.00	—	0.56[c]	0.39	0.089	0.53	0.05	0.32

[a]Sex-average position, centiMorgans, on chromosome 6q, derived from Généthon map.
[b]Estimated value of IBD sharing rate.
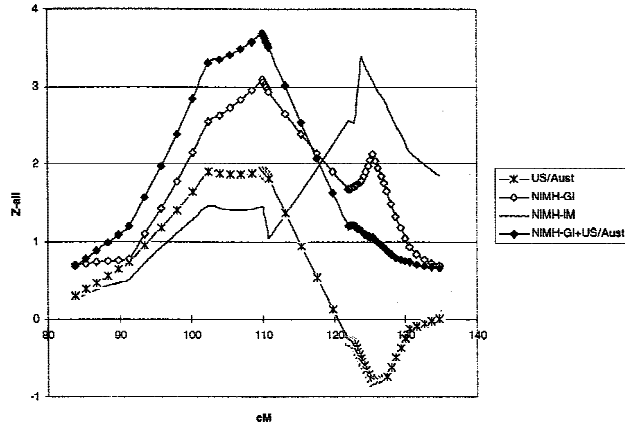[c]Marker data not reported in Cao et al. [1997].

Fig. 2. Multipoint Z-all scores at each point location along the chromosome 6q in the NIMH-IM, NIMH-GI, US/Australia, and the combined NIMH-GI + US/Australia family data sets using the GENEHUNTER-PLUS program (v. 1.1).

or NPL score. As before, the highest evidence for linkage in the replication data sets is located in a more centromeric area than that observed in the screening data set. In the NIMH-IM sample the highest statistic value is located at 124 cM (Z-all = 3.38/Zlr = 3.78). In the NIMH-GI and the US/Australia samples, the highest values are located at 110 cM (Z-all = 3.09/Zlr = 3.06) and 103 cM (Z-all = 1.90/Zlr = 1.97), respectively. In the overall replication data set, the highest evidence for linkage is located at 110 cM (Z-all = 3.7/Zlr = 3.6, p = 0.00014).

All three samples reported here have a heterogeneous ethnic background. Most families are Caucasian in the NIMH-IM and US/Australia samples. The NIMH-GI sample has the highest percentage (35%) of families of African-American descent. Because the GI group has conducted its analyses separately for European-American and African-American subsets of families, and has presented differing results for the two subsets in some regions [Faraone et al., 1997], we also separated the NIMH-GI pedigrees into these two subgroups to determine whether the evidence favoring linkage on chromosome 6q differed by racial subgroup. Table III shows for several of the markers genotyped in the NIMH-GI data set the results of the SIBPAIR program for the Caucasian and African-Americans subsamples. For these calculations, we computed allele frequencies separately in each group. As is evident

from Table III, there is excess sharing in both subsamples, with the significance levels being proportional to the sample size. Thus, we conclude that the putative susceptibility locus for schizophrenia is not specific to one population. The NIMH-IM and US/Australia data sets do not have enough African-American families to allow a test for differences.

DISCUSSION

The major finding in this study is that an additional independent sample of schizophrenia families has shown evidence consistent with excess sharing in this region of 6q but the statistical evidence is weaker (y = 62%, p = 0.013) than in the other two data sets. Because family data sets may not have comparable power (they differ in terms of sample size, marker data information rate of ASPs with missing parental genotypes, etc.), we have evaluated, through computer simulations, the power of each replication data set (given the observed family structures, observed genotypes for parents and unaffected sibs of affected offspring) to detect an excess of marker alleles IBD sharing (y) of 57, 60, or 64%. These are the values estimated under the pairwise affected sib-pair method (SIBPAIR program), in the NIMH-IM data set with the chromosome 6q markers D6S424, D6S454, and D6S474. These y values correspond to locus specific $\lambda s = 1.4$, 1.6, or 1.9, respectively. Table IV presents power estimates of the two replication data sets (rate of replicates out of 2,000 exceeding a maximum LOD score value of 1, 1.5, 2, 2.5, or 3), for the three λs values.

The two replication data sets do not have similar power. The power of NIMH-GI is almost twice that of the US/Australia data set. As expected, both data sets have good power to find LOD scores greater than 1.0 when y = 64%.

To our knowledge, linkage for schizophrenia to this 6q area has been partially tested by one group [Straub and Kendler, see Nurnberger and Foroud, 1998] but negative linkage results were reported in this fourth familial data set.

The issue of what constitutes a statistically significant finding and replication is a challenging problem for the genetic analyses of complex traits and raises controversial debate. Lander and Krugylak [1995] have proposed thresholds derived from genomewide significance thresholds. Following these guidelines, none of our analyzed samples (initial and replication ones) in-

TABLE III. SIBPAIR Results in the NIMH-GI Data Set by Ethnic Background

	African-Americans				Caucasian			
Marker	Shared	Not Shared	y[a]	p	Shared	Not Shared	y[a]	p
D6S474	26.2	23.1	0.55	0.22	46.3	38.2	0.56	0.14
D6S454	24.2	18.4	0.59	0.11	40.4	37.0	0.54	0.21
D6S423	30.9	22.8	0.59	0.054	43.1	23.9	0.66	0.004
D6S416	27.1	19.0	0.60	0.056	31.7	26.3	0.56	0.14
D6S408	23.1	25.4	—	—	29.0	23.2	0.58	0.11
D6S301	29.5	21.2	0.59	0.07	45.4	27.1	0.63	0.011
D6S267	28.4	19.7	0.60	0.056	38.5	33.5	0.54	0.19

[a]Estimated value of IBD sharing rate.

TABLE IV. Power Simulation Results in NIMH-GI and US/Australia Data
Sets, Assuming IBD Sharing Rate, y = 57, 60, or 64%, i.e., Assuming SCZ Loci
With Genetic Effects of 1.4, 1.6, or 1.9

| | | P(LOD≥T) | | | | |
		T = 1 (%)	T = 1.5 (%)	T = 2 (%)	T = 2.5 (%)	T = 3 (%)
y = 57%	NIMH-GI	22	10	5	1.7	0.8
	US/Australia	13	5	2	0.8	0.3
y = 60%	NIMH-GI	45	28	17	9	5
	US/Australia	30	17	7	4	2
y = 64%	NIMH-GI	79	60	44	31	21
	US/Australia	51	33	19	11	7

dependently provided evidence for significant linkage (i.e., *p*-values lower than the threshold of $2.2 10^{-5}$). Note, however, that in our combined replication data set the nominal *p*-value ($1.4 10^{-5}$) exceeds this threshold. For linkage analyses of multifactorial traits, the use of such guidelines, instead of nominal *p*-values, remains controversial [Curtis, 1996; Witte et al., 1996]. Several criticisms have been raised about the accuracy of these guidelines because, for instance, they do not account for the power of the analyzed samples. Furthermore, whether one should apply the same criteria for significance to localize a gene and to replicate its localization is another controversial issue. It has been argued that in a replication study, the nominal *p*-value required is lower than that required in the initial study [Lander and Krugylak, 1995; Lernmark and Ott, 1998].

In the US/Australia data set the peak location, close to D6S424, is consistent with the NIMH-GI data set (close to D6S301). In the total replication data set the overall peak location is still located between these two markers (at 104 cM, LOD = 3.82, *p* = 0.000014), which is about 17 cM centromeric to the peak obtained in the NIMH-IM data set (D6S416). One could question whether this is a consistent replication. However, it has been shown by simulation that the location of the true peak can vary substantially, especially for a locus-specific λs < 2.0 [Hauser and Boehnke, 1997]. Unfortunately, this additional US/Australia replication sample has not allowed us to narrow down this linkage region. In the total replication data set, LOD score values greater than three are obtained within a 13-cM area

(from position 99 cM to 112 cM), which remains at about 9 cM from the peak location obtained in the screening data set (see Fig. 1). When combining all three data sets, the 6q area with LOD score values greater than three is only slightly further narrowed (Fig. 3): this area has now a length of 12 cM (from position 100 cM to 112 cM), and the peak location is at roughly the same position as that obtained in the overall replication data set, that is, at 105 cM. This suggests that very large sample sizes will be needed to narrow down this region by nonparametric linkage methods. Study of the etiological candidate genes in this region is ongoing.

REFERENCES

American Psychiatric Association. 1987. Diagnostic and statistical manual of mental disorders. Washington, DC: APA.

Andreasen NC, Flaum M, Arndt S. 1992. The comprehensive assessment of symptoms and history (CASH): an instrument for assessing diagnosis and psychopathology. Arch Gen Psychiatr 49:615–623.

Cao Q, Martinez M, Zhang J, Sanders AR, Badner JA, Cravchik A, Markey CJ, Beshah E, Guroff JJ, Maxwell ME, Kazuba DM, Whiten R, Goldin LR, Gershon ES, Gejman PV. 1997. Suggestive evidence for a schizophrenia susceptibility locus on chromosome 6q and a confirmation in an independent series of pedigrees. Genomics 43:1–8.

Curtis D. 1996. Genetic dissection of complex traits. Nat Genet 12:356–357.

Davis S, Weeks DE. 1997. Comparison of nonparametric statistics for detection of linkage in nuclear families: single marker evaluation. Am J Hum Genet 61:1431–1444.

Ebers G, Kukay K, Bulman D, Sadovnick A, Rice G, Anderson C, Armstrong H, et al. 1996. A full genome search in multiple sclerosis. Nat Genet 13:472–476.

Endicott J, Spitzer RL. 1978. A diagnostic interview: the schedule for affective disorders and schizophrenia. Arch Gen Psychiatr 35:837–844.

Faraone SV, Blehar M, Pepple J, Moldin SO, Norton J, Nurnberger JI, Malaspina D, Kaufmann CA, Reich T, Cloninger CR, DePaulo JR, Berg K, Gershon ES, Kirch DG, Tsuang MT. 1996. Diagnostic accuracy and confusability analyses: an application to the Diagnostic Interview for Genetic Studies. Psychol Med 26:401–410.

Faraone SV, NIMH Genetics Initiative—Millemnium Schizophrenia Consortium. 1997. A genome scan of the European-American schizophrenia pedigrees of the NIMH Genetics Initiative. Am J Med Genet 81:557.

Gejman PV, Martinez M, Cao Q, Friedman E, Berrettini WH, Goldin LR, Koroulakis P, et al. 1993. Linkage analysis of fifty-seven microsatellite loci to bipolar disorder. Neuropsychopharmacology 9:31–40.

Gershon ES, DeLisi LE, Hamovit J, Nurnberger JI, Maxwell ME, Schreiber J, Dauphinais D, et al. 1988. A controlled family study of chronic psychoses: Schizophrenia and schizoaffective disorder. Arch Gen Psychiatr 45:328–336.

Hauser ER, Boehnke M. 1997. Confirmation of linkage results in affected-sib-pair linkage analysis for complex traits. Am J Hum Genet 61:A278.

Hinds D, Risch N. 1996. The ASPEX package: affected sib-pair mapping. ftp://lahmed.stanford.edu/pub/aspex.
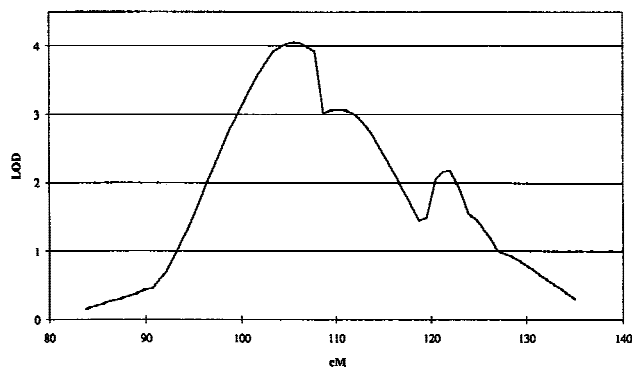
Fig. 3.   LOD score at each point location along the chromosome 6q area obtained from ASPEX sib_phase program (v. 1.12) in the total family sample data sets, NIMH-IM + NIMH-GI + US/Australia.

Knapp M, Seuchter S, Baur M. 1994. Linkage analysis in nuclear families. II. Relationship between affected sib-pair tests and lod score analysis. Hum Hered 44:44–51.

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

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.

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

Lathrop GM, Lalouel JM. 1984. Easy calculations of LOD scores and genetic risks on small computers. Am J Hum Genet 36:460–465.

Lermark A, Ott J. 1998. Sometime it's hot, sometimes it's not. Nat Genet 3:213–214.

Levinson DF, Mahtani MM, Nancarrow DJ, Brown DM, Kruglyak L, Kirby A, Hayward NK, Crowe RR, Andreasen NC, Black DW, Silverman JM, Endicott J, Sharpe L, Mohs RC, Siever LJ, Walters MK, Lennon DP, Jones HL, Nertney DA, Daly MJ, Mowry BJ. 1998. A genome scan of schizophrenia. Am J Psychiatr 6:741–750.

Mowry BJ, Nancarrow DJ, Lennon DP, Sandkuijl LA, Crowe RR, Silver-

man JM, Mohs RC, Siever LJ, Endicott J, Sharpe L, Walters MK, Hayward NK, Levinson DF. 1995. Schizophrenia susceptibility and chromosome 6p24-22. Nat Genet 11:233–234.

Nurnberger J, Blehar MC, Kaufmann CA, York-Cooler C, Simpson SG, Harkavy-Friedman J, Severe JB et al. 1994. Diagnostic interview for genetic studies. Rationale unique features, and training. NIMH Genetics Initiative. Arch Gen Psychiatr 51:849–859.

Nurnberger JI, Forout T. 1998. Chromosome 6 workshop. Psychiatr Genet 8:79–83.

Risch N. 1990. Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 46:222–228.

Roy MA, Lanctot G, Merette C, Cliche D, Fournier JP, Boutin P, Rodrigue C, Charron L, Turgeon M, Hamel M, Montgrain N, Nicole L, Pires A, Wallot H, Ponton AM, Garneau Y, Dion C, Lavallee JC, Potvin A, Szatmari P, Maziade M. 1997. Clinical and methodological factors related to reliability of the best-estimate diagnostic procedure. Am J Psychiatr 154:1726–1733.

Terwilliger J. 1996. Program SIBPAIR: sibpair analysis on nuclear families. ftp://linkage.cpmc.columbia.edu.

Witte JS, Elton RC, Schork NJ. 1996. Genetic dissection of complex traits. Nat Genet 12:355–356.