

# Genomewide Linkage Scan of 409 European-Ancestry and African American Families with Schizophrenia: Suggestive Evidence of Linkage at 8p23.3-p21.2 and 11p13.1-q14.1 in the Combined Sample

Brian K. Suarez,<sup>1,\*</sup> Jubao Duan,<sup>2,\*</sup> Alan R. Sanders,<sup>2</sup> Anthony L. Hinrichs,<sup>1</sup> Carol H. Jin,<sup>1</sup> Cuiping Hou,<sup>2</sup> Nancy G. Buccola,<sup>3</sup> Nancy Hale,<sup>4</sup> Ann N. Weillbaeher,<sup>5</sup> Deborah A. Nertney,<sup>7,8</sup> Ann Olincy,<sup>9</sup> Susan Green,<sup>10</sup> Arthur W. Schaffer,<sup>1</sup> Christopher J. Smith,<sup>11</sup> Dominique E. Hannah,<sup>7,8</sup> John P. Rice,<sup>1</sup> Nancy J. Cox,<sup>6</sup> Maria Martinez,<sup>2,12</sup> Bryan J. Mowry,<sup>7,8</sup> Farooq Amin,<sup>10,13,†</sup> Jeremy M. Silverman,<sup>11</sup> Donald W. Black,<sup>4</sup> William F. Byerley,<sup>14,15,‡</sup> Raymond R. Crowe,<sup>4</sup> Robert Freedman,<sup>9</sup> C. Robert Cloninger,<sup>1</sup> Douglas F. Levinson,<sup>16,§</sup> and Pablo V. Gejman<sup>2</sup>

<sup>1</sup>Departments of Psychiatry and Genetics, Washington University, St. Louis; <sup>2</sup>Center for Psychiatric Genetics, Department of Psychiatry and Behavioral Sciences, Evanston Northwestern Healthcare and Feinberg School of Medicine, Northwestern University, Evanston, IL; <sup>3</sup>School of Nursing, Louisiana State University Health Sciences Center, New Orleans; <sup>4</sup>Mental Health Clinical Research Center and Department of Psychiatry, University of Iowa College of Medicine, Iowa City; <sup>5</sup>Department of Psychiatry and Behavioral Sciences, Feinberg School of Medicine, Northwestern University, and <sup>6</sup>Department of Medicine, University of Chicago, Chicago; <sup>7</sup>Queensland Centre for Schizophrenia Mental Health Research, The Park, Centre for Mental Health, Wacol, Australia; <sup>8</sup>Department of Psychiatry, University of Queensland, Brisbane; <sup>9</sup>Department of Psychiatry and Colorado Psychiatric Health, University of Colorado School of Medicine, Denver; <sup>10</sup>Atlanta VA Medical Center and Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta; <sup>11</sup>Department of Psychiatry, Mount Sinai School of Medicine, New York; <sup>12</sup>Méthodologie Statistique et Epidémiologie Génétique des Maladies Multifactorielles, Institut National de la Recherche et de la Santé Médicale, Evry, France; <sup>13</sup>Baylor College of Medicine, Houston; <sup>14</sup>Department of Psychiatry, University of California at Irvine, Irvine; <sup>15</sup>Department of Psychiatry, University of California–San Francisco, San Francisco; and <sup>16</sup>Department of Psychiatry, University of Pennsylvania, Philadelphia

We report the clinical characteristics of a schizophrenia sample of 409 pedigrees—263 of European ancestry (EA) and 146 of African American ancestry (AA)—together with the results of a genome scan (with a simple tandem repeat polymorphism interval of 9 cM) and follow-up fine mapping. A family was required to have a proband with schizophrenia (SZ) and one or more siblings of the proband with SZ or schizoaffective disorder. Linkage analyses included 403 independent full-sibling affected sibling pairs (ASPs) (279 EA and 124 AA) and 100 all-possible half-sibling ASPs (15 EA and 85 AA). Nonparametric multipoint linkage analysis of all families detected two regions with suggestive evidence of linkage at 8p23.3-q12 and 11p11.2-q22.3 (empirical  $Z$  likelihood-ratio score [ $Z_{lr}$ ] threshold  $\geq 2.65$ ) and, in exploratory analyses, two other regions at 4p16.1-p15.32 in AA families and at 5p14.3-q11.2 in EA families. The most significant linkage peak was in chromosome 8p; its signal was mainly driven by the EA families.  $Z_{lr}$  scores  $>2.0$  in 8p were observed from 30.7 cM to 61.7 cM (Center for Inherited Disease Research map locations). The maximum evidence in the full sample was a multipoint  $Z_{lr}$  of 3.25 (equivalent Kong-Cox LOD of 2.30) near *D8S1771* (at 52 cM); there appeared to be two peaks, both telomeric to *neuregulin 1* (*NRG1*). There is a paracentric inversion common in EA individuals within this region, the effect of which on the linkage evidence remains unknown in this and in other previously analyzed samples. Fine mapping of 8p did not significantly alter the significance or length of the peak. We also performed fine mapping of 4p16.3-p15.2, 5p15.2-q13.3, 10p15.3-p14, 10q25.3-q26.3, and 11p13-q23.3. The highest increase in  $Z_{lr}$  scores was observed for 5p14.1-q12.1, where the maximum  $Z_{lr}$  increased from 2.77 initially to 3.80 after fine mapping in the EA families.

Schizophrenia (SZ [MIM 181500]) is a disorder (or group of disorders) with onset typically in adolescence or young adulthood and characterized by disruption of thinking (e.g., delusions or disorganization), perception (hallucinations), mood, and behavior. The symptoms, if un-

treated, tend to persist, and the course of the disease tends to be chronic, even with treatment. Although psychotic symptoms can be detected in the general population, and a continuum of severity from normality to psychosis has been proposed (Strauss 1969; van Os et

Received November 14, 2005; accepted for publication December 6, 2005; electronically published January 3, 2006.

Address for correspondence and reprints: Dr. Pablo V. Gejman, Evanston Northwestern Healthcare Research Institute, 1001 University Place, Evanston, IL 60201-3137. E-mail: pgejman@northwestern.edu

\* These two authors contributed equally to this work.

† Present affiliation: Atlanta VA Medical Center and Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta.

‡ Present affiliation: Department of Psychiatry, University of California–San Francisco, San Francisco.

§ Present affiliation: Department of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto.

*Am. J. Hum. Genet.* 2006;78:315–333. © 2006 by The American Society of Human Genetics. All rights reserved. 0002-9297/2006/7802-0013\$15.00

al. 2000; Johns et al. 2004), SZ, as defined by the DSM-IV criteria of the American Psychiatric Association (1994), is widely viewed as a discrete illness. Heritability is ~80%, on the basis of twin studies (Cardno et al. 1999). There is a 10-fold increase in risk to first-degree relatives (Gottesman and Shields 1982), as well as a familial coaggregation of SZ with other psychotic disorders and with “schizophrenia spectrum” personality disorders (Kendler et al. 1993*b*, 1993*c*, 1993*d*; Stompe et al. 1998). SZ risk is also increased in adopted-away children who have a biological parent with SZ (Heston 1966; Kety et al. 1971; Rosenthal et al. 1971) and in the offspring of an unaffected MZ cotwin of a proband with SZ (Gottesman and Bertelsen 1989). Complex inheritance is suggested by the high MZ:DZ risk ratio (~4:1) and by the rapid decrease in risk as biological relatedness to a proband decreases (O’Rourke et al. 1982).

For complex disorders such as SZ, it has proven difficult to obtain replicable positive evidence of genetic linkage, but careful analysis of the SZ literature suggests that there is considerable support for a set of potential susceptibility loci. Three types of evidence are available: linkage statistics from published studies and from two recent meta-analyses of genomewide linkage scans (Badner and Gershon 2002; Lewis et al. 2003), evidence from a chromosomal deletion syndrome on chromosome 22q (Lindsay et al. 1995), and reported associations with positional susceptibility genes, some of which have gathered independent support (Owen et al. 2004) and convergent support from independent neuropharmacological findings (Cloninger 2002). We anticipate that many SZ susceptibility genes remain unidentified and that some of these can be characterized using genetic linkage and association methods. Therefore, we have recruited a large sample of affected sibling pairs (ASPs) to search for susceptibility loci by whole-genome linkage screening, to prioritize regions for more-intensive study. The families were ascertained through probands with SZ who have a sibling with SZ or schizoaffective disorder (SA). The sample is part of the National Institute of Mental Health (NIMH) Genetics Initiative for Schizophrenia, and the NIMH will make available the DNA specimens and blinded clinical data (see NIMH Center for Collaborative Genetics Studies of Mental Disorders Web site).

We report here the results of a genome scan of 409 pedigrees with SZ—263 of European ancestry (EA) and 146 of African American ancestry (AA), all collected by the Molecular Genetics of Schizophrenia (MGS1) Collaboration—with the use of a 9-cM map of STRPs. Using nonparametric multipoint linkage analysis, we detected two chromosomal regions with suggestive evidence of linkage (Lander and Kruglyak 1995) on chromosomes 8p23.3-p12 and 11p11.2-q22.3 in the full sample and, in exploratory analyses, two regions that reached similar thresholds on chromosomes 4p16.1-p15.32 and 5p14.3-

q11.2 in the AA and EA samples, respectively. Chromosome 8p is one of the most consistently observed regions in SZ genome scans, as discussed below. There have been several reports of significant association between SZ and polymorphisms in *neuregulin 1* (*NRG1*) in this region (Craddock et al. 2005), but it is not, in fact, clear to what extent this gene accounts for the linkage signals, and replication has not been uniform (Duan et al. 2005). Most such linkage signals have been observed closer to the p-telomere than to *NRG1*. In the present sample, there is evidence of linkage across a broad region, and the two peaks observed primarily in EA families are both telomeric to *NRG1*. Also within this region is a paracentric inversion that is common in Europeans (Broman et al. 2003). A large block of linkage disequilibrium (LD) was observed in this region in our EA sample, presumably because of the absence of crossing-over when a parent is heterozygous for the inversion. The inversion also creates uncertainty about the location and order of certain markers. Previous SZ linkage reports on this region have not specifically addressed the effects of this inversion on the results. We discuss below our efforts to accommodate this inversion.

## Material and Methods

### Family Ascertainment

The study participants were enrolled at nine sites in the United States and one in Australia: University of Chicago, Chicago; University of California, Irvine; University of Colorado Health Sciences Center, Denver; Baylor College of Medicine, Houston; University of Iowa, Iowa City; Washington University, St. Louis; Mount Sinai School of Medicine, New York; University of Pennsylvania, Philadelphia (this site included a subcontract with the Louisiana State University Health Sciences Center, New Orleans); and the University of Queensland, Brisbane. The focus of recruitment efforts was on EA and AA families. EA populations in the United States and Australia have similar ethnic characteristics (Cavalli-Sforza et al. 1994). Families were identified from a variety of sources, including local treatment facilities, physician referrals, the National Alliance for the Mentally Ill and other advocacy groups, Web sites, media announcements, and advertisements. Participants gave informed consent to an interview, provided a blood specimen for DNA and cell lines, and granted permission to obtain their psychiatric records and (usually) to contact a family informant. Available parents were asked to provide a blood specimen. If one or both parents were not available, as many as two unaffected sibs were asked to provide blood specimens. Local institutional review board approval was obtained for each site.

### Clinical Assessment

Subjects (aged  $\geq 18$  years) were interviewed by trained clinicians using the semistructured Diagnostic Interview for Genetic Studies (DIGS) 2.0 (Nurnberger et al. 1994), to elicit informa-

tion required to determine diagnoses of psychotic, mood, and substance-use disorders in accordance with DSM-IV criteria (American Psychiatric Association 1994), the comorbidity of these disorders, medical history, and ratings of positive and negative symptoms of SZ. A family informant was also interviewed, when possible, about each patient's history and about the family psychiatric history, by use of the Family Interview for Genetic Studies (FIGS) (Gershon et al. 1988; Maxwell 1992). Two experienced research clinicians independently reviewed the DIGS, FIGS, interviewer's narrative report, and all available psychiatric records and then assigned all relevant diagnoses (with a confidence level and age at onset identified for each one), a judgment about presence or absence of each DSM-IV criterion for SZ and SA, and the estimated lifetime duration of SZ and of mood syndromes. For all these items, the diagnosticians resolved any disagreements by discussion (or occasionally asked a third diagnostician to serve as a tie-breaker). Thus, a primary best-estimate final diagnosis (BEFD) (Leckman et al. 1982) was assigned on the basis of this consensus procedure. All individuals (interviewed or not) who did not receive an SZ or SA diagnosis were considered to have "diagnosis unknown" for the linkage analyses. One of the two diagnosticians also completed the Lifetime Dimensions of Psychosis Scale (LDPS), a 21-item scale for rating the lifetime duration and severity of positive, bizarre positive, negative, disorganized, and mood symptoms of psychotic disorders, on the basis of all available information for each patient (Levinson et al. 2002). Analyses using LDPS ratings will be presented elsewhere.

#### *Inclusion and Exclusion Criteria*

To be eligible, a family was required to be multiplex—minimally, to have a proband with a BEFD of SZ and one or more siblings with SZ or SA, with a confidence level of "likely" or "definite." Additional first-degree relatives with SZ and SA were recruited when possible. Both SZ-affected and SA-affected relatives have been included in this and most other SZ linkage studies, on the basis of evidence from multiple-family studies showing that both disorders cluster in families ascertained through a proband with SZ (Gershon et al. 1988; Kendler et al. 1993a; Maier et al. 1993; Taylor et al. 1993). For an SA diagnosis, DSM-IV requires that criteria for SZ and for manic, mixed, and/or major depressive episodes be met simultaneously at some point, with persistence of psychotic symptoms without prominent mood symptoms for at least 2 wk, and with mood episodes persisting for "a substantial portion of the total duration" (Cloninger et al. 1998, p. 278) of illness, defined here as 30% for consistency with the first NIMH Genetics Initiative for Schizophrenia study. Subjects were excluded from receiving a final diagnosis of SZ or SA if psychosis was limited to periods of likely substance intoxication or withdrawal, if persistent psychotic symptoms were considered likely to be related to substance use (e.g., increasing paranoia after years of amphetamine use or symptoms limited to visual hallucinations and "flashbacks" after hallucinogen use), if psychosis might have been caused by another disorder (e.g., epilepsy predating SZ onset) as determined by consensus judgment, or if the individual had moderate or severe mental retardation.

#### *Diagnostic Reliability and Supplemental Clinical Measures*

The cross-site reliability of diagnoses was measured by comparing the original site's consensus BEFD with a new consensus BEFD produced blindly by a second site using blinded copies of all available case material for 32 cases from the first NIMH Genetics Initiative for Schizophrenia from Washington University (Cloninger et al. 1998) and 36 cases from the present study (total  $N = 68$ ). The 68 total cases included 40 with a diagnosis of SZ from the original site, 10 with an SA diagnosis, and 18 with mood, personality, and/or substance-use diagnoses, most with at least some symptoms suggestive of an SZ spectrum disorder. Cohen's kappas (Cohen 1960) were 0.88 for SZ and 0.89 for SA, indicating no notable variation across sites.

#### *STRP Genotyping*

DNA was extracted from lymphoblastoid cell lines by the Rutgers University Cell and DNA Repository and was delivered in two "waves" to the Center for Inherited Disease Research (CIDR) for genotyping. CIDR's methods and details of the CIDR map are available online (see CIDR Web site). The marker set consisted of ~400 STRPs, with mean spacing of 9 cM (the largest gap was 18 cM) and a mean marker heterozygosity of 0.76. Family relationships were checked prior to the scan with a forensic panel of 12 markers. Samples with failed or very low amplification and contaminated samples were identified, and sex checks and a review of Mendelian discrepancies were performed when required. Failed wave I specimens were replaced when possible, and the entire family was re-genotyped (for 23 families) in wave II. Each gel included four CEPH controls, blind duplicates, and blank PCR products. Over 5% of wave I specimens were genotyped a second time in wave II. Analyses with RELCHECK (Boehnke and Cox 1997; Broman and Weber 1998) to assess family relationships were performed with the full set of genotypes. Genotypes were removed in cases of Mendelian inconsistencies or unlikely genotypes ( $P < .01$ ), as determined by MERLIN (multipoint engine for rapid likelihood inferece) (Abecasis et al. 2002). Quality control analyses for waves I and II, respectively, showed missing data rates of 3.96% and 3.48%, genotypewise error rates of 0.1% and 0.03% based on blind duplicates, and Mendelian inconsistency rates of 0.72% and 0.67%.

#### *Analysis of Population Substructure*

Accurate estimation of marker-allele frequencies is critical when DNA is unavailable for some parents and when population substructure exists (Curtis and Sham 1996), and allele frequencies can be significantly different for EA and AA subjects (Goddard et al. 2000; Grigull et al. 2001). It was decided that separate allele-frequency estimates for these two ethnic groups would be established for all analyses. Therefore, we used the program STRUCTURE (Pritchard et al. 2000) to identify and exclude families that were outliers with respect to the two main ethnic clusters, because these families would otherwise have been analyzed with inappropriate frequency estimates. Self-reported ancestries were EA or AA, plus 10 families with Hispanic ancestry. STRUCTURE analyses were completed assuming  $K = 2, 3$ , or 4 clusters by use of all autosomal

**Table 1**  
**Fine-Mapping Summary**

VALUE	CHROMOSOME									TOTAL AVERAGE										
	4	5	8	10p	10q	11	EA+AA	EA+AA	EA+AA											
STRP $Z_{lr}$ peak group	AA	EA	EA+AA	EA	EA+AA	EA+AA	EA	EA+AA	EA+AA	EA+AA	EA+AA	EA+AA	EA+AA	EA+AA	EA+AA	EA+AA	EA+AA	EA+AA	EA+AA	EA+AA
Peak STRP $Z_{lr}$	2.72	2.77	3.25	2.04	2.23	2.74	2.04	2.23	2.74	2.74	2.74	2.74	2.74	2.74	2.74	2.74	2.74	2.74	2.74	2.74
Approximate 1-LOD linkage interval of STRP scan (cM) <sup>a</sup> :																				
Start	0	33.6	2.1	0	120.3	61.9	0	120.3	61.9	61.9	61.9	61.9	61.9	61.9	61.9	61.9	61.9	61.9	61.9	61.9
End	35.5	74.3	60.3	20.9	170.9	94.6	20.9	170.9	94.6	94.6	94.6	94.6	94.6	94.6	94.6	94.6	94.6	94.6	94.6	94.6
Total approximate 1-LOD interval (cM) <sup>a</sup>																				
Fine-mapping region chosen (cM) <sup>b</sup> :																				
Start	.0	24.9	.0	.0	134.7	43.2	.0	134.7	43.2	43.2	43.2	43.2	43.2	43.2	43.2	43.2	43.2	43.2	43.2	43.2
End	43.6	84.0	77.9	20.0	170.9	110.1	20.0	170.9	110.1	110.1	110.1	110.1	110.1	110.1	110.1	110.1	110.1	110.1	110.1	110.1
Total distance fine mapped (cM) <sup>b</sup>	43.6	59.1	77.9	20.0	36.2	67.0	20.0	36.2	67.0	67.0	67.0	67.0	67.0	67.0	67.0	67.0	67.0	67.0	67.0	67.0
Cytogenetic position <sup>c</sup>	4p16.3-4p15.2	5p15.2-5q13.3	8p23.3-8q12.1	10p15.3-10p14	10q25.3-10q26.3	11p13-11q23.3														
Physical distance (Mb) <sup>d</sup> :																				
Start	.3	14.0	.6	.3	116.4	34.9	.3	116.4	34.9	34.9	34.9	34.9	34.9	34.9	34.9	34.9	34.9	34.9	34.9	34.9
End	27.2	75.8	59.5	6.7	134.0	116.6	6.7	134.0	116.6	116.6	116.6	116.6	116.6	116.6	116.6	116.6	116.6	116.6	116.6	116.6
Total distance fine mapped (Mb)	26.8	61.8	58.9	6.4	17.7	81.8	6.4	17.7	81.8	81.8	81.8	81.8	81.8	81.8	81.8	81.8	81.8	81.8	81.8	81.8
No. of genotyped SNPs	91	120	131	33	71	139	33	71	139	139	139	139	139	139	139	139	139	139	139	139
Initial SNP interval (cM) <sup>b</sup>	.48	.49	.59	.61	.51	.48	.61	.51	.48	.48	.48	.48	.48	.48	.48	.48	.48	.48	.48	.48
No. of SNPs failed for assay	4	4	4	1	3	3	1	3	3	3	3	3	3	3	3	3	3	3	3	3
No. of SNPs failed for low (<90%) genotyping rates	6	5	5	3	0	4	3	0	4	4	4	4	4	4	4	4	4	4	4	4
No. of SNPs with excessive Mendelian errors	7	0	1	0	2	1	0	2	1	1	1	1	1	1	1	1	1	1	1	1
No. of SNPs remaining with clean genotypes	74	111	121	29	66	131	29	66	66	66	66	66	66	66	66	66	66	66	66	66
Cleaned SNP interval (cM) <sup>b</sup>	.59	.53	.64	.69	.55	.51	.69	.55	.55	.55	.55	.55	.55	.55	.55	.55	.55	.55	.55	.55
Average genotyping rate of cleaned SNPs (%)	.987	.988	.980	.988	.988	.988	.988	.988	.988	.988	.988	.988	.988	.988	.988	.988	.988	.988	.988	.988
Average genotyping rate (per genotype) of cleaned SNPs (%)	.00015	.00023	.00012	.00012	.00017	.00018	.00012	.00017	.00017	.00018	.00018	.00018	.00018	.00018	.00018	.00018	.00018	.00018	.00018	.00018
Mendelian error rate (per genotype) of cleaned SNPs (%)	.00048	.00030	.00042	.00069	.00075	.00033	.00069	.00075	.00075	.00033	.00033	.00033	.00033	.00033	.00033	.00033	.00033	.00033	.00033	.00033
Unlikely-recombinants rate (per genotype) of cleaned SNPs (%)																				
Average MAF of cleaned SNPs:																				
EA	.37	.35	.36	.34	.35	.38	.34	.35	.35	.38	.38	.38	.38	.38	.38	.38	.38	.38	.38	.38
AA	.35	.34	.35	.37	.34	.35	.37	.34	.34	.35	.35	.35	.35	.35	.35	.35	.35	.35	.35	.35
No. of clean SNPs removed because of deviation from HWE:																				
EA	1	2	0	0	1	2	0	1	1	1	2	1	1	1	2	1	1	1	1	1
AA	0	1	3	1	0	1	1	0	0	0	1	1	0	0	1	0	0	0	0	0

No. of clean SNPs removed because of adjacent SNP LD:										
EA	9	21	12	6	11	13	72	...	...	...
AA	6	5	9	3	7	12	42	...	...	...
No. of clean SNPs removed because of inversion (only chromosome 8):										
EA	0	0	5	0	0	0	5	...	...	...
AA	0	0	5	0	0	0	5	...	...	...
No. of final SNPs analyzed for linkage:										
EA	64	88	104	23	54	116	449	...	...	...
AA	68	105	104	25	59	118	479	...	...	...
Final SNP interval (cM) <sup>b</sup> :										
EA	.68	.67	.75	.87	.67	.58	...	...	.68	...
AA	.64	.56	.75	.80	.61	.57	...	...	.63	...
Final fine-map IC:										
EA	.83	.85	.85	.80	.83	.85	...	...	.84	...
AA	.77	.78	.79	.77	.78	.78	...	...	.76	...
Fine-mapping $Z_{it}$ peak group										
Peak fine-mapping $Z_{it}$	AA	EA	EA	EA	EA	EA+AA	...	...	...	...
SNP nearest to peak fine-mapping $Z_{it}$	3.09	3.80	3.33	2.52	2.08	3.08	...	...	...	...
Approximate 1-LOD interval after fine mapping (cM) <sup>b</sup> :	<i>rs7681266</i>	<i>rs1027164</i>	<i>rs7834209</i>	<i>rs1155931</i>	<i>rs1343418</i>	<i>rs4275647</i>	...	...	...	...
Start	1.0	44.9	5.1	.0	134.1	69.7	...	...	...	...
End	32.0	76.7	46.0	23.2	153.5	89.2	...	...	...	...
Total approximate 1-LOD interval (cM) <sup>b</sup>	31.0	31.8	40.9	23.2	19.4	19.5	166	...	...	...
Total narrowing of fine-mapping peak achieved (cM) <sup>b</sup>	4.5	8.9	17.3	-2.3	31.2	13.2	73	...	...	...
Cytogenetic position of approximate 1-LOD interval after fine mapping <sup>c</sup>	4p16.3-4p15.32	5p14.1-5q12.1	8p23.2-8p21.2	10p15.3-10p14	10q25.3-10q26.13	11q13.1-11q14.1	...	...	...	...
Physical distance of approximate 1-LOD interval after fine mapping (Mb) <sup>c</sup> :	...	...	...	...	...	...	...	...	...	...
Start	.00	25.22	2.54	.00	115.78	63.69	...	...	...	...
End	16.65	61.62	25.93	9.27	126.53	82.59	...	...	...	...
Total approximate 1-LOD interval (Mb) <sup>c</sup>	16.7	36.4	23.4	9.3	10.8	18.9	115	...	...	...

<sup>a</sup> cM values refer to the STRP map of CIDR (used for linkage scan).

<sup>b</sup> cM values refer to the deCODE map (used for fine mapping).

<sup>c</sup> Cytogenetic locations and Mb values are from the UCSC July 2003 freeze.

**Table 2**

**Counts of Families and Individuals Genotyped**

SAMPLE AND NO. OF AFSPs PER FAMILY	AFSPs <sup>a</sup>	AHAPs <sup>b</sup>	Affected Nonsibs <sup>c</sup>	Unaffected Sibs <sup>d</sup>	NO. OF GENOTYPED PARENTS				
					0	1	2	3	4
EA sample:									
0 AFSPs	0	15	2	5	5	8	1	0	0
1 AFSPs	226	0	6	77	81	75	70	0	0
2 AFSPs	18	0	1	5	7	7	4	0	0
3 AFSPs	4	0	1	9	2	0	1	0	1
4 AFSPs	0	0	0	0	0	0	0	0	0
5 AFSPs	1	0	0	0	1	0	0	0	0
No. of families	249	NA	NA	NA	96	90	76	0	1
No. of individuals or pairs	279	15	10	96	0	90	152	0	4
AA sample:									
0 AFSPs	0	53	7	15	27	20	0	0	0
1 AFSPs	80	17	8	29	57	18	5	0	0
2 AFSPs	15	11	1	9	6	7	1	0	1
3 AFSPs	2	4	1	2	0	2	0	0	0
4 AFSPs	2	0	0	5	1	1	0	0	0
5 AFSPs	0	0	0	0	0	0	0	0	0
No. of families	99	NA	NA	NA	91	48	6	0	1
No. of individuals or pairs	124	85	16	60	0	48	12	0	4
Total sample:									
No. of families	348	NA	NA	NA	187	138	82	0	2
No. of individuals or pairs	403	100	26	156	0	138	164	0	8

NOTE.—All counts reflect genotyped individuals only. A total of 1,380 individuals were genotyped in 409 families. The number of all (genotyped or not) parents was 920 (552 EA and 368 AA). NA = not applicable.

<sup>a</sup> Genotyped independent AFSPs ( $n - 1$ ).

<sup>b</sup> Genotyped AHAPs, all possible pairs; see main text for description of this counting method.

<sup>c</sup> Genotyped affected nonsiblings; this includes persons who are not siblings or half siblings of affected individuals, so it includes parents, aunts, uncles, and offspring.

<sup>d</sup> Genotyped unaffected siblings (and also unaffected half siblings) of affected individuals.

STRP genotypes. Each run consisted of 10,000 burn-ins and 10,000 subsequent iterations. There were clear EA and AA clusters regardless of the value of  $K$ , with no additional clusters that reflected self-reported ethnicity. Eight of the 10 Hispanic families consistently clustered (posterior  $P > .85$ ) with either the EA or AA group (or showed approximately equal EA or AA admixture). Thus,  $K = 2$  was assumed, and 10 runs of  $K = 2$  with different starting seeds produced no changes in family assignments to the clusters. Thirteen families with 1 or 2 ungenotyped parents were removed from linkage analyses because they did not attain the selected threshold of a posterior probability of group membership of at least 0.85. Two families with self-reported Asian ancestry (without genotyped parents) were also removed.

#### Confirmation of Biological Relationships

All participant-reported genetic relationships were verified via RELCHECK (Boehnke and Cox 1997; Broman and Weber 1998) for autosomal and X-linked STRPs from the forensic tests, and Mendelian consistency was verified via PEDCHECK (O'Connell and Weeks 1998) for autosomal and X-linked STRPs from the genome scan, respectively. On the basis of these checks, 26 families were excluded because they no longer contained an ASP (full or half) or because the "siblings" proved to be not related, and 6 families were excluded because the siblings had identical genotypes (because of either an unre-

ported MZ twin relationship or an error during phlebotomy, blood submission, or processing). For the retained families, two nuclear pedigrees contained self-reported full-sibling ASPs that were proven to be half-sibling ASPs in the EA sample; 18 nuclear pedigrees contained self-reported full-sibling ASPs that were proven to be half-sibling ASPs in the AA sample. Finally, one large pedigree was split into subfamilies so that they could be analyzed by GENEHUNTER-PLUS (GH+) (Kong and Cox 1997), resulting in the creation of a "new" family; the subfamilies were treated as if they were unrelated to one another.

#### STRP Maps and Allele Renumbering

The STRP genotyping was performed in two waves. Briefly, in addition to three Y-linked markers and one pseudoautosomal marker not used in linkage analysis, 393 autosomal and X-linked STRPs were typed in wave I, and 400 in wave II. Genotypes for the 387 markers common to both waves were hand-checked for consistency of allele designations, and alleles were renumbered when necessary, including 114 for which renumbering was complex (see tables A1 and A2 in appendix A for detailed examples). For the genome-scan analysis, we assumed map positions provided by CIDR, which are similar to Marshfield map positions (Broman et al. 1998). For the linkage fine-mapping analysis, all positions were drawn or interpolated from the deCODE map (Kong et al. 2002), as discussed below.

**Table 3**

**Clinical Characteristics of All Genotyped Individuals**

CHARACTERISTIC	PROBANDS		SIBLINGS		FATHERS		MOTHERS		OTHERS <sup>a</sup>	
	EA	AA	EA	AA	EA	AA	EA	AA	EA	AA
No. of individuals	263	145	368	193	92	10	148	48	38	78
Percentage male	68.8	53.4	57.7	48.2	100	100	0	0	60.5	44.3
Mean (SD) age at enrollment (years)	39.1 (10.8)	41.4 (10.5)	40.6 (10.8)	43.3 (10.4)	65.5 (10.2)	69.6 (11.1)	63.3 (11.1)	60.8 (12.9)	33.9 (17.5)	38.5 (12.3)
Percentage with SZ	100	100	62.7	63.2	4.4	10.0	2.7	18.8	50.0	64.6
Percentage with SA	0	0	12.1	10.4	0	0	.7	2.1	13.2	6.3
Percentage with substance dependence:										
Any:										
Males	24.6	42.3	24.7	31.2	4.3	20.0	...	...	17.4	34.3
Females	16.9	23.5	15.3	20.0	...	...	4.1	8.3	13.3	25.0
Alcohol	15.8	26.7	12.4	15.5	4.3	20.0	2.7	4.2	5.3	16.5
Cannabis	7.5	8.2	7.0	7.3	0	0	0	2.1	5.3	6.3
Cocaine	2.3	12.3	1.6	10.9	0	0	0	0	0	8.8
Amphetamine	2.6	1.4	1.9	1.0	0	0	1.0	0	2.6	0

<sup>a</sup> The “Others” group consists of half siblings and other non–first-degree relatives.

*STRP Allele-Frequency Estimation*

Maximum-likelihood allele-frequency estimates were obtained from the USERM13 subroutine of MENDEL (Lange et al. 1988; Boehnke 1991) separately for the EA and AA families. By likelihood-ratio testing, EA and AA allele frequencies differed at  $P < 10^{-16}$  for 82.5% of the markers, at  $10^{-16} \leq P \leq .01$  for 17% of the markers, and at  $P > .01$  for only 0.5% (three) of the markers. Accordingly, for the linkage analysis of the combined EA and AA sample, we used separate allele-frequency estimates. For instance, if there were  $N$  alleles in the EA sample, we renumbered the AA alleles, starting with  $N + 1$ . The sum of the EA and AA allele frequencies totaled 200% (table A2).

*Linkage Analyses of STRPs and Determination of Empirical Significance Thresholds*

Linkage analyses were performed using GH+ (Kong and Cox 1997) under the exponential-model option and the  $S_{ALL}$  scoring function, to compute  $Z$  likelihood-ratio ( $Z_{lr}$ ) scores for each map position.  $Z_{lr}$  follows a normal distribution asymptotically under the null hypothesis of no linkage. All results are presented on the  $Z_{lr}$  scale. The correspondence between  $Z_{lr}$  and the  $\log_{10}$  likelihood ratio of linkage is  $LOD = (Z_{lr})^2 / 2 \ln(10)$ . Information content (IC) was computed by GH+. The algorithm computes entropy for the probability distribution for the inheritance vectors of all pedigrees with or without genotypic data. IC is then computed by subtracting the ratio of these terms from 1. Map locations were drawn from the CIDR map (Broman et al. 1998) or were interpolated into that map.

The linkage analysis of all families combined was considered to be the primary analysis. Exploratory analyses of the EA and AA families separately were also performed. To determine empirical significance thresholds for  $Z_{lr}$  scores, 5,000 replicates of the actual sample, map, and allele frequencies (including patterns of missing data) were generated using SIMULATE (Terwilliger et al. 1993) under the assumption of an absence of linkage. Each replicate was analyzed separately for all families (the planned primary analysis) and then for the AA and EA families separately. For the primary analysis, the  $Z_{lr}$  thresh-

old was 3.60 for “significant” linkage (expected in  $\leq 5\%$  of genome scans) and 2.65 for “suggestive” linkage (expected  $\leq 1$  times per genome scan) (Holmans et al. 2004). In comparison, the thresholds for EA and AA samples (without correction for multiple testing) were similar: 3.62 and 3.63, respectively, for significant linkage and 2.71 and 2.69, respectively, for suggestive linkage.

Also, as discussed below, a common paracentric inversion has been reported (Broman et al. 2003) in the region of chromosome 8p that produced the greatest evidence of linkage in the present study. One STRP in the CIDR map (*D8S1469*) is located in the middle of the typical inversion segment, and another STRP in the CIDR map (*D8S1130*) is close to the boundary of the typical inversion segment. The exact positions of these two STRPs, and the probability of recombination around them, would differ in individuals with and individuals without the inversion. The primary genome scan analysis results reported below excluded *D8S1469*, the less informative of the two inversion STRPs.

*SNP Genotyping and Data Cleaning*

For linkage fine mapping, genotyping of SNP markers was performed at Evanston Northwestern Healthcare’s Center for Psychiatric Genetics (Evanston, IL) by use of SNPlex (Applied Biosystems [ABI]). Using SNPbrowser software, version 1.0 (ABI), 585 SNPs (13 SNPlex pools) were selected to form a 0.5–0.6-cM map across 304 cM (255 Mb [Matisse et al. 2003]) of chromosomes 4p16.3-p15.2 (44 cM; 27 Mb), 5p15.2-q13.3 (59 cM; 62 Mb), 8p23.3-q12.1 (78 cM; 59 Mb), 10p15.3-p14 (20 cM; 7 Mb), 10q25.3-q26.3 (36 cM; 18 Mb), and 11p13-q23.3 (67 cM; 82 Mb) from HapMap, Celera (Venter et al. 2001), or public databases (see National Center for Biotechnology Information [NCBI] dbSNP Web site, build 34) (table 1). Validated SNPs were selected if their minor-allele frequency (MAF) was reported to be  $>25\%$  in databases for both EA and AA samples, and they were tested using bioinformatics for suitability for the SNPlex assay. Forty nanograms of fragmented DNA was dried down on each well of a 384-well plate. After phosphorylation of oligonucleotide ligation assay probes

**Table 4**

**Additional Clinical Characteristics of Genotyped Affected Subjects**

AFFECTED GROUP AND CHARACTERISTIC	PROBANDS				SIBLINGS			
	EA		AA		EA		AA	
	Males	Females	Males	Females	Males	Females	Males	Females
<b>SZ:</b>								
No. of individuals	182	81	78	67	155	77	66	54
Mean age at enrollment (years)	37.7	42.2	40.0	43.1	38.2	42.1	43.2	42.5
Mean age at illness onset (years)	20.3	20.3	19.6	22.2	20.1	21.1	19.3	20.3
Mean duration of nonaffective psychosis (mo)	207.5	258.9	238.1	253.9	215.4	246.6	276.4	253.9
Percentage of subjects with mood episodes	46.4	64.6	38.2	61.8	51.3	64.9	42.4	69.1
Mean duration of mood episodes (mo)	11.6	23.5	13.1	13.1	13.2	13.1	14.0	16.8
<b>SA:</b>								
No. of individuals					24	21	8	12
Mean age at enrollment (years)					40.0	41.8	35.4	41.8
Mean age at illness onset (years)					21.1	21.9	18.1	19.3
Mean duration of nonaffective psychosis (mo)					218.4	234.3	182.6	222.4
Mean duration of mood episodes <sup>a</sup> (mo)					94.8	123.6	89.3	115.7

<sup>a</sup> All subjects with SA had one or more mood episodes (manic, mixed, and/or major depressive), by definition.

and universal linkers, allele-specific ligation and enzymatic clean-up were performed. PCR was performed with universal primers, and biotinylated amplicons were captured on streptavidin-coated plates. Single-stranded PCR products were hybridized with a set of fluorescently labeled, universal ZipChute probes that have a unique sequence corresponding to each SNP. ZipChute probes were eluted and separated on a 3730 DNA Analyzer (ABI), and genotypes were called by GeneMapper 3.5 (ABI), blind to diagnosis.

Mendelian errors and unlikely genotypes were evaluated using MERLIN (Abecasis et al. 2002), as described above. We first excluded 19 SNPs because of failed assays and 23 SNPs because of low call rates (<90%). Then, we excluded 11 SNPs because of excessive Mendelian errors (>1%). The remaining 532 SNPs had 0.017% Mendelian errors per SNP (range 0.00%–0.66%) and 0.044% unlikely recombinants (range 0.00%–0.58%) for all fine-mapping regions (table 1). Two families were excluded—one because of excessive Mendelian errors (due to a specimen swap) and one without genotyped parents, because of excessive unlikely recombinants. The final call rate for all 532 cleaned SNPs was 98.6% (range 92.1%–100.0%), with a total of 729,158 cleaned genotypes remaining for further analysis. The average interval between neighboring SNPs was 0.59 Mb for chromosome 4, 0.53 Mb for chromosome 5, 0.64 Mb for chromosome 8, 0.69 Mb for chromosome 10p, 0.55 Mb for chromosome 10q, and 0.51 Mb for chromosome 11. There were several sizable centromeric gaps where the genome sequence information was incomplete (a 3.9-Mb gap for chromosome 5, a 3.8-Mb gap for chromosome 8, and a 4.1-Mb gap for chromosome 11). The average MAF was 0.35 (range 0.04–0.50) for AA subjects and 0.36 (range 0.00–0.50) for EA subjects. We also evaluated the repeatability with 121 SNPs from the chromosome 8 fine mapping, in which we observed repeatability of 99.9% when 268 DNA samples were blindly re-genotyped. Of 31,931 genotypes that were nonzero in both experiments, 31 were discrepant, with 3 changes involving both alleles in a homozygous subject and 28 others

involving a change of only one of the two alleles in a heterozygous individual.

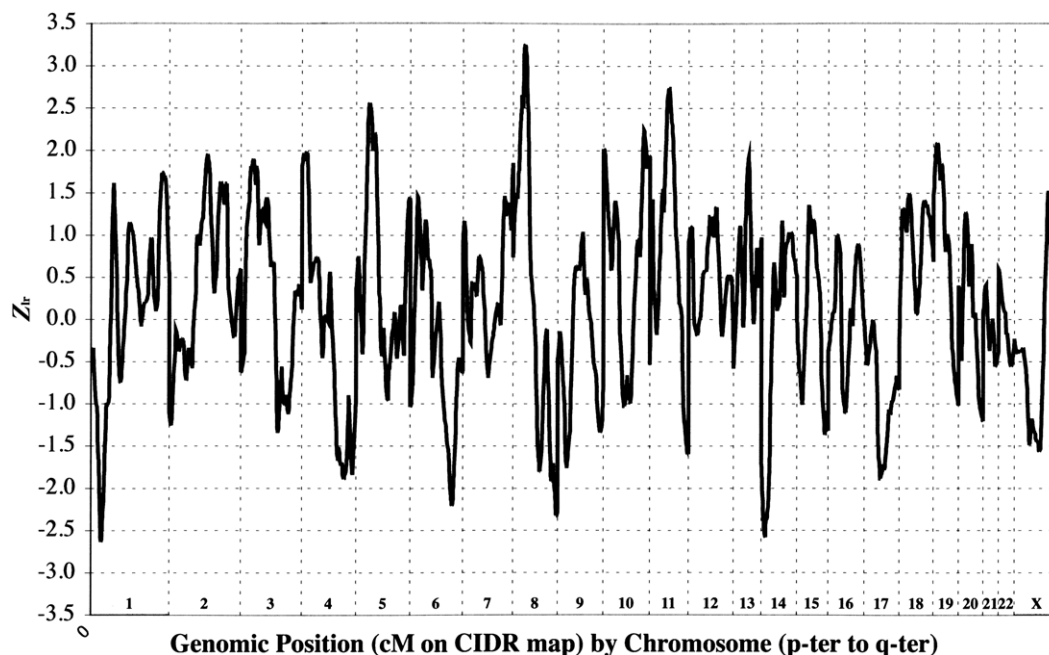
Deviations from Hardy-Weinberg equilibrium (HWE) were analyzed separately in EA and AA samples in a set of as many unrelated individuals as possible. Six SNPs were removed because of significant deviations at  $P < .01$  in the EA sample, and 6 SNPs were likewise removed in the AA sample, including 2 SNPs with highly significant deviations in the AA sample ( $P = .00006$  for *rs1355305*;  $P = .0002$  for *rs896044*) (table 1).

*Selection of a Map for Linkage Fine Mapping*

We therefore undertook to select a subset of markers for fine-mapping analyses that had the highest heterozygosity (and hence the highest IC). Pairwise marker LD between the cleaned fine-mapping SNPs was analyzed with a set of unrelated individuals, separately for EA and AA samples, by use of ASSOCIATE (Ott 1985), which computes maximum-likelihood estimates of the LD parameter by the expectation-maximization algorithm (Dempster et al. 1977). For example, of 121 SNPs in the chromosome 8 fine-mapping region, significant ( $P < .05$ ) pairwise LD was observed between 40 and 20 adjacent marker pairs in the EA and AA samples, respectively, which is consistent with other reports of greater LD in EA than in AA samples (Hinds et al. 2005).

A contiguous block of highly significant LD was observed in the EA sample across 3.31 Mb (~6 cM), with six adjacent SNPs extending from *rs2980438* (at 8.132 Mb) to *rs7824640* (at 11.442 Mb) on chromosome 8. It appears likely that this LD block is due to the large polymorphic inversion that is observed with a frequency >20% in the European (Giglio et al. 2001; Broman et al. 2003) and Japanese (Sugawara et al. 2003) populations. The inversion typically spans 4.7 Mb between two low-copy-repeat regions, each of which contains several olfactory receptor genes (Giglio et al. 2001). Direct demonstration of the inversion by use of FISH is beyond the scope of the present study, but we inferred that inversions could be present in some unknown but substantial proportion of our





**Figure 1** Results of the STRP genome scan for SZ.  $Z_{tr}$  scores for the entire sample are plotted on the Y-axis, and genomic position (cM on CIDR map) by chromosome (p-ter to q-ter) is plotted on the X-axis.

EA subjects on the basis of the large LD block and the fact that recombination cannot occur between chromosomes that are heterozygous for an inversion, leading to LD if the inversion is common. We detected no significant LD in this region in AA subjects, which suggests that this inversion is infrequent in sub-Saharan Africa, similar to the 17q21.31 paracentric inversion reported recently (Stefansson et al. 2005).

In families with missing parental genotypes, linkage scores can be inflated if substantial LD is present between marker pairs (Huang et al. 2004; Dunn et al., in press). Recent work by one of us (Levinson and Holmans, in press) indicates, however, that there is no appreciable inflation of linkage scores if dense SNP maps are trimmed to eliminate marker pairs whose LD (as measured by  $r^2$ ) is  $>0.05$ . Accordingly, we used this criterion to trim our SNP maps separately for the EA and AA families by removing one of each adjacent pair of SNPs that showed this level of LD. Removal of a SNP will create a new adjacent pair, and these new pairs were also evaluated for LD, until all adjacent pairs had  $r^2$  estimates  $<0.05$ . For chromosome 8, a total of 17 SNPs were trimmed from both EA and AA families. In the EA families, 12 SNPs were removed because  $r^2$  was  $>0.05$ , and 5 were removed because they are located in the inversion (where there was strong LD). In the AA families, three SNPs were removed because they deviated from HWE (at  $P < .01$ ), nine were removed because  $r^2$  was  $>0.05$ , and five were removed because they are located in the inversion. We retained a single SNP (hCV1965865) from the inversion region. However, we deleted this SNP in analyses that included both SNPs and STRPs, since an STRP in the inversion (*D8S1130*) is more polymorphic. Similarly, we had already deleted the less-informative inversion STRP (*D8S1469*) for the combined SNP and STRP analysis, leaving a total of 104 SNPs

and 9 STRPs for the final chromosome 8p fine map. Fine-mapping linkage results are reported for the EA and AA families separately and combined.

For other fine-mapping regions, we followed similar rules to exclude SNPs that may be in LD with adjacent SNPs (table 1). The final numbers of SNPs used in fine-mapping analyses for the EA and AA samples, respectively, were 64 and 68 for chromosome 4, 88 and 105 for chromosome 5, 104 and 104 for chromosome 8, 23 and 25 for chromosome 10p, 54 and 59 for chromosome 10q, and 116 and 118 for chromosome 11, for a total of 449 and 479 cleaned fine-mapping SNPs without significant intermarker LD that proceeded to linkage analyses (table 1).

Correct specifications of marker-marker genetic distances are also critical for accurate multipoint linkage analysis. Genetic-mapping information was not available for most of our fine-mapping SNPs, and our sample did not include the kinds of complete and multigenerational pedigrees that ideally should be used to estimate genetic distances. The deCODE map includes genetic locations for a larger number of markers, on the basis of more meioses, than does the CIDR map. Therefore, we created a genetic and a physical map of the STRPs in the region, using the UniSTS database (see UniSTS Web site), deCODE map information for genetic locations, and the NCBI sequence map for physical positions. Genetic locations were assigned to SNPs by use of linear interpolation of physical positions in relation to the genetic locations of flanking STRPs.

#### *Sharing of Biomaterials and Clinical Data*

Biological materials, genotypes, and blinded clinical data will be made available to the scientific community by the NIMH

**Table 5****STRPs with a Multipoint  $Z_{lr}$  of Nominal Significance at  $P < .01$** 

CHROMOSOME AND MARKER	ALIAS	POSITION (cM) <sup>a</sup>	$Z_{lr}$		
			EA Sample	AA Sample	Total Sample
3:					
D3S2406	GGAT2G03	103	2.33	...	...
4:					
D4S2366	GATA22G05	13	...	2.66	...
D4S403	AFM157xg3	26	...	2.43	...
5:					
D5S2848	GATA145D09	40	2.65	...	...
D5S1470	GATA7C06	45	2.72	...	2.52
D5S2500	GATA67D03	69	2.37	...	...
6:					
D6S2427	GGAA15B08	54	...	2.54	...
8:					
D8S264	AFM143xd8	1	2.56	...	...
D8S262	AFM127xh2	7	3.33	...	...
D8S1130	GATA25C10	22	3.06	...	...
D8S1106	GATA23D06	26	2.72	...	...
D8S1145	GATA72C10	37	3.09	...	2.47
D8S560	AFMa127ye5	43	2.96	...	2.51
D8S1771	AFMb320va5	50	2.95	...	3.24
D8S1477	GGAA20C10	60	...	...	2.64
11:					
D11S2371	GATA90D07	76	2.66	...	2.74
D11S2002	GATA30G01	85	...	...	2.36
20:					
D20S851	AFMa218yb5	25	2.34	...	...

NOTE.—A nominal  $P < .01$  is a  $Z_{lr} > 2.326$  here.<sup>a</sup> cM values refer to the STRP map of CIDR.

Center for Collaborative Genetics Studies of Mental Disorders (see the Center's Web site).

## Results

We report here results of the clinical characteristics of the sample, a genomewide linkage scan of SZ, and the fine mapping of 4p16.3-p15.2, 5p15.2-q13.3, 8p23.3-q12.1, 10p15.3-p14, 10q25.3-q26.3, and 11p13-q23.3.

### *Families and Individuals Included in Linkage Analyses*

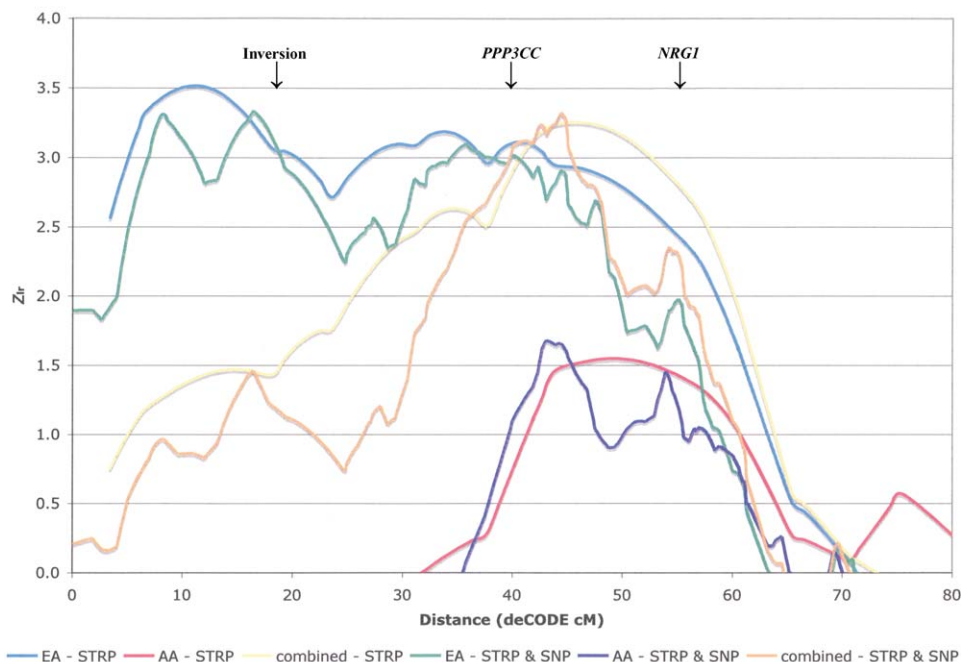
A total of 459 families met eligibility criteria and were included in the two waves of CIDR genotyping, although 4 of these families were removed because one of the siblings was no longer considered to be affected after further diagnostic review in preparation for dimensional analyses. After removal of families as described above, 409 families (408 independent families) were included in the linkage analysis: 263 (64.3%) EA and 146 (35.7%) AA.

**Table 6**

**STRPs by Map Position and Multipoint  $Z_{lr}$  for Linkage Scan**

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

The 409 families resulted from the splitting of 1 family that was too large to be analyzed by GH+. A full description of the DNA specimens available from this data set, including those excluded from the present analysis, will be made available on the NIMH Human Genetics Initiative for Schizophrenia Web site. Table 2 describes the number of statistically independent affected full-sib pairs (AFSPs) and the number of all possible affected half-sib pairs (AHAPs) used in the linkage analysis. Statistically independent AFSPs were counted as  $S - 1$  for  $S$  affected full sibs in a sibship with only full sibs; all possible AHAPs were counted as  $S(S - 1)/2$  for  $S$  affected half sibs in a sibship with only half sibs. Because there is no optimal method for counting independent AHAPs when there are full and half sibs in the same family, all possible half-sib pairs have been counted in these cases (see fig. A1 in appendix A for an illustration and explanation). Linkage analyses included 403 independent AFSPs (279 EA and 124 AA), 100 AHAPs (15 EA and 85 AA), and 26 other genotyped affected relatives (parents, aunts, uncles, and offspring). Genotypes were available for 33% of all parents (45% EA and 17% AA), and 156 unaffected full or half sibs were genotyped (96 EA and 60 AA). Linkage analyses included genotypes from 1,380 subjects.



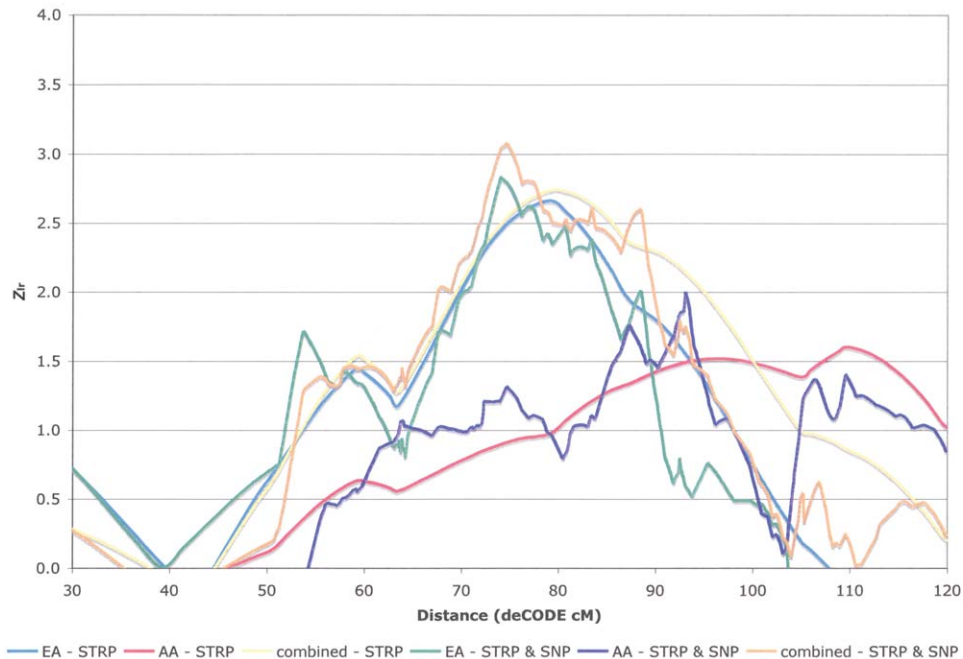
**Figure 2** Results of the STRP genome scan for SZ on chromosome 8 and the subsequent SNP fine mapping. Positive  $Z_{ir}$  scores for the STRP scan for the entire sample (yellow line), EA subsample (blue line), and AA subsample (red line) are plotted on the Y-axis, and genomic position (CIDR cM converted to deCODE cM from p-ter, to allow plotting of the fine-mapping data) on chromosome 8 is plotted on the X-axis (0–80 cM). Similarly, positive  $Z_{ir}$  scores for the scanning STRPs plus the fine-mapping SNPs for the entire sample (orange line), EA sample (green line), and AA sample (purple line) are plotted on the Y-axis, and genomic position (deCODE cM from p-ter) on chromosome 8 is plotted on the X-axis (0–80 cM). The *PPP3CC* gene is located at 39 deCODE cM, and the *NRG1* (GGF2 isoform) gene spans 53–54 deCODE cM. The common inversion on chromosome 8p spans ~18–22 deCODE cM and is marked.

Descriptions of the 1,380 individuals who were genotyped are given in tables 3 and 4. Of the 409 families, 263 were EA, and 146 (from 145 independent families) were AA. The mean age at enrollment was 39.1 and 41.4 years for the EA and AA probands, respectively, with onset at ~20 years of age for all groups of male and female probands, both EA and AA. As noted in table 3, 35% of fathers and 56% of mothers of the EA probands and 7% of fathers and 33% of mothers of the AA probands were genotyped. Overall, 4.9% of fathers and 6.6% of mothers received a diagnosis of SZ, and 1% of mothers received a diagnosis of SA. The “other” group consisted of 61 half sibs plus other non–first-degree relatives. In table 4, we display the characteristics of probands and siblings with SZ. The percentage of SZ-affected subjects with a mood disorder did not differ by ethnicity or between probands and relatives, but females had a higher percentage than males did (64.5% vs. 46.2%;  $\chi^2_1 = 24.8$ ;  $P < .0001$ ). All SA-affected subjects had mood episodes (manic, mixed, and/or major depressive), by definition. We used a general linear model to examine differences in age at enrollment, age at onset, duration of psychosis, and duration of mood episodes

in the subjects with SZ and those with SA. The independent variables were ethnicity, sex, and familial relationship (probands vs. siblings). Both sex and ethnicity were significant predictors of duration of nonaffective psychosis ( $P < .01$ ), with females and AA subjects having longer durations. No other variables were significant at the  $P < .01$  level. There was a trend toward females having a longer duration of mood episodes ( $P = .03$ ). Next, we compared the siblings with SZ and those with SA. The subjects with SA had a longer duration of mood episodes ( $P < .0001$ ). The only other significant variable was ethnicity as a predictor of the duration of nonaffective psychosis ( $P = .02$ ).

#### Genome Scan

For the STRP genome scan, the CIDR map was used, although, for purposes of display alongside the fine-mapping SNPs, we converted to the deCODE map. Figure 1 displays the multipoint  $Z_{ir}$  graph across the genome for all families, and table 5 lists the STRPs that attained a multipoint  $Z_{ir}$  with nominal significance at  $P < .01$  (results for all STRPs are found in table 6). We detected two



**Figure 3** Results of the STRP genome scan for SZ on chromosome 11 and the subsequent SNP fine mapping. Positive  $Z_{ir}$  scores for the STRP scan for the entire sample (yellow line), EA subsample (blue line), and AA subsample (red line) are plotted on the Y-axis, and genomic position (CIDR cM converted to deCODE cM from p-ter, to allow plotting of the fine-mapping data) on chromosome 11 is plotted on the X-axis (30–120 cM). Similarly, positive  $Z_{ir}$  scores for the scanning STRPs plus the fine-mapping SNPs for the entire sample (orange line), EA subsample (green line), and AA subsample (purple line) are plotted on the Y-axis, and genomic position (deCODE cM from p-ter) on chromosome 11 is plotted on the X-axis (30–120 cM).

chromosomal regions with suggestive evidence ( $Z_{ir} \geq 2.65$ ) of linkage in the primary analyses of all families: 8p23.3-p12 and 11p11.2-q22.3. The mean IC for the STRP scan was 0.60 for the EA sample and 0.55 for the AA sample.

On chromosome 8p23.3-p12 (fig. 2),  $Z_{ir}$  scores  $>2.0$  were observed across 31 cM, from 30.7 to 61.7 cM on the CIDR map (~26.6–58.5 cM on the deCODE map used for fine-mapping analyses; see below). The maximum multipoint evidence of linkage was a  $Z_{ir}$  score of 3.25 (equivalent Kong-Cox LOD of 2.30) near *D8S1771* (at 52 cM). Although the meaning of a 1-LOD interval is not clear for complex disorders, we note that here the 1-LOD interval (on the CIDR map) extended from 37.0 cM (18.36 Mb, at *D8S1145*) to 60.3 cM (32.12 Mb, at *D8S1477*). The linkage signal in this region was observed primarily in EA families, in which two peaks of similar magnitude were observed, at 11.5 cM ( $Z_{ir} = 3.52$ ) and at 39.6 cM ( $Z_{ir} = 3.19$ ), with a 1-LOD interval extending from 2.15 cM (~3.03 Mb, centromeric to *D8S264*) to 58.3 cM (~30.11 Mb, telomeric to *D8S1477*). In AA families, negative  $Z_{ir}$  scores were observed from 8pter to 37.0 cM, with a maximum  $Z_{ir}$  of 1.55 at 54.2 cM.

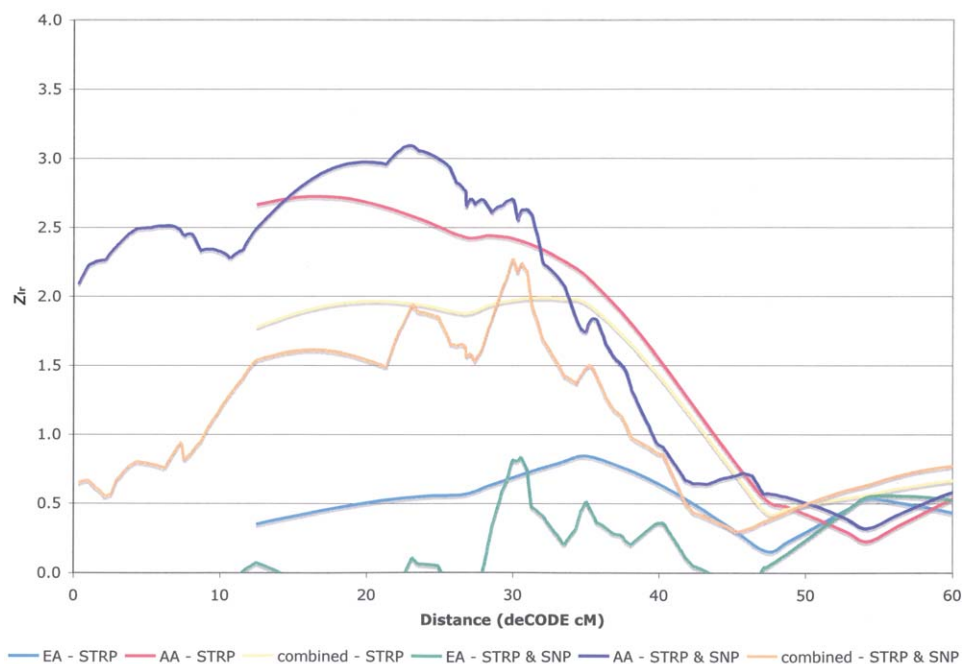
Suggestive linkage was also observed in the entire sample at 11p11.2-q22.3 (fig. 3), with a  $Z_{ir}$  of 2.74 at

*D11S2371* (76.1 cM; equivalent LOD of 1.63), with a 1-LOD interval from 61.9 cM (~60.55 Mb, centromeric to *D11S1344*) to 94.6 cM (~94.39 Mb, centromeric to *D11S2000*). Again, greater evidence of linkage was observed in EA families ( $Z_{ir} = 2.66$  at *D11S2371*, 76.1 cM) than in AA families ( $Z_{ir} = 1.61$  at *D11S1391*, 104.6 cM).

In the exploratory analysis, uncorrected suggestive evidence of linkage was observed in AA families on chromosome 4p16.1-p15.32 (fig. 4) and in EA families on chromosome 5p14.3-q11.2 (fig. 5). For chromosome 4, the maximum  $Z_{ir}$  of 2.72 was at 15.5 cM (6.92 Mb, centromeric to *D4S2366*), with a 1-LOD interval from 4pter to 35.5 cM (~22.65 Mb, centromeric to *D4S2639*) in AA families.  $Z_{ir}$  did not exceed 1.0 in the EA families on chromosome 4. For chromosome 5, the maximum  $Z_{ir}$  of 2.77 was at 43.2 cM (~31.61 Mb, telomeric to *D5S1470*), with a 1-LOD interval from 33.6 cM (~17.50 Mb, telomeric to *D5S2845*) to 74.3 cM (~66.62 Mb, telomeric to *D5S2500*) in EA families.  $Z_{ir}$  did not exceed 1.0 in the AA families in this region of chromosome 5.

#### Linkage Fine-Mapping Analysis

For the fine-mapping analysis, the deCODE map was used. We performed fine mapping for the four suggestive



**Figure 4** Results of the STRP genome scan for SZ on chromosome 4 and the subsequent SNP fine mapping. Positive  $Z_{ir}$  scores for the STRP scan for the entire sample (yellow line), EA subsample (blue line), and AA subsample (red line) are plotted on the Y-axis, and genomic position (CIDR cM converted to deCODE cM from p-ter, to allow plotting of the fine-mapping data) on chromosome 4 is plotted on the X-axis (0–60 cM). Similarly, positive  $Z_{ir}$  scores for the scanning STRPs plus the fine-mapping SNPs for the entire sample (orange line), EA subsample (green line), and AA subsample (purple line) are plotted on the Y-axis, and genomic position (deCODE cM from p-ter) on chromosome 4 is plotted on the X-axis (0– 60 cM).

regions (4p16.3-p15.2, 5p15.2-q13.3, 8p23.3-q12.1, and 11p13-q23.3) and two other regions that had multiple markers with a  $Z_{ir} > 2.0$  (10p15.3-p14 and 10q25.3-q26.3) (table 1). The average IC across the fine-mapping region for EA and AA families was 0.64 and 0.59, respectively, for the STRP map and was 0.85 and 0.79, respectively, for the map of 104 SNPs and 9 STRPs on chromosome 8p. The average IC for other fine-mapping regions for EA and AA families, respectively, was 0.83 and 0.77 for chromosome 4, 0.85 and 0.78 for chromosome 5, 0.80 and 0.77 for chromosome 10p, 0.83 and 0.78 for chromosome 10q, and 0.85 and 0.78 for chromosome 11 (table 1).

Figure 2 shows  $Z_{ir}$  scores for the full sample and for EA and AA families separately, with a combination of STRPs and SNPs on chromosome 8p. In the full sample, the maximum  $Z_{ir}$  score of 3.32 was observed at 44.4 cM (deCODE), just telomeric to *D8S1771*, where a similar peak ( $Z_{ir} = 3.24$ ) was observed in the scan analysis. In EA families, two peaks were again observed, although at slightly different locations than in the scan analysis: a telomeric peak ( $Z_{ir} = 3.33$ ) at 16.4 cM (deCODE), near the boundary of the inversion region (*D8S1130*), and a more centromeric peak ( $Z_{ir} = 3.10$ ) at 35.7 cM (deCODE), telomeric to *D8S560*. In AA families, a maximum  $Z_{ir}$  score of 1.68 was observed at 43.1 cM (de-

CODE), contributing to the peak result for the full sample. Results were essentially identical for analyses with the SNPs alone.

Therefore, fine mapping did not significantly increase or decrease the overall significance or length of the linkage peak in chromosome 8, despite an increase in IC. We note that the effect of the inversion on the analysis remains unclear. The genome-scan map contains two STRPs in the inversion region, as discussed above. We removed *D8S1469* from the linkage analysis of STRPs alone (and also in the fine mapping) so that only one STRP (*D8S1130*) remained in this region. In EA families, the peak  $Z_{ir}$  in this analysis remains very close to the inversion region. Until a more comprehensive analysis of the presence and sequence of inversions in this region can be performed, their effect on linkage results cannot be clearly understood. We consider the results of our analysis of 9 STRPs and 104 SNPs, with only 1 STRP within the inversion region, to be the most conservative estimation possible from the present data, and we conclude that genomewide suggestive linkage has been observed in the full sample, with two somewhat distinct peaks observed in the EA families.

However, we observed a more noticeable increase in significance for fine mapping of 5p15.2-q13.3 (fig. 5). In EA families, the 1-LOD interval was narrowed from

40.7 cM to 31.8 cM (table 1). It became clear that the centromere separated the linkage region into two peaks, and the maximum  $Z_{lr}$  score of the more significant peak increased from 2.77, telomeric to *D5S1470* (43.2 cM and ~31.6 Mb) (5p peak), to 3.80 at *rs1027164* (73.8 cM and 57.3 Mb) (5q peak) (fig. 5).  $Z_{lr}$  did not exceed 1 before or after fine mapping in the AA families (data not shown).

With regard to the fine mapping of chromosome 4p16.3-p15.2 that showed suggestive linkage in AA families and of chromosome 11p13-q23.3 that showed suggestive linkage in both the EA and the combined samples, the significance remained almost the same as before fine mapping (table 1 and figs. 3 and 4). However, we were able to narrow the 1-LOD interval for chromosome 11 from 32.7 cM to 19.5 cM in the full sample. For chromosome 10 regions, fine mapping did not improve any significance of linkage beyond a  $Z_{lr}$  score of 3 (table 1).

## Discussion

Genetic factors appear to predominate in the etiology of SZ, given a heritability estimated at 0.8. Most studies of the molecular genetics of SZ have investigated pos-

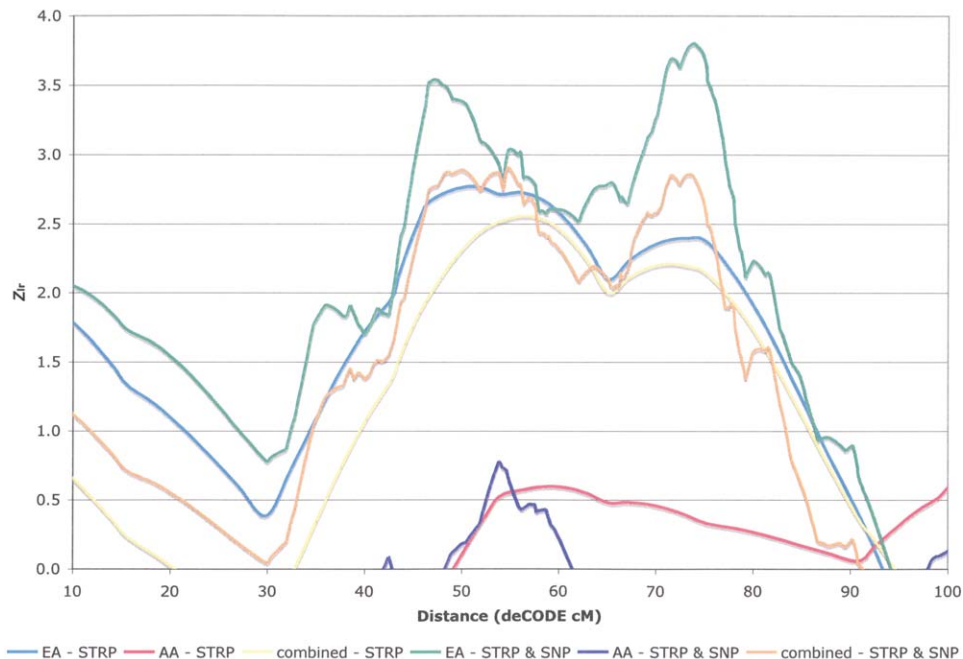
sible associations with functional polymorphisms in loci relevant to monoamine neurotransmission, which have been considered candidate genes on the basis of simple pharmacological models of SZ. Linkage studies have the advantage of not depending on any knowledge of the pathophysiology of the disorder. The most recent meta-analysis of SZ genome scans, including 20 analyses from 16 individual projects, showed significant evidence, across these studies, of linkage in a number of chromosomal regions (Lewis et al. 2003). In several of these regions, plausible candidate genes have been identified for which substantial evidence of association with SZ has been reported more than once (Craddock et al. 2005).

We have reported here one of the largest genome scans of SZ to date. The primary linkage analyses included predominantly EA or AA families. To take into account the possibility that these groups might differ in the effects of different SZ susceptibility loci, separate exploratory

**Table 7**

**STRPs and SNPs by Map Position and Multipoint  $Z_{lr}$  for Fine-Mapping Regions**

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.



**Figure 5** Results of the STRP genome scan for SZ on chromosome 5 and the subsequent SNP fine mapping. Positive  $Z_{lr}$  scores for the STRP scan for the entire sample (yellow line), EA subsample (blue line), and AA subsample (red line) are plotted on the Y-axis, and genomic position (CIDR cM converted to deCODE cM from p-ter, to allow plotting of the fine-mapping data) on chromosome 5 is plotted on the X-axis (10–100 cM). Similarly, positive  $Z_{lr}$  scores for the scanning STRPs plus the fine-mapping SNPs for the entire sample (orange line), EA subsample (green line), and AA subsample (purple line) are plotted on the Y-axis, and genomic position (deCODE cM from p-ter) on chromosome 5 is plotted on the X-axis (10–100 cM).



analyses of EA and AA families were also performed. Genomewide suggestive evidence of linkage was observed in the primary analysis on chromosomes 8p23.3-p12 and 11p11.2-q22.3, and, in the exploratory analyses, this threshold was also exceeded (without correction for multiple tests) on chromosome 4p16.1-p15.32 in AA families and on chromosome 5p14.3-q11.2 in EA families.

There are some limitations to our conclusions. For example, it is not known how to best define the phenotype for genetic studies of SZ. There is considerable support for the approach of combining the DSM-IV categories of SZ and SA, on the basis of their high relative risk in families of probands with SZ compared with the general population, their similar clinical characteristics in our (and other) samples, and the high reliability with which these diagnoses have been assigned. However, more-optimal phenotypes or endophenotypes might be identified in the future. Furthermore, despite the fact that this is a very large single sample of pedigrees with SZ, it lacks sufficient power to reliably detect loci with small effects.

The strongest evidence of linkage was at chromosome 8p23.3-p12. However, the presence of a common paracentric inversion in our strongest linkage region has complicated our analyses and the interpretation of results. It is possible, for example, that SZ is associated either with the presence of certain specific inversions or with one or more genes within the inversion segment. Also, failure to account for the presence of the inversion might have influenced the results of previous linkage analyses in this region (e.g., the location of the peak signals) in ways that are not yet clear. We evaluated the sensitivity of our linkage analyses to the genetic map assumptions, and linkage results remained basically unchanged whatever the assumed order of the SNPs located within the inverted region. Our results appeared to be more sensitive to the map length. Unsurprisingly,  $Z_{lr}$  scores varied with the assumed length, but these variations remained relatively moderate. For instance, a twofold decrease in map length led to a 6% reduction in  $Z_{lr}$  score. Conversely, the assumption of a twofold longer map led to a 5% increase in  $Z_{lr}$  score. In spite of these perhaps modest limitations related to the map effects of the inversion, evidence of a susceptibility locus for SZ in 8p has been detected in other family samples (table A3), and two positional candidate genes—*NRG1* and *protein phosphatase 3* (formerly *2B*), *catalytic subunit, gamma isoform (calcineurin A gamma)* (*PPP3CC*)—have been proposed to account for the linkages (Stefansson et al. 2002; Gerber et al. 2003), although neither can be considered definitively established (for example, see the detailed analysis of the reported association between *NRG1* and SZ [Duan et al. 2005]). The linkage signal in the present sample was extremely broad: nine STRP markers across 60 cM produced linkage signals at nominal  $P < .01$  ( $Z_{lr} > 2.326$ ) in the combined sample and in the EA fam-

ilies separately (table 5). We continued to observe suggestive evidence of linkage in the fine-mapping analysis, and it is improbable that a denser SNP map would substantially increase the linkage signal. In the EA families, which are largely responsible for the positive signal in this region, there appear to be two peaks. Given this, plus the repeated observation of evidence of SZ linkage in this region and the diverse locations of the peak evidence of linkage across studies, there might be more than one SZ susceptibility gene in this region. The 8p23-p12 region contains 239 annotated genes (UCSC Genome Bioinformatics, May 2004 freeze); this includes 137 genes that are annotated as full length, with an initiating ATG and valid stop-codon, that can be translated from the genome without frameshifts, and that use consensus splice sites (Consensus Coding DNA Sequence [CCDS] database, March 2005 report; see CCDS Web site). In addition to *NRG1* and *PPP3CC* in 8p, a substantial number of the genes in this region might be considered positional candidates because of their roles in brain function, and, given that the pathophysiology of SZ is largely unknown, there could be additional relevant genes. Therefore, large and systematic case-control and/or family-based LD mapping experiments should be undertaken in this region.

In summary, a genome scan of a large sample of pedigrees with SZ has produced suggestive evidence of linkage in two regions and evidence of two other suggestive regions in exploratory analyses of AA and EA families (see table 7 for results of fine mapping in our sample). Some of these regions have support from multiple other SZ linkage scans (see tables A3 and A4). The present study increases support for the hypothesis that one, or possibly more than one, SZ susceptibility gene is located on chromosome 8p. The SZ candidate regions observed in multiple samples should be studied systematically in large samples with the use of LD mapping methods.

## Acknowledgments

We thank the patients and families for their participation. Data and biomaterials from the NIMH Genetics Initiative for Schizophrenia MGS1 families were collected in nine projects. From 1999 to 2004, the principal investigators (PIs) and co-investigators were: University of Chicago, Chicago (grant R01 MH59571): P.V.G. (collaboration coordinator and PI) and A.R.S.; Baylor College of Medicine, Houston (grant R01 MH59587): F.A. (PI); University of California at Irvine, Irvine (grant R01 MH60870): W.F.B. (PI); University of Iowa, Iowa City (grant R01 MH59566): D.W.B. (PI) and R.R.C.; Washington University, St. Louis (grant R01 MH60879): C.R.C. (PI); University of Colorado, Denver (grant R01 MH59565): R.F. (PI) and A.O.; University of Pennsylvania, Philadelphia

(grant R01 MH61675): D.F.L. (PI), and subcontract to Louisiana State University, New Orleans: N.G.B. (subcontract PI); University of Queensland, Brisbane (grant R01 MH59588): B.J.M. (PI); and Mt. Sinai School of Medicine, New York (grant R01 MH59586): J.M.S. (PI). We also thank the individuals at each participating institution for their laboratory, database, and/or clinical contributions, especially Roberta Fishman, Layla Kassem, Eric B. Carpenter, Gregory J. Burrell, Christos Pantelis, Robert J. Barrett, and Douglas A. Fugman. Genotyping was performed at CIDR (for STRPs) and at Evanston Northwestern Healthcare Research Institute (for SNPs).

## Appendix A

**Table A1**

**Example of Renumbering to Assure that Allelic Designations Are Comparable between Waves I and II for *D6S2427***

FRAGMENT SIZE (bp)	ALLELE FREQUENCY		ORIGINAL ALLELIC DESIGNATION		RENUMBERED ALLELIC DESIGNATION	
	Wave I	Wave II	Wave I	Wave II	Wave I	Wave II
187	.003	.003	1	1	1	1
195	.004	...	2	...	2	...
199	.025	.019	3	2	3	3
200	...	.002	...	3	...	4
202	.057	...	4	...	5	...
203	...	.028	...	4	...	5
206	.116	...	5	...	6	...
207	.003	.148	6	5	7	6
209	...	.001	...	6	...	8
210	.283	.247	7	7	9	9
211	.003	.003	8	8	10	10
213	...	.001	...	9	...	11
214	.229	.256	9	10	12	12
215	.005	.002	10	11	13	13
218	.153	.153	11	12	14	14
219	...	.001	...	13	...	15
222	.060	.072	12	14	16	16
226	.038	.045	13	15	17	17
227	.001	...	14	...	18	...
230	.021	.016	15	16	19	19
234	.001	.002	16	17	20	20

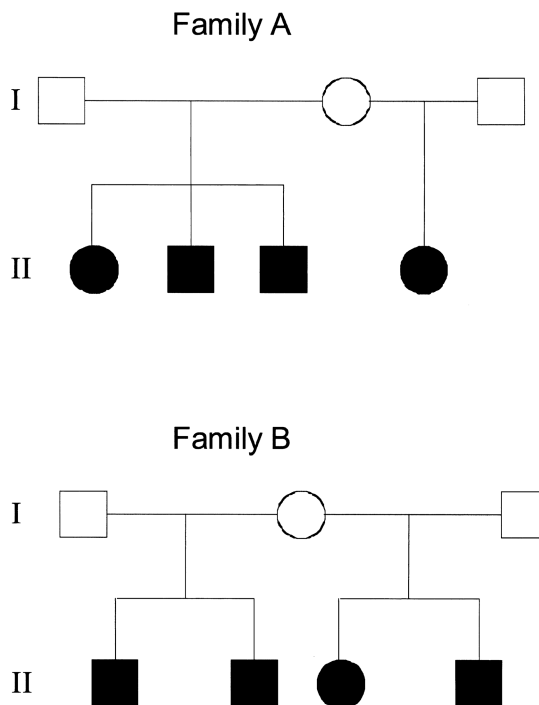
NOTE.—The presence of rare alleles in one wave but not the other means that the original allelic designations from CIDR cannot be used when the waves are combined. For instance, allele 10 in wave I has a frequency of 0.5%, whereas, in wave II, it has a frequency of 25.6%. In fact, the fragment of size 214 bp (relabeled as allele 12 when the waves are combined) has a frequency of 24.7% in this sample. When a large proportion of families have ungenotyped parents, as is the case here, the results of a linkage analysis necessarily will depend on the estimates of the allele frequencies. Accordingly, when genotypes are produced in different waves, by different laboratories, or by use of different technologies, it is critically important to assure that the alleles align correctly.

**Table A2**

**Example of Allelic Renumbering for *D15S128* Used to Combine the EA and AA Families**

FRAGMENT SIZE (bp)	ALLELE NUMBER	ALLELE FREQUENCY		RENUMBERED ALLELE (AA ONLY)
		EA	AA	
196	1	.0005	...	...
198	2	.0011	...	...
200	3	.0643	.0141	14
202	4	.0050	.1141	15
204	5	.0050	.1043	16
206	6	.2494	.3424	17
208	7	.2629	.0826	18
210	8	.1661	.1065	19
212	9	.1275	.0576	20
213/214	10	.0833	.0620	21
215/216	11	.0263	.0185	22
217	12	.0062	.0011	23
219	13	.0022	.0032	24
221	14	...	.0924	25
225	15	...	.0011	26

NOTE.—Shown are the allele frequencies for *D15S128* in the EA and AA families. The original recoding of this STRP resulted in 15 alleles. Alleles 1–13 are the only alleles present in the EA sample. Accordingly, we renumbered the alleles of the genotypes in the AA families, as indicated. In the linkage analysis of the combined sample, the allele frequencies sum to 200%, resulting in the correct computation of conditional probabilities of the genotypes of missing parents.



**Figure A1** Illustration of the counting of sib pairs. Because there is no optimal method for counting independent affected half sib pairs when there are full and half sibs in the same family, all possible half sib pairs (AHAPs) have been counted in these cases. Consider, for instance, the two families shown here. Pedigree A is counted as two statistically independent AFSPs and three AHAPs. Pedigree B is counted as two AFSPs and four AHAPs.



## Table A3

### Previous Chromosome 8p SZ Linkage Reports

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

## Table A4

### Previous Chromosome 5 Pericentromeric SZ Linkage Reports

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

## Web Resources

The URLs for data presented herein are as follows:

- Center for Inherited Disease Research (CIDR), <http://www.cidr.jhmi.edu/>
- Consensus Coding DNA Sequence (CCDS) database, <http://www.ncbi.nlm.nih.gov/projects/CCDS/> (for March 2005 report)
- dbSNP, <http://www.ncbi.nlm.nih.gov/SNP/> (for build 34)
- HapMap, <http://www.hapmap.org/>
- NIMH Center for Collaborative Genetics Studies of Mental Disorders, <http://zork.wustl.edu/nimh/>
- NIMH Human Genetics Initiative for Schizophrenia, [http://zork.wustl.edu/nimh/NIMH\\_initiative/NIMH\\_initiative\\_link.html](http://zork.wustl.edu/nimh/NIMH_initiative/NIMH_initiative_link.html)
- Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for SZ)
- UCSC Genome Bioinformatics, <http://genome.ucsc.edu/> (for May 2004 assembly of the human genome)
- UniSTS, <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unists> (for comprehensive database of sequence-tagged sites)

## References

- Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30:97–101
- American Psychiatric Association (APA) (1994) *Diagnostic and Statistical Manual of Mental Disorders*. APA, Washington, DC
- Badner JA, Gershon ES (2002) Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. *Mol Psychiatry* 7:405–411
- Bespalova IN, Angelo GW, Durner M, Smith CJ, Siever LJ, Buxbaum JD, Silverman JM (2005) Fine mapping of the 5p13 locus linked to schizophrenia and schizotypal personality disorder in a Puerto Rican family. *Psychiatr Genet* 15:205–210
- Blouin JL, Dombroski BA, Nath SK, Lasseter VK, Wolyniec PS, Nestadt G, Thornquist M, et al (1998) Schizophrenia susceptibility loci on chromosomes 13q32 and 8p21. *Nat Genet* 20:70–73
- Boehnke M (1991) Allele frequency estimation from data on relatives. *Am J Hum Genet* 48:22–25
- Boehnke M, Cox NJ (1997) Accurate inference of relationships in sib-pair linkage studies. *Am J Hum Genet* 61:423–429
- Broman KW, Matsumoto N, Giglio S, Martin CL, Roseberry JA, Zuffardi O, Ledbetter DH, Weber JL (2003) Common long human inversion polymorphism on chromosome 8p. In: Goldstein DR (ed) *Science and statistics: a festschrift for Terry Speed*. Vol 40. IMS lecture notes—monograph series. Institute of Mathematical Statistics, Beachwood, OH, pp 237–245
- Broman KW, Murray JC, Sheffield VC, White RL, Weber JL (1998)

- Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am J Hum Genet* 63:861–869
- Broman KW, Weber JL (1998) Estimation of pairwise relationships in the presence of genotyping errors. *Am J Hum Genet* 63:1563–1564
- Brzustowicz LM, Hodgkinson KA, Chow EW, Honer WG, Bassett AS (2000) Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22. *Science* 288:678–682
- Cardno AG, Marshall EJ, Coid B, Macdonald AM, Ribchester TR, Davies NJ, Venturi P, Jones LA, Lewis SW, Sham PC, Gottesman II, Farmer AE, McGuffin P, Reveley AM, Murray RM (1999) Heritability estimates for psychotic disorders: the Maudsley twin psychosis series. *Arch Gen Psychiatry* 56:162–168
- Cavalli-Sforza LL, Menozzi P, Piazza A (1994) *The history and geography of human genes*. Princeton University Press, Princeton, NJ
- Cloninger CR (2002) The discovery of susceptibility genes for mental disorders. *Proc Natl Acad Sci USA* 99:13365–13367
- 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
- Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20:37–46
- Cooper-Casey K, Mesen-Fainardi A, Galke-Rollins B, Llach M, LaPrade B, Rodriguez C, Riondet S, Bertheau A, Byerley W (2005) Suggestive linkage of schizophrenia to 5p13 in Costa Rica. *Mol Psychiatry* 10:651–656
- Craddock N, O'Donovan MC, Owen MJ (2005) The genetics of schizophrenia and bipolar disorder: dissecting psychosis. *J Med Genet* 42:193–204
- Curtis D, Sham PC (1996) Population stratifications can cause false positive linkage results if founders are untyped. *Ann Hum Genet* 60:261–263
- DeLisi LE, Mesen A, Rodriguez C, Bertheau A, LaPrade B, Llach M, Riondet S, Razi K, Relja M, Byerley W, Sherrington R (2002) Genome-wide scan for linkage to schizophrenia in a Spanish-origin cohort from Costa Rica. *Am J Med Genet* 114:497–508
- Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc* 39:1–38
- Duan J, Martinez M, Sanders AR, Hou C, Krasner AJ, Schwartz DB, Gejman PV (2005) Neuregulin 1 (NRG1) and schizophrenia: analysis of a US family sample and the evidence in the balance. *Psychol Med* 35:1599–1610
- Dunn G, Bertelsen S, Bierut L, Hinrichs A, Jin C, Kauwe J, Suarez BK. Microsatellites versus SNPs in linkage analysis for quantitative and qualitative measures. *BMC Genet* (in press)
- Garver DL, Holcomb J, Mapua FM, Wilson R, Barnes B (2001) Schizophrenia spectrum disorders: an autosomal-wide scan in multiplex pedigrees. *Schizophr Res* 52:145–160
- Gerber DJ, Hall D, Miyakawa T, Demars S, Gogos JA, Karayiorgou M, Tonegawa S (2003) Evidence for association of schizophrenia with genetic variation in the 8p21.3 gene, *PPP3CC*, encoding the calcineurin gamma subunit. *Proc Natl Acad Sci USA* 100:8993–8998
- Gershon ES, DeLisi LE, Hamovit J, Nurnberger JI Jr, Maxwell ME, Schreiber J, Dauphinais D, Dingman CW, Guroff JJ (1988) A controlled family study of chronic psychoses: schizophrenia and schizoaffective disorder. *Arch Gen Psychiatry* 45:328–336
- Giglio S, Broman KW, Matsumoto N, Calvari V, Gimelli G, Neumann T, Ohashi H, Voullaire L, Larizza D, Giorda R, Weber JL, Ledbetter DH, Zuffardi O (2001) Olfactory receptor-gene clusters, genomic-inversion polymorphisms, and common chromosome rearrangements. *Am J Hum Genet* 68:874–883
- Goddard KA, Hopkins PJ, Hall JM, Witte JS (2000) Linkage disequilibrium and allele-frequency distributions for 114 single-nucleo-

- tide polymorphisms in five populations. *Am J Hum Genet* 66:216–234
- Gottesman II, Bertelsen A (1989) Confirming unexpressed genotypes for schizophrenia: risks in the offspring of Fischer's Danish identical and fraternal discordant twins. *Arch Gen Psychiatry* 46:867–872
- Gottesman II, Shields J (1982) *Schizophrenia: the epigenetic puzzle*. Cambridge University Press, Cambridge, United Kingdom
- Grigull J, Alexandrova R, Paterson AD (2001) Clustering of pedigrees using marker allele frequencies: impact on linkage analysis. *Genet Epidemiol* 21:S61–S66
- Gurling HM, Kalsi G, Brynjolfsson J, Sigmundsson T, Sherrington R, Mankoo BS, Read T, Murphy P, Blaveri E, McQuillin A, Petursson H, Curtis D (2001) Genomewide genetic linkage analysis confirms the presence of susceptibility loci for schizophrenia, on chromosomes 1q32.2, 5q33.2, and 8p21–22 and provides support for linkage to schizophrenia, on chromosomes 11q23.3–24 and 20q12.1–11.23. *Am J Hum Genet* 68:661–673
- Heston LL (1966) Psychiatric disorders in foster home reared children of schizophrenic mothers. *Br J Psychiatry* 112:819–825
- Hinds DA, Stuve LL, Nilsen GB, Halperin E, Eskin E, Ballinger DG, Frazer KA, Cox DR (2005) Whole-genome patterns of common DNA variation in three human populations. *Science* 307:1072–1079
- Holmans P, Zubenko GS, Crowe RR, DePaulo JR Jr, Scheftner WA, Weissman MM, Zubenko WN, Boutelle S, Murphy-Eberenz K, MacKinnon D, McInnis MG, Marta DH, Adams P, Knowles JA, Gladis M, Thomas J, Chellis J, Miller E, Levinson DF (2004) Genomewide significant linkage to recurrent, early-onset major depressive disorder on chromosome 15q. *Am J Hum Genet* 74:1154–1167
- Huang Q, Shete S, Amos CI (2004) Ignoring linkage disequilibrium among tightly linked markers induces false-positive evidence of linkage for affected sib pair analysis. *Am J Hum Genet* 75:1106–1112
- Johns LC, Cannon M, Singleton N, Murray RM, Farrell M, Brugha T, Bebbington P, Jenkins R, Meltzer H (2004) Prevalence and correlates of self-reported psychotic symptoms in the British population. *Br J Psychiatry* 185:298–305
- Kaufmann CA, Suarez B, Malaspina D, Pepple J, Svrakic D, Markel PD, Meyer J, Zambuto CT, Schmitt K, Matisse TC, Harkavy Friedman JM, Hampe C, Lee H, Shore D, Wynne D, Faraone SV, Tsuang MT, Cloninger CR (1998) NIMH Genetics Initiative Millennium Schizophrenia Consortium: linkage analysis of African-American pedigrees. *Am J Med Genet* 81:282–289
- Kendler KS, McGuire M, Gruenberg AM, O'Hare A, Spellman M, Walsh D (1993a) The Roscommon Family Study. I. Methods, diagnosis of probands, and risk of schizophrenia in relatives. *Arch Gen Psychiatry* 50:527–540
- (1993b) The Roscommon Family Study. III. Schizophrenia-related personality disorders in relatives. *Arch Gen Psychiatry* 50:781–788
- (1993c) The Roscommon Family Study. IV. Affective illness, anxiety disorders, and alcoholism in relatives. *Arch Gen Psychiatry* 50:952–960
- Kendler KS, McGuire M, Gruenberg AM, Spellman M, O'Hare A, Walsh D (1993d) The Roscommon Family Study. II. The risk of nonschizophrenic nonaffective psychoses in relatives. *Arch Gen Psychiatry* 50:645–652
- Kety SS, Rosenthal D, Wender PH, Schulzinger F (1971) Mental illness in the biological and adoptive families of adopted schizophrenics. *Am J Psychiatry* 128:302–306
- 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, Jonsson GM, Gudjonsson SA, Richardsson B, Sigurdardottir A, 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
- Lange K, Weeks D, Boehnke M (1988) Programs for pedigree analysis: MENDEL, FISHER, and dGENE. *Genet Epidemiol* 5:471–472
- Leckman JF, Sholomskas D, Thompson WD, Belanger A, Weissman MM (1982) Best estimate of lifetime psychiatric diagnosis: a methodological study. *Arch Gen Psychiatry* 39:879–883
- Levinson DF, Holmans P. The effect of linkage disequilibrium on linkage analysis of incomplete pedigrees. *BMC Genet* (in press)
- Levinson DF, Mowry BJ, Escamilla MA, Faraone SV (2002) The Lifetime Dimensions of Psychosis Scale (LDPS): description and interrater reliability. *Schizophr Bull* 28:683–695
- Lewis CM, Levinson DF, Wise LH, DeLisi LE, Straub RE, Hovatta I, Williams NM, et al (2003) Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: schizophrenia. *Am J Hum Genet* 73:34–48
- Lindholm E, Ekholm B, Shaw S, Jalonen P, Johansson G, Pettersson U, Sherrington R, Adolfsson R, Jazin E (2001) A schizophrenia-susceptibility locus at 6q25, in one of the world's largest reported pedigrees. *Am J Hum Genet* 69:96–105
- Lindsay EA, Morris MA, Gos A, Nestadt G, Wolyniec PS, Lasseter VK, Shprintzen R, Antonarakis SE, Baldini A, Pulver AE (1995) Schizophrenia and chromosomal deletions within 22q11.2. *Am J Hum Genet* 56:1502–1503
- Maier W, Lichtermann D, Minges J, Hallmayer J, Heun R, Benkert O, Levinson DF (1993) Continuity and discontinuity of affective disorders and schizophrenia: results of a controlled family study. *Arch Gen Psychiatry* 50:871–883
- Matisse TC, Sachidanandam R, Clark AG, Kruglyak L, Wijsman E, Kakol J, Buyske S, et al (2003) A 3.9-centimorgan-resolution human single-nucleotide polymorphism linkage map and screening set. *Am J Hum Genet* 73:271–284
- Maxwell ME (1992) Family Interview for Genetic Studies (FIGS): a manual for FIGS. Clinical Neurogenetics Branch, Intramural Research Program, NIMH, Bethesda, MD
- Moises HW, Yang L, Kristbjarnarson H, Wiese C, Byerley W, Macciardi F, Arolt V, et al (1995) An international two-stage genome-wide search for schizophrenia susceptibility genes. *Nat Genet* 11:321–324
- 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. *Arch Gen Psychiatry* 51:849–859
- O'Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63:259–266
- O'Rourke DH, Gottesman II, Suarez BK, Rice J, Reich T (1982) Refutation of the general single-locus model for the etiology of schizophrenia. *Am J Hum Genet* 34:630–649
- Ott J (1985) A chi-square test to distinguish allelic association from other causes of phenotypic association between two loci. *Genet Epidemiol* 2:79–84
- Owen MJ, Williams NM, O'Donovan MC (2004) The molecular genetics of schizophrenia: new findings promise new insights. *Mol Psychiatry* 9:14–27
- Park N, Joo SH, Cheng R, Liu J, Loth JE, Lilliston B, Nee J, Grunn A, Kanyas K, Lerer B, Endicott J, Gilliam TC, Baron M (2004) Linkage analysis of psychosis in bipolar pedigrees suggests novel putative loci for bipolar disorder and shared susceptibility with schizophrenia. *Mol Psychiatry* 9:1091–1099
- Paunio T, Ekelund J, Varilo T, Parker A, Hovatta I, Turunen JA, Rinard K, Foti A, Terwilliger JD, Juvonen H, Suvisaari J, Arajärvi R, Suokas J, Partonen T, Lonnqvist J, Meyer J, Peltonen L (2001) Genome-wide scan in a nationwide study sample of schizophrenia families

- in Finland reveals susceptibility loci on chromosomes 2q and 5q. *Hum Mol Genet* 10:3037–3048
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Pulver AE, Lasseter VK, Kasch L, Wolyniec P, Nestadt G, Blouin JL, Kimberland M, Babb R, Vourlis S, Chen H, Lalioti M, Morris MA, Karayiorgou M, Ott J, Meyers D, Antonarakis SE, Housman D, Kazazian HH (1995) Schizophrenia: a genome scan targets chromosomes 3p and 8p as potential sites of susceptibility genes. *Am J Med Genet* 60:252–260
- Rosenthal D, Wender PH, Kety SS, Welner J, Schulsinger F (1971) The adopted-away offspring of schizophrenics. *Am J Psychiatry* 128:307–311
- Silverman JM, Greenberg DA, Altstiel LD, Siever LJ, Mohs RC, Smith CJ, Zhou G, Hollander TE, Yang XP, Kedache M, Li G, Zaccario ML, Davis KL (1996) Evidence of a locus for schizophrenia and related disorders on the short arm of chromosome 5 in a large pedigree. *Am J Med Genet* 67:162–171
- SLCG (1996) Additional support for schizophrenia linkage on chromosomes 6 and 8: a multicenter study. Schizophrenia Linkage Collaborative Group for Chromosomes 3, 6 and 8. *Am J Med Genet* 67:580–594
- Stefansson H, Helgason A, Thorleifsson G, Steinthorsdottir V, Masson G, Barnard J, Baker A, et al (2005) A common inversion under selection in Europeans. *Nat Genet* 37:129–137
- Stefansson H, Sigurdsson E, Steinthorsdottir V, Bjornsdottir S, Sigurdsson T, Ghosh S, Brynjolfsson J, et al (2002) *Neuregulin 1* and susceptibility to schizophrenia. *Am J Hum Genet* 71:877–892
- Stompe T, Willinger U, Fischer G, Meszaros K, Berger P, Strobl R, Berger K, Isenberg E, Todd RD, Cloninger CR, Reich T, Aschauer HN (1998) The unified biosocial model of personality in schizophrenia families and controls. *Psychopathology* 31:45–51
- Straub RE, MacLean CJ, Ma Y, Webb BT, Myakishev MV, Harris-Kerr C, Wormley B, Sadek H, Kadambi B, O'Neill FA, Walsh D, Kendler KS (2002) Genome-wide scans of three independent sets of 90 Irish multiplex schizophrenia families and follow-up of selected regions in all families provides evidence for multiple susceptibility genes. *Mol Psychiatry* 7:542–559
- Strauss JS (1969) Hallucinations and delusions as points on continua function: rating scale evidence. *Arch Gen Psychiatry* 21:581–586
- Sugawara H, Harada N, Ida T, Ishida T, Ledbetter DH, Yoshiura K, Ohta T, Kishino T, Niikawa N, Matsumoto N (2003) Complex low-copy repeats associated with a common polymorphic inversion at human chromosome 8p23. *Genomics* 82:238–244
- Taylor MA, Berenbaum SA, Jampala VC, Cloninger CR (1993) Are schizophrenia and affective disorder related? Preliminary data from a family study. *Am J Psychiatry* 150:278–285
- Terwilliger JD, Speer M, Ott J (1993) Chromosome-based method for rapid computer simulation in human genetic linkage analysis. *Genet Epidemiol* 10:217–224
- van Os J, Hanssen M, Bijl RV, Ravelli A (2000) Strauss (1969) revisited: a psychosis continuum in the general population? *Schizophr Res* 45:11–20
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, et al (2001) The sequence of the human genome. *Science* 291:1304–1351