Novel Linkage to Chromosome 20p Using Latent Classes of Psychotic Illness in 270 Irish High-Density Families


Background: Several lines of evidence suggest that the clinical heterogeneity of schizophrenia is due to genetic heterogeneity. Genetic heterogeneity may decrease the signal-to-noise ratio in linkage and association studies. Therefore, linkage studies of clinically homogeneous classes of psychotic illness may result in greater power to detect at least some loci.

Methods: Latent class analysis was used to divide psychotic subjects from 270 Irish high-density families (N = 755) into six classes based on the Operational Criteria Checklist for Psychotic Illness. We heuristically call them Bipolar, Schizoaffective, Mania, Schizomania, Deficit Syndrome, and Core Schizophrenia. The latter four had prevalences of greater than .08 and were individually tested for linkage in a 20-cM nonparametric autosomal genomewide scan. Empirical significance was determined using 200 simulated genome scans.

Results: Seven regions achieved empirical criteria for suggestive significance for at least one latent class: 5q23.2-q35.3, 8q13.1-q23.1, 10q23.33-q26.3, 12q21.2-q24.32, 19q13.32-q13.43, 20p13-q22.3, and 21q11.2-q22.3. Five of 200 simulated scans resulted in seven suggestively significant loci (experiment-wide p = .03). Furthermore, at 20p13-p12.2, the Mania and Schizomania classes individually achieved criteria, whereas Deficit Syndrome had a suggestive logarithm of the odds peak 28 cM centromeric to this locus.

Conclusions: Using empirically derived, clinically homogeneous phenotypes, four chromosomal regions were suggestively linked but provided little evidence of linkage using traditional operationalized criteria. This approach was particularly fruitful on chromosome 20, which had previously yielded little evidence of linkage. Future studies of psychiatric illness may increase their ability to detect linkage or association by using clinically homogeneous phenotypes.

Key Words: Chromosome 20; genetic; linkage; schizophrenia; subtypes

Schizophrenia is a devastating neuropsychiatric illness with an estimated annual cost of $32 billion in the United States (1). Before the advent of Kraepelin’s unifying concept of dementia praecox, the nosology of psychotic illness embraced several distinct entities. Since Bleuler (2), controversy over whether schizophrenia is a single or multiple diseases has persisted. Although it is unchallenged that the illness can comprise a range of more or less clinically distinct presentations, it is unknown whether these result from more than one etiopathogenic processes. This question has been of greater interest over the last several years because susceptibility genes have been identified through positional cloning and subsequently replicated in independent samples (reviewed by Harrison and Weinberger) (3).

Schizophrenia, like other complex diseases, has long been suspected to be genetically heterogeneous, and a body of recent studies provides some empirical evidence for this. For example, although genetic variation in DTNB1 (dysbindin) has been demonstrated to impart risk in several samples (4–10), different haplotypes have been associated in different populations. Family-based and case-control samples even from the same country imply different risk haplotypes, suggesting allelic heterogeneity (6,9).

If clinical heterogeneity is indeed a consequence of genetic heterogeneity, this could have considerable implications both for psychiatric nosology and for elucidating the genetic architecture of schizophrenia. From a more pragmatic perspective in the interest of gene finding, examining clinical subforms of illness as phenotypes of interest could increase the signal-to-noise ratio in linkage and association studies and provide increased power to identify susceptibility genes.

Several observations suggest that this may be the case. Two studies have reported that subjects from families linked to specific genomic regions may differ clinically from other subjects (11,12). In addition, subjects with high-risk alleles in DTNB1 may have higher levels of negative symptoms (13,14). However, a discovery-based approach that attempts to detect genetic linkage to empirically determined, symptomatically homogeneous, nonoverlapping classes of illness has not been implemented. In this report, we describe genomewide linkage analyses of four classes of psychotic illness that were empirically derived from clinical and course features using latent class analysis in the Irish Study of High-Density Schizophrenia Families (ISHDSF). A genome scan has been previously published using this sample based on traditional but clinically heterogeneous diagnostic criteria (15), allowing us to examine whether additional loci could be detected using clinically homogeneous phenotypes.

Methods and Materials

Subjects and Assessment

Fieldwork for the Irish Study of High Density Schizophrenia Families (ISHDSF) was performed between April 1987 and November 1992 and described previously (16). Interviews were conducted by Irish psychiatrists and social scientists with a background in mental health or survey work after consent was obtained by using procedures approved by the ethical review panel.
panels at the Health Research Board and the Queen's University. The original linkage sample consisted of 1425 individuals from 270 families who were ascertained on the basis of having more than one member with DSM-III-R schizophrenia or poor outcome schizoaffective disorder.

Diagnoses were generated using modified sections of the Structured Interview for DSM-III-R (SCID) for selected Axis I disorders (17). All relevant diagnostic information for each individual relative was reviewed, blind to pedigree assignment and marker genotypes, independently by KSK and DM. Each diagnostician made up to three best-estimate DSM-III-R diagnoses. For each subject with a probable lifetime history of hallucinations or delusions, the Operational Criteria Checklist for Psychotic Illness (OPCRIT) (18) was completed by KSK (N = 755, of which 722 were genotyped) based on review of hospital records and personal interviews. This is an instrument designed for use in a best-estimate procedure that codes symptom and course features as assessed over the entire course of illness. It was designed to allow the experienced clinician to integrate the relative prominence of clinical features over the entire course of illness.

Latent Class Analysis

Latent class analysis (LCA) is a method that is sometimes described as a categorical analog of factor analysis. The premise of this method is that all instances of illness in a population belong to a finite set of mutually exclusive and exhaustive classes. Class membership completely determines and explains the distribution of responses to scale items, as well as correlations between items.

We developed our latent class typology from five OPCRIT factor-derived scales that have been previously described (13). These included scales for hallucinations, delusions, and negative, manic, and depressive symptoms. In this study, we also included age of onset, and scales for Schneiderian symptoms, and premorbid and social dysfunction. A scale was constructed based on OPCRIT items representing Schneiderian symptoms, which were selected on the basis of a priori definitions. These included the following items: delusions of influence, bizarre delusions, delusions of passivity, thought insertion, thought withdrawal, thought broadcast, third-person auditory hallucinations, running commentary voices, and abusive, accusatory, or persecutory voices. All of these loaded on one of two of our original OPCRIT factors—hallucinations or delusions. Premorbid and social dysfunction was derived from six items of the OPCRIT representing psychosocial and course variables. These included being single; being unemployed; duration of illness of at least 6 months; duration of prodromal, acute, and residual stages of at least 6 months; poor premorbid social adjustment; and premorbid personality disorder. These items loaded on one factor in factor analysis using Proc Factor in SAS (SAS Institute, Cary, North Carolina) (19).

We dichotomized all eight measures using a median split in each, resulting in a matrix of 755 individual subjects and eight dichotomous variables. The program LEM (20) was used to assign class membership based on the likelihood of subjects’ particular item profiles. We examined solutions consisting of two to eight classes. Subjects were assigned to a particular class if they had a probability of .5 or greater of membership in it. Because our goals were not only to reduce the clinical heterogeneity of our phenotypes but also to maximize the power to detect linkage, we aimed to have the largest sample sizes possible per class while excluding subjects with low probabilities of membership in any class. Because this approach, compared with entering the OPCRIT items themselves directly into LCA, entails data reduction, it is inevitable that some information is lost. However, entering all OPCRIT items into analysis would have resulted in a large number of classes, which would have been difficult to interpret on clinical grounds and would have had low power to detect linkage because of the small sample size.

We determined familial aggregation of each latent class as external validation of the particular latent class solution used to classify affected subjects. In addition, it was important to demonstrate that familial factors determined class membership to justify a genome scan of the individual classes. This was done by calculating a simple kappa, as well as tetrachoric correlations for sibling pair membership in each class. Both tests were operationalized in Proc Freq in SAS (19).

Genotyping

A detailed description of the genome scan strategy has been published (15). Briefly, a sample of 270 families in which at least two members were affected with psychotic illness was divided into three equal subsets. The scan used 684 markers, with 196 being shared across individual subsets. Average intermarker distances were 9.3, 9.8, and 9.6 cM for the three sets. There were several follow-up regions in which intermarker distance was less and in which all three subsets were genotyped. Finally, in the course of this and other analyses, we “filled in” markers that had been genotyped in only one or two family sets by genotyping them in the remaining families. This occurred at chromosomes 4p, 9q, 14q, 20p, and 21q. Markers were predominantly tri- and tetranucleotide repeat microsatellites generated by the Cooperative Human Linkage Center, and many of them are in the Weber screening set, version 8.0 (http://research.marshfieldclinic.org/genetics/GeneticResearch/sets/scrsset8.txt). Genotyping methods have been previously described (21,22).

Linkage Analysis

We employed autosomal genomewide multipoint nonparametric linkage analysis of four of six latent classes. The nonparametric logarithm of the odds (NPLq) was calculated using all relative pairs at every centimorgan using all available information. NPL-Z scores were converted to Kong and Cox logarithm of the odds (LOD) scores (23) in Merlin (http://www.sph.umich.edu/csg/abecasis/Merlin/index.html) (24). Allele frequencies were calculated from pedigree founders only. Before linkage analysis, all genotypes were checked for non-Mendelian inheritance and excessive recombination events using Merlin. Erroneous genotypes were removed. NPL scores were calculated at 1-cM intervals throughout the genome. This was to rule out the possibility of inflation of LOD scores due to possible inbreeding (25). We used the “--rsq” option in Merlin to rule out the inflation of linkage statistics due to marker-to-marker linkage disequilibrium (LD). This treats markers that are in LD with each other, as defined by an r² threshold, as a “cluster.” We used a threshold of .05, as recently suggested by Levinson and Holmans (26). However, no markers in the map used in this study formed clusters meeting this r² threshold.

Determining Empirical Significance of Linkage Results

Genomewide significance levels were calculated from 200 simulated genomes constructed using the gene-dropping method in Merlin (24). Each simulated data set used the original phenotypes and generated simulated genotypes with the same allele frequen-
Moving from a one-class (AIC total sample. Of note, we observed the greatest change in AIC in actively common classes, each comprising 17.69% or more of the than the eight-class solution (H11002). This solution resulted in the most highly negative AIC, at 165.41. Although it was less than 1 unit of AIC more negative than the eight-class solution (H11002), it contained four rela-
tionships of cases of DSM-III-R schizophrenia, disorganized type, were in Table 2. These proportions were largely as expected. For example, the vast majority of DSM-III-R bipolar disorder cases were in the Schizomaniac and Manic classes, and the vast majority of cases of DSM-III-R schizophrenia, disorganized type, were in the Deficit Syndrome and Core Schizophrenia classes.

**Familial Aggregation of Latent Classes**
A simple (unweighted) $\kappa$ indicated modest but significant sibling resemblance for class overall ($\kappa = .12$, 95% confidence intervals [CI] = .06–.17). Among individual classes, the only ones with significant tetrarchic correlations were Schizomania ($r = .19$, $p = .04$), Deficit Syndrome ($r = .29$, $p = .003$), and Core Schizophrenia ($r = .22$, $p = .02$). Bipolar ($r = .25$, $p = .06$), and Mania ($r = .18$, $p = .08$) attained trend significance, but Schizoaffective ($r = .13$, $p = .42$) did not.

**Empirical Significance Levels**
The LOD scores corresponding to “suggestive significance” (i.e., less than one expected per genome scan under the null hypothesis) were calculated on the basis of the simulated genome scans of Schizomania, Mania, Deficit Syndrome, and Core Schizophrenia. The Schizoaffective and Bipolar classes were not analyzed because they each represented less than 8% of the sample. These LOD scores were, respectively, 1.58, 1.32, 1.49, and 1.60. However, as we tested multiple phenotypes (four latent classes), it was important to model this multiple testing into

---

**Table 1.** Latent Classes of Psychotic Illness and Their Probabilities of Being Above and Below the Sample: Medians of Eight Symptomatic and Course Variables

<table>
<thead>
<tr>
<th>Latent Class</th>
<th>Schizoaffective</th>
<th>Bipolar</th>
<th>Schizomania</th>
<th>Mania</th>
<th>Deficit Syndrome</th>
<th>Core Schizophrenia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delusions Below</td>
<td>.003</td>
<td>.683</td>
<td>.167</td>
<td>.907</td>
<td>.975</td>
<td>.135</td>
</tr>
<tr>
<td>Hallucinations Below</td>
<td>.997</td>
<td>.318</td>
<td>.833</td>
<td>.093</td>
<td>.025</td>
<td>.865</td>
</tr>
<tr>
<td>Negative Below</td>
<td>.277</td>
<td>.473</td>
<td>.211</td>
<td>.815</td>
<td>.570</td>
<td>.103</td>
</tr>
<tr>
<td>Depressive Below</td>
<td>.723</td>
<td>.527</td>
<td>.789</td>
<td>.185</td>
<td>.430</td>
<td>.897</td>
</tr>
<tr>
<td>Manic Below</td>
<td>.427</td>
<td>.224</td>
<td>.960</td>
<td>.818</td>
<td>.264</td>
<td>.217</td>
</tr>
<tr>
<td>Schneiderian Below</td>
<td>.573</td>
<td>.777</td>
<td>.040</td>
<td>.182</td>
<td>.736</td>
<td>.783</td>
</tr>
<tr>
<td>Above</td>
<td>.000</td>
<td>.196</td>
<td>.515</td>
<td>.530</td>
<td>.843</td>
<td>.762</td>
</tr>
<tr>
<td>Above</td>
<td>1.000</td>
<td>.804</td>
<td>.485</td>
<td>.470</td>
<td>.157</td>
<td>.238</td>
</tr>
<tr>
<td>Above</td>
<td>.145</td>
<td>.000</td>
<td>.307</td>
<td>.340</td>
<td>.755</td>
<td>.736</td>
</tr>
<tr>
<td>Above</td>
<td>.855</td>
<td>1.000</td>
<td>.693</td>
<td>.661</td>
<td>.245</td>
<td>.264</td>
</tr>
<tr>
<td>Above</td>
<td>.000</td>
<td>1.000</td>
<td>.982</td>
<td>1.000</td>
<td>.114</td>
<td></td>
</tr>
<tr>
<td>Above</td>
<td>1.000</td>
<td>.000</td>
<td>.920</td>
<td>.018</td>
<td>.000</td>
<td>.886</td>
</tr>
<tr>
<td>Course Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premorbid/Social Dysfunction Below</td>
<td>.118</td>
<td>.186</td>
<td>.678</td>
<td>.803</td>
<td>.299</td>
<td>.361</td>
</tr>
<tr>
<td>Above</td>
<td>.882</td>
<td>.815</td>
<td>.322</td>
<td>.198</td>
<td>.701</td>
<td>.639</td>
</tr>
<tr>
<td>Age at Onset Below</td>
<td>.704</td>
<td>.822</td>
<td>.406</td>
<td>.236</td>
<td>.471</td>
<td>.584</td>
</tr>
<tr>
<td>Above</td>
<td>.297</td>
<td>.178</td>
<td>.594</td>
<td>.764</td>
<td>.529</td>
<td>.416</td>
</tr>
</tbody>
</table>

The six latent classes derived in this study are given, along with their respective clinical and course profiles. All latent classes are defined by their probabilities of falling above or below the sample medians of the five symptomatic Operational Criteria Checklist for Psychotic Illness factors, as well as both course variables. For example, the subjects in the Schizoaffective class have a high probability of being above the sample median on the delusions, hallucinations, negative, depressive, manic, and Schneiderian factors. Subjects in the Schizomania class have a similar profile, except they have a high probability of being below the sample median on the negative and depressive factors.
our tests for empirical significance. Therefore, in each of our 200 simulated data sets, the same four classes were retained and tested. Experiment-wide significance was calculated by comparing the number of suggestive peaks seen in our data to the number expected based on the simulations. We observed nine chromosomes achieving this level of significance, for any latent number expected based on the simulations. We observed nine ing the number of suggestive peaks seen in our data to the tested. Experiment-wide significance was calculated by compar-
simulated data sets, the same four classes were retained and our tests for empirical significance. Therefore, in each of our 200

**Linkage Analysis of Individual Latent Classes**

We performed multipoint NLA using the Schizomania, Mania, Deficit Syndrome, and Core Schizophrenia classes as the phenotypes of interest. All loci achieving suggestive significance for any individual class, including marker names, centimorgan positions, and LODs, are presented in Table 3.

We were first interested in the impact of the latent class typology on evidence for linkage previously obtained in studies of schizophrenia and other psychotic syndromes in this sample. Chromosome 5q was the only region with considerable previous evidence of linkage both in this sample and several others and that also contained a locus suggestively linked to a specific latent class. Core Schizophrenia attained a LOD of 1.76 at 5q35.1-q35.2. However, this was about 45 cM telomeric to marker D5S804, which was the most highly linked marker in the region using traditional diagnostic schemes.

Table 2. DSM-III-R Diagnoses Within Four Latent Classes Used as Phenotypes in Genomewide Linkage Analysis

<table>
<thead>
<tr>
<th>DSM-III-R Diagnosis</th>
<th>Latent Class (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bipolar Disorder</td>
<td>2.61</td>
</tr>
<tr>
<td>Major Depression</td>
<td>0.65</td>
</tr>
<tr>
<td>Opioid Dependence</td>
<td>0.00</td>
</tr>
<tr>
<td>Schizoaffective Disorder</td>
<td>25.49</td>
</tr>
<tr>
<td>Psychosis, Not Otherwise Specified</td>
<td>0.65</td>
</tr>
<tr>
<td>Schizophreniform Disorder</td>
<td>2.61</td>
</tr>
<tr>
<td>Delusional Disorder</td>
<td>0.00</td>
</tr>
<tr>
<td>Undifferentiated Schizophrenia</td>
<td>39.22</td>
</tr>
<tr>
<td>Paranoid Schizophrenia</td>
<td>28.76</td>
</tr>
<tr>
<td>Catatonic Schizophrenia</td>
<td>0.00</td>
</tr>
<tr>
<td>Disorganized Schizophrenia</td>
<td>0.00</td>
</tr>
<tr>
<td>Simple Schizophrenia</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**Distribution of DSM-III-R diagnoses within four latent classes used as phenotypes in genome-wide linkage analysis.** Figures represent percent of total cases within each respective latent class.

Table 3. Genomic Regions Exceeding Criteria for Suggestive Linkage

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Class</th>
<th>Maximum LOD</th>
<th>Cytogenetic Position</th>
<th>Markers</th>
<th>Classwise p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>4</td>
<td>1.76</td>
<td>5q23.2-q35.3</td>
<td>DSS818-DSS408</td>
<td>.025</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>1.84</td>
<td>8q13.1-q23.1</td>
<td>DSS136-DSS180</td>
<td>.032</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>1.64</td>
<td>10q23.33-q26.3</td>
<td>DTS677-DTS212</td>
<td>.0334</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>1.59</td>
<td>12q21.2-q24.32</td>
<td>DTS219-DTS278</td>
<td>.0439</td>
</tr>
<tr>
<td>19</td>
<td>3</td>
<td>1.71</td>
<td>19q13.32-q13.43</td>
<td>DTS652-DTS254</td>
<td>.0341</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>2.01</td>
<td>20p13-q11.23</td>
<td>DTS473-DTS478</td>
<td>.0153</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>1.59</td>
<td>20p13-q12.2</td>
<td>DTS473-DTS160</td>
<td>.0382</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>2.57</td>
<td>20p12.3-q13.12</td>
<td>DTS155-DTS481</td>
<td>.0036</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>1.93</td>
<td>21q11.2-q22.3</td>
<td>DTS191-DTS144</td>
<td>.018</td>
</tr>
</tbody>
</table>

Genomic regions exceeding empirical criteria for suggestive linkage are listed, along with their respective latent class. LOD = logarithm of the odds.

*Corresponds to d + 1/n + 1, where d = number of chromosomes achieving a higher LOD in n = 4400 chromosomes tested, in all 200 simulated genome scans. Not corrected for the testing of multiple classes.
Discussion

This is the first linkage study of schizophrenia that we know of to use categories of affection that are based on empirically derived clinical subtypes defined by symptomatic profiles rather than operationalized diagnostic criteria. This was done to attempt to detect chromosomal regions harboring susceptibility genes for specific psychotic syndromes but that may go undetected in a genome scan using traditional diagnostic criteria.

Linkage results obtained in this study were only partly consistent with results previously reported in this sample using operationalized diagnostic criteria. The one previously linked region that was supported was 5q. However, the peak attained by Mania on this chromosome was considerably telomeric to the original peak reported using traditional phenotypes (15,31). Because of the relatively sparse genotyping of this sample and the exploratory and post hoc nature of this analysis. Nevertheless, the presence of two independent classes with suggestive linkage to the same markers seems to provide considerably more evidence of a susceptibility gene in that region than one linked class alone.

Finally, we relaxed the conservative condition that individual class’s LOD peaks must be at the same position, which may be unrealistic, given results of previous simulation work (38). There were therefore three latent classes on chromosome 20 that individually exceeded thresholds for suggestive significance within a 28-cM region. This distance is well within the 95% confidence intervals for stochastic variation of linkage peaks reported by Roberts et al. (38). In our 200 simulated genomes, no chromosomes contained three suggestive peaks. Of these three classes, Deficit Syndrome differed phenotypically from both Mania and Schizomania in having low levels of manic symptoms and high levels of negative symptoms. However, there have now been several reports of bipolar disorder being associated with susceptibility genes first identified in positional cloning studies of schizophrenia. These include DISC1 (39), neuregulin-1 (40), and DAOA (formerly known as G72) (41–43).

The excess linkage of Deficit Syndrome compared with other classes is consistent with an established literature strongly suggesting that negative symptoms may have a greater biological substrate than other illness features. This is supported by findings of 2.57, 1.84, 1.71, and 1.57, we observed one, two, three, and four chromosomes, respectively, while we would expect .09, .50, .66, and .84 chromosomes on the basis of the simulated scans. This was followed by Schizomania and Core Schizophrenia (two chromosomes each), then Mania (one chromosome). Interestingly, this order mirrors that for statistical significance of sibling pair resemblance for the class.

Evidence of linkage to chromosome 20p achieved suggestive significance for more than one individual class, whereas this region had little previous support in this sample (15). Of note, the two classes linked to the same locus in this region, Mania and Schizomania, are both associated with high levels of manic symptoms. One previously published genome scan of bipolar disorder, of which manic symptoms are the hallmark, produced evidence of linkage to chromosome 20p12.1 (34). This region was also supported in a meta-analysis of most of the world’s published schizophrenia genome scans (35), whereas in one scan, this was the most highly linked region in the genome (36).

A recent study in a Japanese sib-pair sample using a high-density single nucleotide polymorphism map reported linkage to 20p11.2 (37).

At a second locus, 28 cM centromeric to the first peak on chromosome 20, a third class, Deficit Syndrome, had the highest LOD of any latent class in the genome at 2.57 (Figure 1). Furthermore, Schizomania had a second peak in this location with a LOD of 1.78, about 4 cM centromeric to the peak for Deficit Syndrome. However, it had a LOD of only 1.53, slightly less than its threshold of 1.59, at precisely the same position as the Deficit Syndrome peak. To shed more light on the linkage behavior of these two classes, we collapsed both into one class and performed linkage analysis on the combined class, in an exploratory manner. This resulted in a maximum LOD of 5.53 (nominal p value calculated by Merlin as .0000, i.e., p < .000005) at the peak for Deficit Syndrome. Because this was a post hoc test, we employed a permutation procedure to shed more light on its empirical significance. Random subsets of the affected subjects, each having a sample size equal to that of the combined class, were selected, and linkage was subsequently performed using the original genotypes. Of 5000 permutations, none achieved a LOD of 5.53. However, we would not interpret this as genomewide significant linkage because of the exploratory and post hoc nature of this analysis. Nevertheless, the presence of two independent classes with suggestive linkage to the same markers seems to provide considerably more evidence of a susceptibility gene in that region than one linked class alone.

Figure 1. Plot of logarithm of the odds (LOD) scores for Schizomania, Mania, Deficit Syndrome, and Core Schizophrenia classes on chromosome 20. Kong and Cox LOD scores for the latent classes Schizomania, Mania, Deficit Syndrome, and Core Schizophrenia are plotted against centimorgan position on chromosome 20. Both Schizomania and Mania achieved suggestive significance at marker at 20p13-p12.2, closest to marker D20S473. The genomewide empirical significance of this was .0498. At 28 cM centromeric to this locus, Deficit Syndrome also achieved suggestive significance. None of the simulated genome scans contained chromosomes in which three latent classes were suggestively linked.

Because of the relatively sparse genotyping of this sample and the high levels of negative symptoms. However, there have now been several reports of bipolar disorder being associated with susceptibility genes first identified in positional cloning studies of schizophrenia. These include DISC1 (39), neuregulin-1 (40), and DAOA (formerly known as G72) (41–43).

The excess linkage of Deficit Syndrome compared with other classes is consistent with an established literature strongly suggesting that negative symptoms may have a greater biological substrate than other illness features. This is supported by findings...
of greater family history (44) and more severe brain structural abnormalities (45–48). Of more immediate relevance is a recent report of linkage to chromosome 6p being explained by a class of schizophrenia associated with cognitive dysfunction (49), which has been associated with negative symptoms in both chronic (50) and first-episode (51) patients. Furthermore, we have recently demonstrated that association with a DTNB1 high-risk haplotype was greater in subgroups of our sample characterized by high levels of negative symptoms (13).

Three more chromosomal regions were suggestively linked in this study but had produced little evidence of linkage previously: 12q (Deficit Syndrome), 19q (Deficit Syndrome), and 21q (Schizomania). Chromosome 21q22, although not previously demonstrating linkage to schizophrenia, has been linked to bipolar disorder in several studies (34,52–55). Furthermore, there has been a recent report of association between melatonin 2 (TRPM2) in this region and bipolar disorder (56). Our finding of linkage to Schizomania in this region suggests the presence of a gene that predisposes to a broader phenotype that includes not only bipolar disorder but also psychotic disorders with prominent manic symptoms.

These results support previous studies suggesting that clinical heterogeneity is due to genetic heterogeneity. Specifically, they provide preliminary evidence of the existence of genes that predispose to more or less specific clinical subtypes of illness. We have previously described these as Mixed Susceptibility–Modifier genes (57). These results strengthen the case recently made by Jablensky (58) that subclassing strategies may be useful in gene-finding studies of schizophrenia.

There are two main limitations specific to the analytic strategy pursued in this study, which make it even more critical to attempt replication in independent samples. First, multiple tests were performed by using several nonoverlapping liability classes, thereby increasing the probability of false positives. Although we attempted to evaluate their effects on the overall false-positive rate by examining 200 null data sets, none of the loci we tested achieved genomewide significance individually in any one class. Therefore, loci not previously supported are in need of replication in other samples. Second, our latent class typology may not generalize to other samples or populations, possibly complicating efforts at replication. Furthermore, the typology may have been influenced by the clinical rating instrument or the specifics of the statistical methods used. These could include the methods of selecting the number of factors and their rotation in factor analysis (59), as well as our dichotomization of factor-derived scores in latent class analysis and so on. We in fact attempted to enter all single OPCRIT items directly into latent class analysis (60(suppl 1):4 – 6).

15.Straub RE, MacLean CJ, Ma Y, Webb BT, Myakishev MV, Harris-Kerr C, et al. (2002): Genome-wide scans of three independent sets of 90 Irish multiplex schizophrenia families and follow-up of selected regions in all...

www.sobp.org/journal