# No Significant Association of 14 Candidate Genes With Schizophrenia in a Large European Ancestry Sample: Implications for Psychiatric Genetics

Alan R. Sanders, M.D.
Jubao Duan, Ph.D.
Douglas F. Levinson, M.D.
Jianxin Shi, Ph.D.
Deli He, B.S.
Cuiping Hou, B.S.
Gregory J. Burrell, B.S.
John P. Rice, Ph.D.
Deborah A. Nertney, B.S.
Ann Olincy, M.D.
Pablo Rozic, M.D.
Sophia Vinogradov, M.D.
Nancy G. Buccola, A.P.R.N.B.C.
Bryan J. Mowry, M.D.
Robert Freedman, M.D.
Farooq Amin, M.D.
Donald W. Black, M.D.
Jeremy M. Silverman, Ph.D.
William F. Byerley, M.D.
Raymond R. Crowe, M.D.
C. Robert Cloninger, M.D.
Maria Martinez, Ph.D.
Pablo V. Gejman, M.D.

**Objective:** The authors carried out a genetic association study of 14 schizophrenia candidate genes (*RGS4, DISC1, DTNBP1, STX7, TAAR6, PPP3CC, NRG1, DRD2, HTR2A, DAOA, AKT1, CHRNA7, COMT*, and *ARVCF*). This study tested the hypothesis of association of schizophrenia with common single nucleotide polymorphisms (SNPs) in these genes using the largest sample to date that has been collected with uniform clinical methods and the most comprehensive set of SNPs in each gene.

**Method:** The sample included 1,870 cases (schizophrenia and schizoaffective disorder) and 2,002 screened comparison subjects (i.e. controls), all of European ancestry, with ancestral outliers excluded based on analysis of ancestry-informative markers. The authors genotyped 789 SNPs, including tags for most common SNPs in each gene, SNPs previously reported as associated, and SNPs located in functional domains of genes such as promoters, coding exons (including nonsynonymous SNPs), 3′ untranslated regions, and conserved noncoding sequences. After extensive data cleaning, 648 SNPs were analyzed for association of single SNPs and of haplotypes.

**Results:** Neither experiment-wide nor gene-wide statistical significance was observed in the primary single-SNP analyses or in secondary analyses of haplotypes or of imputed genotypes for additional common HapMap SNPs. Results in SNPs previously reported as associated with schizophrenia were consistent with chance expectation, and four functional polymorphisms in *COMT, DRD2*, and *HTR2A* did not produce nominally significant evidence to support previous evidence for association.

**Conclusions:** It is unlikely that common SNPs in these genes account for a substantial proportion of the genetic risk for schizophrenia, although small effects cannot be ruled out.

*(Am J Psychiatry 2008; 165:497–506)*

The intensive search for DNA sequence variation underlying susceptibility to schizophrenia has been motivated by evidence that etiology is predominantly genetic: heritability is ~80% based on twin studies (1), with overlapping risks of schizophrenia and schizoaffective disorder in families and a pattern of illness in families that suggests complex mechanisms involving multiple genes of small effect (2, 3). Currently, the genetic mechanisms remain unknown.

A decade ago, genes involved in monoaminergic neurotransmission were the most widely studied schizophrenia "candidate genes" because drugs that blocked dopamine receptors were the best available treatments. A new set of mostly "positional" candidate genes has now emerged—disease-related genes identified by their location in relation to DNA markers or cytogenetic abnormalities (4, 5). These genes are involved in pathways that can plausibly be related to mechanistic hypotheses of schizophrenia. We present here a study of the association of schizophrenia to DNA sequence variants in 14 of the best-supported of these current candidate genes selected on the basis of our reading of the literature; others might cre-

This article is featured in this month's AJP Audio and is discussed in an editorial by Dr. Hamilton (p. 420).

ate a slightly different list, but these genes would be given consideration by most investigators.

The study has two important features. First, the large sample was collected by uniform methods, whereas most schizophrenia samples are smaller or were assembled from separate studies. We studied subjects of European ancestry (the larger of the two ancestry groups in our sample) because most previous support for these associations has come from this population, and findings can be confounded by the varying frequencies of many DNA sequence variants across populations. Second, we tested dense sets of single nucleotide polymorphisms (SNPs) in each gene (rather than a few), including "tags" for most known common SNPs, plus additional SNPs in critical gene elements such as those that change amino acid sequence.

The selected genes include the following: *RGS4, DISC1, DTNBP1, STX7, TAAR6, PPP3CC, NRG1, DRD2, HTR2A, DAOA, AKT1, CHRNA7, COMT*, and *ARVCF* (data supplement Table 2 available at http://ajp.psychiatryonline.org). The strength of previous evidence for association varies among these genes (6, 7). Generally, one or more studies reported an experiment-wide significant result, but no finding has been consistently observed.

The interpretation of genetic association studies depends on the hypothesis. The reported associations in these genes are generally for "common" SNPs (typically defined as those present on at least 5% of chromosomes). Although we cannot yet test all rare and common DNA variation by direct sequencing, we can systematically study common SNPs because the HapMap project (hapmap.org) has catalogued a large proportion of common SNPs genome-wide and shown how to "tag" them with subsets and because high-throughput technologies can now test them accurately (8). We agree with Todd (9) that to firmly establish a finding of association with a common SNP, one should observe evidence across studies that clearly exceeds a statistical threshold (probably near $p=10^{-7}$ for tests of single SNPs [10]) that takes into account all common variation in the genome. (Note that very low p values are often reported for combinations of SNP alleles [haplotypes], but testing haplotypes requires many more statistical tests and therefore an even more stringent threshold.) None of these 14 genes has produced association evidence at this level in a single study or across studies.

We tested these genes in a large sample using SNPs that tagged most common variants plus SNPs previously reported as associated and additional known SNPs in functional elements. While recognizing that ultimately more stringent statistical thresholds must be achieved to account for testing SNPs throughout the genome with a low prior probability of association for any one SNP and given the low prior probability of any single candidate gene association being "true," in the context of the absence of well-established pathophysiological hypotheses for schizophrenia, we have applied two empirically derived criteria of significance: one that accounts for all tests in this experiment (considered the primary criterion here) and one that accounts for all tests in each gene. A gene-wide threshold would be most appropriate for an association that had been rigorously established by previous studies. This does not appear to be the case, but given that this is the first large-scale systematic study of most of these genes, it is important to avoid false negative as well as false positive results. We have also considered whether in SNPs or haplotypes with previous experiment-wide evidence for association we observed nominally significant results more frequently than expected by chance. We were unable to detect association of any one SNP with schizophrenia by any of these criteria. The implications of these findings are further discussed below.

## Method

Complete details about the method and results are available in the data supplement (text, tables, figures). We provide here a summary of the most pertinent information.

### *Subjects*

The study sample included 1,952 unrelated individuals with a diagnosis of schizophrenia or schizoaffective disorder and 2,126 comparison subjects. After the quality control checks described below, 1,870 case and 2,002 comparison subjects were included in analyses.

Cases were recruited in three related studies (Table 1). Most were recruited by the Molecular Genetics of Schizophrenia Part 2 study. The present investigators are currently completing the recruitment of this large case-control sample of European ancestry and African American individuals for genetic association studies of schizophrenia, which is part of an NIMH repository program (nimhgenetics.org). The present study includes approximately two-thirds of the Molecular Genetics of Schizophrenia Part 2 European ancestry sample. The remaining subjects are from the Molecular Genetics of Schizophrenia Part 1 (11) and Schizophrenia Genetics Initiative (12, 13) studies of multiply affected pedigrees, with one case from each eligible family included here. Approximately 20% of cases had a first- or second-degree relative with a known or suspected history of schizophrenia. Recruitment sites are listed in data supplement Table 3. These subjects of European ancestry (by self-report) included some cases from Australia, where European ancestry is similar to the United States (14). Cases (ages 18 and over) were identified from clinics, hospitals, physician referrals, advocacy and support organizations, and Internet and media announcements and advertisements.

All case subjects signed institutional review board-approved written informed consent forms that authorized deposition of their biological materials and nonidentifying clinical information in NIMH repository for use in genetic studies. Cases were interviewed by trained clinicians with the Diagnostic Interview for Genetic Studies 2.0 (15) to elicit DSM-IV diagnostic and symptom information for psychotic, mood, and substance use disorders. In 98.6% of the cases, two of three possible types of information were obtained: Diagnostic Interview for Genetic Studies, Family Interview for Genetic Studies interview with an informant (16), and psychiatric records; 26 cases who could not be meaningfully interviewed were diagnosed with high confidence by means of psychiatric records alone. Two senior clinicians independently reviewed all information and then assigned a primary consensus best-estimate final diagnosis (17) and comorbid diagnoses. Eligible cases received a "definite" or "likely" consensus best-estimate final diag-

**TABLE 1. Characteristics of the Case Sample[a]**

| | Cases | | Male | | Schizophrenia | | Schizoaffective Disorder, Manic Type | | Schizoaffective Disorder, Depressed Type | |
|---|---|---|---|---|---|---|---|---|---|---|
| Study | N | % | N | % by Study | N | Study % | N | Study % | N | Study % |
| NIMH Schizophrenia Genetics Initiative | 56 | 3.0 | 42 | 75.0 | 56 | 100.0 | 0 | 0.0 | 0 | 0.0 |
| Molecular Genetics of Schizophrenia Part 1 | 239 | 12.8 | 168 | 70.3 | 239 | 100.0 | 0 | 0.0 | 0 | 0.0 |
| Molecular Genetics of Schizophrenia Part 2 | 1,575 | 84.2 | 1,088 | 69.1 | 1,383 | 87.8 | 121 | 7.7 | 71 | 4.5 |
| Total cases | 1,870 | 100.0 | 1,298 | 69.4 | 1,678 | 89.7 | 121 | 6.5 | 71 | 3.8 |

[a] Shown are the numbers of cases (after all data cleaning and ancestry exclusions) subdivided by sex and diagnosis and by the recruiting study. All cases were of self-reported European ancestry and clustered with other individuals of European ancestry in analysis of ancestry-informative SNP markers (see text). The NIMH Schizophrenia Genetics Initiative (12) and the Molecular Genetics of Schizophrenia Part 1 (11) studies recruited multiply affected pedigrees for linkage analysis; one case per eligible family has been included here. There was a small overlap between the NIMH Schizophrenia Genetics Initiative sample (accounted for 3% of the cases analyzed here) and the samples in which associations were first reported in samples of European ancestry for five of these 14 genes, although this did not affect the results (see data supplement). The Molecular Genetics of Schizophrenia Part 2 study recruited unrelated cases and comparison subjects (see text).

nosis of schizophrenia or schizoaffective disorder, with psychosis judged unlikely to have been caused by substance use or medical illness and without moderate or severe mental retardation.

Blood specimens for U.S. participants were shipped overnight to the Rutgers University Cell and DNA Repository for transformation to lymphoblastic cell lines and DNA extraction; in Australia, lymphoblastic cell lines were established at Queensland Institute for Medical Research and aliquots shipped to Rutgers.

A marketing research company, Knowledge Networks (Menlo Park, Calif.), recruited the comparison subjects (Molecular Genetics of Schizophrenia Part 2). Knowledge Network's national online participant panel, recruited by random-digit dialing of residential phone numbers, is demographically similar to the U.S. population (age, sex, education, metropolitan/nonmetropolitan residence) (data supplement Table 4). A member of approximately 30% of targeted households joined the panel. Those without Internet access were given a web TV. Approximately 60,000 individuals of European ancestry were in the panel at some point during Molecular Genetics of Schizophrenia Part 2 recruitment; 15,485 were randomly selected, sent a letter explaining the study, then sent an e-mail message pointing to a web site to learn more about the study, given preliminary online informed consent, and completed a self-report clinical assessment; 3,364 (21.7%) completed these procedures and gave a blood sample, collected by Examination Management Services Inc. (Irving, Tex.), which also obtained written informed consent authorizing NIMH to use the biological materials and clinical information for any medical research study. We anonymized the comparison sample by destroying any hard copy materials (e.g., written informed consent forms) or computer files with links between identification numbers and personal identifiers.

The online assessment included the Composite International Diagonostic Interview—Short Form (18), modified for lifetime common mood, anxiety, and substance use disorders; items for lifetime diagnosis or treatment of psychosis or bipolar disorder; a nicotine dependence screen; neuroticism and extraversion personality scales (19); and items for sexual orientation, current height and weight, highest lifetime weight, and ancestral background, plus previously collected demographic information. We excluded 9.4% of the comparison subjects (0.4% endorsed more than 50 of 69 screening or personality items; 0.4% failed to answer five or more of these questions; 0.6% were not fully screened owing to software failures; and 8% endorsed or failed to deny previous treatment or diagnosis of schizophrenia, schizoaffective disorder, auditory hallucinations, delusions, or bipolar disorder).

### *Self-Reported Ancestry*

Cases reported up to four ancestries for each parent, and comparison subjects reported ancestries for each grandparent. We excluded cases mentioning non-European ancestry (except partial Native American ancestry, which was overreported). Reported ancestries were similar for cases and comparison subjects (data supplement Figure 1).

### *SNP Selection, Genotyping, and Quality Control*

We genotyped 224 ancestry-informative SNP markers (SNPlex genotyping system, Applied Biosystems, Foster City, Calif.) reported to differentiate European ancestry from African, Amerindian, or Asian ancestries (20–22), including rs4988235 (located ~14 kilobases upstream from lactase and associated with lactase persistence), whose frequency varies north-south across Europe (23).

We genotyped 789 SNPs (756 by SNPlex only, 20 by TaqMan [Applied Biosystems] only, and 13 by both methods) in 14 candidate gene regions (2.38 Mb of sequence), including SNPs previously reported as associated with schizophrenia, SNPs tagging most known common variation (based on HapMap I when this study was planned), and additional SNPs in putative functional gene elements, i.e., promoters, coding exons, 3′ untranslated regions, and conserved noncoding sequences (Table 2).

We excluded 164 SNPs (30 ancestry-informative SNP markers, 134 in candidate genes) whose genotypes were called in less than 90% of samples or had inconsistent clustering by inspection, including seven monomorphic candidate gene SNPs and nine chromosome X ancestry-informative SNP markers. Five candidate gene SNPs showed departures from Hardy-Weinberg equilibrium ($p<0.0001$ by exact statistics) (PLINK [24]); none of these five SNPs showed evidence for association. Valid SNPs included 185 autosomal ancestry-informative SNP markers and 648 SNPs for association tests of candidate genes (including 433 tag SNPs to assess common variation). Pair-wise tagging analysis (Tagger [25]) at an $r^2$ threshold of 0.8 showed that candidate gene SNPs captured 94% of HapMap I common variants (minor allele frequency>0.05) (range=84%–100% for individual genes) and 83% for HapMap II (range=62%–93%) (data supplement Figure 2).

### *Sample Quality Control and Ancestry Analyses*

Of genotyped cases (1,952) and comparison subjects (2,126) (data supplement Table 3), we excluded 12 cases and 51 comparison subjects with aggregate genotype call rates less than 95% across all valid SNPs, two cases and 23 comparison subjects with unresolved sex typing (amelogenin) discrepancies, five case and two comparison samples that were duplicates of another sample, 13 comparison subjects who were apparently related to another

**TABLE 2. Single Nucleotide Polymorphism (SNP) Coverage of 14 Schizophrenia Candidate Genes[a]**

| | Length of Kilobases | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Gene Symbol | RefSeq gene | Tagged gene | Promoters | Exon | Intron | Conserved Noncoding Sequence | Untranslated Region | Number of Exons |
| *RGS4* | 7 | 12 | 6 | 2 | 4 | — | 2.2 | 5 |
| *DISC1* | 414 | 421 | 6 | 14 | 395 | 7.5 | — | 20 |
| *DTNBP1* | 140 | 144 | 8 | 3 | 137 | 0.1 | 0.2 | 16 |
| *STX7* | 53 | 62 | 6 | 3 | 50 | — | 6 | 10 |
| *TAAR6* | 1 | 5 | 2 | 1 | — | — | — | 1 |
| *PPP3CC* | 100 | 104 | 6 | 3 | 97 | — | 0.3 | 16 |
| *NRG1* | 1104 | 1132 | 16 | 8 | 1103 | 13.3 | 3.6 | 19 |
| *DRD2* | 66 | 70 | 8 | 3 | 61 | 1.1 | 0.8 | 9 |
| *HTR2A* | 63 | 69 | 4 | 2 | 60 | 0.8 | 3.3 | 4 |
| *DAOA* | 25 | 44 | 6 | 4 | 36 | 0.5 | – | 13 |
| *AKT1* | 26 | 30 | 6 | 4 | 22 | 0.1 | 1.0 | 11 |
| *CHRNA7* | 139 | 142 | 4 | 2 | 131 | 0.5 | 4.6 | 10 |
| *COMT* | 27 | 32 | 18 | 9 | 19 | – | 0.2 | 16 |
| *ARVCF* | 47 | 54 | 12 | 8 | 38 | 0.1 | 4.6 | 21 |
| Total | 2,212 | 2,323 | 108 | 65 | 2,151 | 24.1 | 26.8 | 171 |

[a] Gene symbols are from the Human Genome Organisation (HUGO) Gene Nomenclature Committee. Sizes (in kilobases) are quoted from RefSeq. Tagged gene length includes an additional 2 kilobases on each side of each gene plus splice variants extending beyond RefSeq boundaries. Isoform GGF2 was used for *NRG1*. For *DAOA*, part of overlapping *G30* gene was tagged. Sizes and numbers of exons are derived from RefSeq, University of California at Santa Cruz, AceView, and Visualization Tool for Alignment, taking into account alternative splicing and partial overlaps among classes of gene segments (so that the total of the lengths of single elements is slightly longer than the tagged gene length). The number of SNPs in the nonsynonymous column are a subset of the number of SNPs in the exon column. Exon refers to the translated region. Nonsynonymous means change in coded amino acid sequence. Previous association=previously reported as associated with schizophrenia (these SNPs are also counted in the column for each gene domain). "Cleaned SNPs"=SNPs passing all quality control filters. All of the 70 SNPs previously reported to be associated with schizophrenia were successfully genotyped except rs4262285 in *NRG1*. Of the attempted 789 SNPs, the genotyping failures were distributed among the genes roughly proportional to the number of SNPs attempted for each gene (related most strongly to gene length), with lower minor allele frequency SNPs and less well-validated SNPs being overrepresented in the failures.

**TABLE 3. Single SNP Association Tests With Empirical Pointwise p<0.05[a]**

| | | | Allele | | Minor Allele Frequency | |
|---|---|---|---|---|---|---|
| Gene | SNP | Position | Minor/Major | Associated | Cases | Comparison Subjects |
| *RGS4* | rs2661319*^ | 159,771,435 | A/G | G | 0.470 | 0.493 |
| *RGS4* | rs2842030 | 159,772,153 | G/T | T | 0.420 | 0.443 |
| *DISC1* | rs10864695^ | 228,191,952 | G/A | G | 0.198 | 0.179 |
| *DISC1* | rs9431997^ | 228,234,115 | G/C | G | 0.099 | 0.083 |
| *DISC1* | rs9432010 | 228,247,572 | G/A | G | 0.101 | 0.085 |
| *DISC1* | rs17768115^ | 228,279,546 | G/C | C | 0.199 | 0.217 |
| *DISC1* | rs2038636^ | 228,361,984 | C/T | T | 0.400 | 0.423 |
| *STX7* | rs3183732^ | 132,823,032 | A/G | G | 0.291 | 0.321 |
| *STX7* | rs12207033^ | 132,824,835 | T/C | C | 0.087 | 0.105 |
| *STX7* | rs3757298 | 132,826,709 | C/A | A | 0.086 | 0.103 |
| *STX7* | rs4470875^ | 132,877,343 | T/C | C | 0.077 | 0.093 |
| *TAAR6* | rs8192625*^ | 132,934,025 | A/G | G | 0.066 | 0.079 |
| *NRG1* | rs3802158^ | 32,524,438 | A/G | G | 0.439 | 0.466 |
| *NRG1* | rs7834206 | 32,525,690 | A/C | C | 0.440 | 0.466 |
| *NRG1* | rs4733130 | 32,526,536 | C/T | T | 0.440 | 0.468 |
| *NRG1* | rs2439305 | 32,549,006 | C/T | T | 0.433 | 0.463 |
| *NRG1* | rs7825588^ | 32,623,943 | A/G | A | 0.134 | 0.118 |
| *DRD2* | rs17529477^ | 112,822,277 | A/G | G | 0.297 | 0.322 |
| *DRD2* | rs4245147 | 112,823,217 | C/T | T | 0.461 | 0.484 |
| *DRD2* | rs7131056^ | 112,834,984 | A/C | A | 0.452 | 0.424 |
| *HTR2A* | rs4941573^ | 46,362,858 | G/A | G | 0.425 | 0.402 |
| *HTR2A* | rs6313* | 46,367,941 | T/C | T | 0.425 | 0.400 |
| *HTR2A* | rs6311* | 46,369,479 | T/C | T | 0.425 | 0.400 |
| *DAOA* | rs1539070^ | 104,922,458 | G/C | G | 0.193 | 0.175 |
| *DAOA* | rs1557072^ | 104,925,835 | T/C | T | 0.012 | 0.007 |
| *AKT1* | rs2498794^ | 104,316,296 | C/T | T | 0.468 | 0.491 |
| *CHRNA7* | rs10438342^ | 30,189,338 | A/G | G | 0.347 | 0.370 |
| *CHRNA7* | rs2221223^ | 30,198,685 | C/A | A | 0.144 | 0.161 |
| *COMT* | rs3788319^ | 18,304,105 | A/G | A | 0.492 | 0.468 |
| *ARVCF* | rs2012714^ | 18,352,201 | A/G | G | 0.318 | 0.344 |

[a] Shown are all the 30 SNPs with an empirical pointwise p<0.05 value (Armitage trend test; p<0.01 bolded). Nominal pointwise and empirical gene-wide p values are also shown for this test, as well as for the classical allelic $\chi^2$ test and for the EIGENSTRAT $\chi^2$ test that corrects for population substructure. Carets indicate tag SNPs, and asterisks indicate SNPs that have previously been reported as associated with schizophrenia.

| Number of Cleaned SNPs Analyzed | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Exon | Non-synonymous | Untranslated Region | Promoter | Conserved Noncoding Sequence | Intron | Previous Association | Number of SNPs | Proportion |
| 4 | 0 | 1 | 3 | 0 | 4 | 4 | 12 | 0.02 |
| 21 | 3 | 0 | 2 | 33 | 59 | 14 | 115 | 0.18 |
| 15 | 3 | 1 | 4 | 1 | 17 | 12 | 38 | 0.06 |
| 10 | 1 | 4 | 5 | 0 | 8 | 2 | 27 | 0.04 |
| 7 | 7 | 0 | 3 | 0 | 7 | 4 | 17 | 0.03 |
| 13 | 0 | 0 | 2 | 0 | 6 | 4 | 21 | 0.03 |
| 17 | 4 | 1 | 7 | 62 | 130 | 11 | 217 | 0.33 |
| 11 | 2 | 4 | 5 | 6 | 6 | 2 | 32 | 0.05 |
| 8 | 5 | 2 | 2 | 6 | 16 | 2 | 34 | 0.05 |
| 17 | 3 | 0 | 2 | 3 | 8 | 6 | 30 | 0.05 |
| 11 | 0 | 0 | 2 | 1 | 3 | 2 | 17 | 0.03 |
| 1 | 1 | 0 | 2 | 3 | 12 | 1 | 18 | 0.03 |
| 18 | 1 | 0 | 10 | 0 | 1 | 3 | 29 | 0.04 |
| 21 | 3 | 2 | 8 | 1 | 9 | 2 | 41 | 0.06 |
| 174 | 33 | 15 | 57 | 116 | 286 | 69 | 648 | 1.00 |

| | Armitage Trend Test p Values | | | | |
|---|---|---|---|---|---|
| | Nominal | Empirical | | | |
| Odds Ratio | Pointwise | Pointwise | Gene-Wide | Pointwise Allelic $\chi^2$ | EIGENSTRAT $\chi^2$ |
| 1.10 | <0.05 | <0.05 | 0.26 | 0.05 | 0.05 |
| 1.10 | <0.05 | <0.05 | 0.23 | 0.05 | 0.05 |
| 1.13 | <0.04 | <0.04 | 0.87 | 0.04 | 0.02 |
| 1.22 | <0.02 | <0.02 | 0.58 | 0.01 | 0.03 |
| 1.21 | <0.02 | <0.02 | 0.69 | 0.02 | 0.03 |
| 1.12 | <0.05 | <0.05 | 0.94 | 0.05 | 0.08 |
| 1.10 | <0.04 | <0.04 | 0.91 | 0.04 | 0.02 |
| 1.15 | 0.004 | **0.004** | 0.051 | 0.00 | 0.00 |
| 1.23 | 0.009 | **0.009** | 0.11 | 0.01 | 0.01 |
| 1.22 | <0.02 | <0.02 | 0.15 | 0.01 | 0.01 |
| 1.22 | <0.02 | <0.02 | 0.18 | 0.02 | 0.02 |
| 1.22 | <0.03 | 0.026 | 0.28 | 0.03 | 0.03 |
| 1.12 | <0.02 | <0.02 | 0.83 | 0.02 | 0.03 |
| 1.11 | <0.02 | 0.02 | 0.86 | 0.02 | 0.03 |
| 1.12 | <0.02 | <0.02 | 0.77 | 0.02 | 0.02 |
| 1.13 | 0.008 | **0.008** | 0.60 | 0.01 | 0.01 |
| 1.15 | <0.05 | <0.05 | 0.98 | 0.05 | 0.07 |
| 1.12 | <0.02 | <0.02 | 0.24 | 0.02 | 0.01 |
| 1.10 | <0.05 | <0.05 | 0.47 | 0.05 | 0.04 |
| 1.12 | <0.02 | <0.02 | 0.18 | 0.01 | 0.01 |
| 1.10 | <0.05 | <0.05 | 0.54 | 0.04 | 0.07 |
| 1.11 | <0.03 | <0.03 | 0.41 | 0.03 | 0.05 |
| 1.11 | <0.04 | <0.04 | 0.44 | 0.03 | 0.05 |
| 1.13 | 0.051 | <0.05 | 0.53 | 0.05 | 0.04 |
| 1.84 | <0.02 | <0.02 | 0.18 | 0.01 | 0.01 |
| 1.10 | <0.05 | <0.05 | 0.32 | 0.04 | 0.03 |
| 1.10 | <0.05 | <0.05 | 0.42 | 0.04 | 0.05 |
| 1.14 | <0.05 | <0.05 | 0.41 | 0.04 | 0.05 |
| 1.10 | <0.05 | <0.05 | 0.47 | 0.04 | 0.03 |
| 1.13 | <0.03 | <0.02 | 0.26 | 0.02 | 0.02 |

comparison subject, and 63 case and 35 comparison subject specimens that lay outside the European ancestry cluster in a principle components analysis of ancestry-informative SNP marker data (EIGENSTRAT [26]). Thus, 1,870 case and 2,002 comparison samples were available for association analyses. The groups were well matched for the first two EIGENSTRAT principal component scores (data supplement Figure 3). The first score was correlated (r=0.87) with rs4988235 (lactase persistence) genotypes (data supplement Figure 4), presumably reflecting north-south European genetic variation (23).

### *Association Tests for Single SNPs*

Armitage trend tests of association were computed (PLINK) for each of 648 SNPs. For comparison, we also computed classical $\chi^2$ tests of allele counts in cases versus comparison subjects and association tests (EIGENSTRAT) that controlled for possible differential ancestry between cases and comparison subjects (the squared correlation between the ancestry-adjusted genotypes at the tested SNP and the ancestry-adjusted phenotypes under the null hypothesis of no association, following a $\chi^2$ distribution with 1° of freedom [26]). Empirical significance of single-SNP Armitage tests was estimated by permuting phenotype status to generate 100,000 data sets of all 648 SNPs under the null hypothesis of no association. Experiment-wide and gene-wide empirical significance were defined, respectively, as the probability of observing at least one SNP in the experiment or in its gene with an Armitage trend test at least as large as the observed one. Empirical significance tests are necessary here to correct for the correlations in association tests for SNPs, which are in linkage disequilibrium with each other (data supplement Figures 5–16); i.e., there are pairs of alleles at nearby SNPs that are usually found on the same chromosomes because of their evolutionary history.

### *Haplotypic Analyses*

We carried out tests of combinations of SNPs (haplotypes) (data supplement Tables 5–7), including haplotypes that had previously been reported as associated with schizophrenia and additional exploratory analyses with UNPHASED (27) to compute a global p value accounting for all possible haplotypes (with a frequency of 3% or greater) for each set of SNPs and PLINK to compute a p value and odds ratio for the most associated haplotype. Exploratory analyses included "sliding windows" of two and three SNPs and also an "anchored" stepwise procedure, starting with SNPs with nominal pointwise (p<0.05) association and then searching for two- and three-SNP combinations within each gene with greater evidence for association. Empirical p values were determined for these tests by permutation if the nominal global p value was less than $10^{-3}$.

### *Imputation of Nongenotyped HapMap SNPs*

As an additional exploratory analysis, genotypes were imputed for all ungenotyped HapMap II SNPs in candidate gene regions (if present in at least three HapMap CEPH Utah chromosomes) using MACH 1.0 (www.sph.umich.edu/csg/abecasis/MACH/). MACH uses Markov chain models to infer the probability of each possible genotype of an SNP in each subject based on a training data set (haplotyped CEPH Utah HapMap II data). Score tests were used to test allelic association with the sums of these probabilities in cases versus comparison subjects. Imputed data can suggest additional SNPs that merit genotyping for more precise association tests (28).

### *Power Analyses*

Data supplement Table 8 (all SNPs) and Table 9 (tag SNPs) show the power of this sample to detect experiment-wide empirical significance across a range of genetic models, assuming 1% disease prevalence and weak or strong linkage disequilibrium between the true susceptibility variant and the associated SNP. Data supplement Table 10 shows the minimum genotypic relative risk at which empirical gene-wide significance can be detected with 80% power for each gene. Genotypic relative risk is the increase in risk produced by carrying one risk allele for dominant or multiplicative models or two risk alleles for recessive models. The sample has excellent power to detect gene-wide significant association for genotypic relative risk values of 1.25–1.50 in the presence of strong linkage disequilibrium or 1.3–1.7 SNP with weak linkage disequilibrium, except for less common alleles with recessive effects (a well-known limitation of case-control studies).

## Results

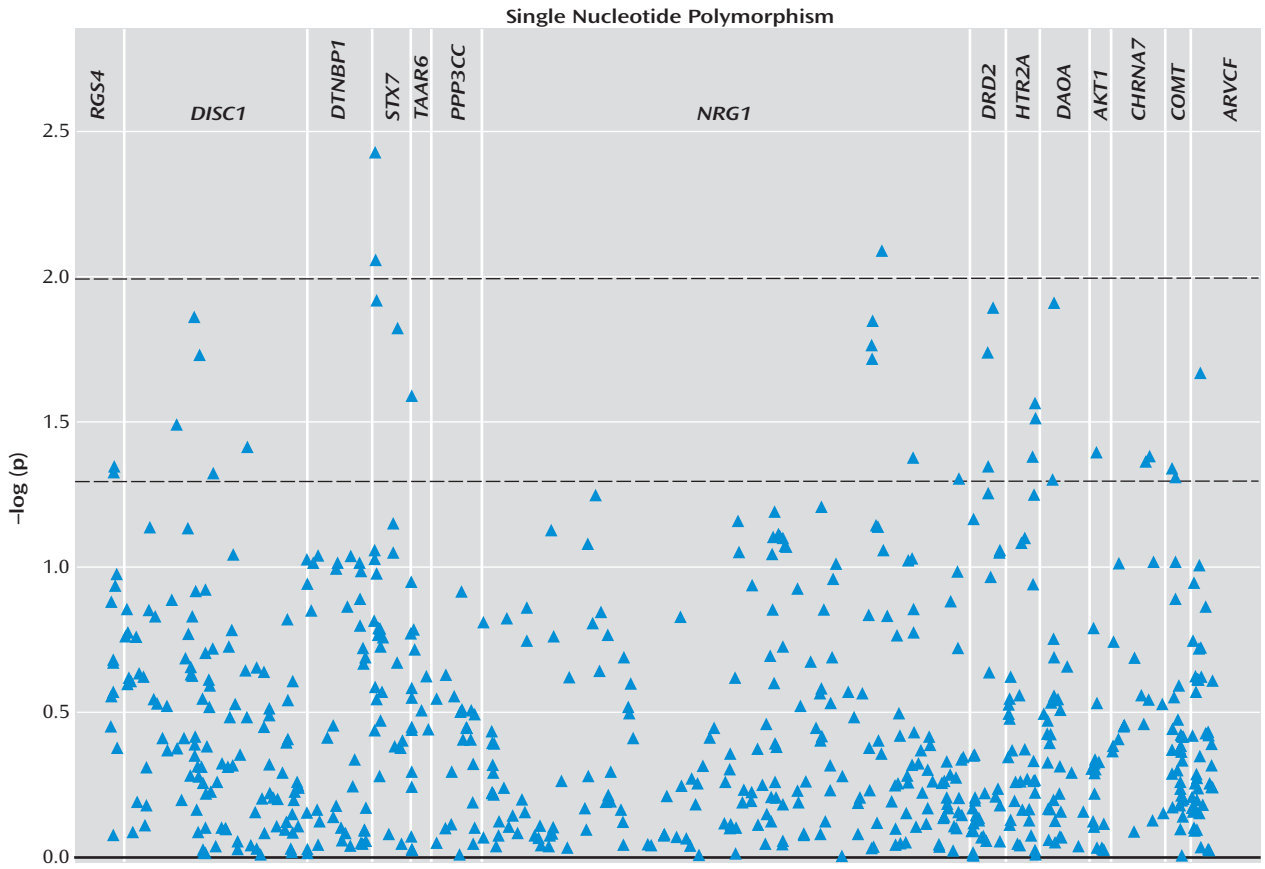
### *Association Analyses for Single SNPs*

The results are shown in Table 3 and Table 4, Figure 1, and data supplement Table 11. Pointwise empirical p<0.05 values (expected 5% of the time by chance) were observed in 4.6% of Armitage trend tests for all SNPs (30 of 648 tests) and 4.8% for tag SNPs (21 of 433 tests); and p<0.01 values (expected 1% of the time by chance) were observed in 0.5% of tests (three of 648 and two of 433, respectively), with the lowest value observed in tag SNP rs3183732 in *STX7* (empirical pointwise p=0.004). No SNP achieved empirical experiment-wide significance. Thresholds for the 5% significance level were nominal p<0.00008 for all SNPs or p<0.0002 if the analysis was limited to tag SNPs only; thresholds for "suggestive" association (i.e., expected once per experiment of this size) were p<0.002 or p<0.003, respectively. One SNP (rs3183732) in *STX7* showed gene-wide significance based only on tag SNPs (empirical gene-wide p<0.05) but not based on all SNPs (gene-wide p=0.051), a marginal result for an SNP for which association has not been previously reported. Exclusion of schizoaffective disorder cases did not alter these conclusions. Analyses using a correction for the possible effects of case-control ancestry differences (EIGENSTRAT) produced the same results. Tests of haplotypes and of imputed genotypes for additional HapMap common SNPs did not produce additional positive findings (data supplement Tables 5–7 and 12–13, data supplement Figure 17).

Figure 2 shows the quantile-quantile distribution of observed versus expected p values for tag SNPs. A straight line indicates good fit of a theoretically uniform distribution. There is a small departure below the null line, within the 95% confidence interval (perhaps reflecting modest linkage disequilibrium among the tag SNPs), consistent with a lack of evidence for association.

Our tests of the 70 SNPs, chosen because of previous positive reports of association with schizophrenia, can be viewed as a particularly interesting subset of tests (although not clearly as "replication" tests because previous evidence for association in these genes did not achieve a very strong threshold of significance). Of these, 69 (all except rs4262285 in *NRG1*) were successfully genotyped, and four (5.8%) produced p<0.05 values (Table 3), consistent with chance expectation. No nominally significant p values were observed for three functional polymorphisms that have been reported to be associated with schizophre-

**FIGURE 1. Association Results for Single Nucleotide Polymorphisms (SNPs) in Candidate Genes[a]**




[a] Shown are the results of Armitage trend tests of association for all 648 candidate gene SNPs as the $-\log_{10}$ of the pointwise nominal p value for each SNP (y axis) ordered along the x axis by analyzed SNP relative physical position within each gene region (Mb in hg17, i.e., National Center for Biotechnology Information Build 35), each gene region demarcated by a vertical line. The horizontal line at $-\log_{10}=1.3$ corresponds to a pointwise nominal p=0.05, and the horizontal line at $-\log_{10}=2.0$ to a pointwise nominal p=0.01.

nia: rs4680 (Val/Met, *COMT*), rs1801028 (Cys311Ser, *DRD2*), and rs1799732 (–141C Ins/Del, *DRD2*). Two functional SNPs in *HTR2A* that are in strong linkage disequilibrium (rs6313, T102C; rs6311, –1,438A/G) were nominally associated (p<0.03–0.04), but for the opposite rs6313 allele (T) than the previously reported association (C).

### *Comparison Sample*

Epidemiological and clinical characteristics of the comparison sample are described in data supplement Table 4. This comparison sample is currently used for the analysis of association with multiple psychiatric disorders by numerous research groups.

## Discussion

We did not detect a significant association of schizophrenia with SNPs in 14 candidate genes that have been of great interest to the field in a large sample of case and comparison subjects with closely comparable ancestry, studied with analysis of single SNPs, haplotypes, and imputed genotypes with a comprehensive map of common SNPs, additional SNPs with known or putative functional effects, and SNPs in these genes that had been previously reported as associated with schizophrenia.

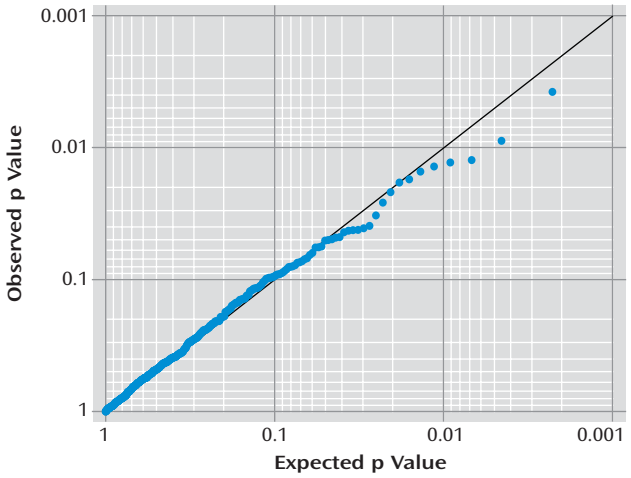
Our sample could possibly be in some way atypical, although we doubt that our findings can be explained in this way. It is possible, for example, that a sample limited to known familial cases would produce different results, given that pedigree-based linkage studies identified the regions in which many of these genes are located. However, most of the original and subsequent association reports have been in European ancestry case-control samples. Our cases were also likely to be clinically representative of other samples based on our collective experience in multicenter studies and our own data: we demonstrated high cross-site interrater reliability for schizophrenia and schizoaffective disorder diagnoses (kappas of 0.88 and 0.89) (11). Although we interviewed highly screened subjects (i.e., subjects with eligible clinical diagnoses were further screened by study clinicians), the best-estimate final diagnosis process still excluded approximately 10% of interviewed cases, indicat-

**TABLE 4. Single SNP Association Results for SNPs by Gene[a]**

| | Tested SNPs | | | Tag SNPs | | |
|---|---|---|---|---|---|---|
| Gene | N | p<0.05 | p<0.01 | N | p<0.05 | p<0.01 |
| *RGS4* | 12 | 2 | 0 | 6 | 1 | 0 |
| *DISC1* | 115 | 5 | 0 | 78 | 4 | 0 |
| *DTNBP1* | 38 | 0 | 0 | 22 | 0 | 0 |
| *STX7* | 27 | 4 | 2 | 15 | 3 | 2 |
| *TAAR6* | 17 | 1 | 0 | 15 | 1 | 0 |
| *PPP3CC* | 21 | 0 | 0 | 8 | 0 | 0 |
| *NRG1* | 217 | 5 | 1 | 159 | 2 | 0 |
| *DRD2* | 32 | 3 | 0 | 17 | 2 | 0 |
| *HTR2A* | 34 | 3 | 0 | 27 | 1 | 0 |
| *DAOA* | 30 | 2 | 0 | 19 | 2 | 0 |
| *AKT1* | 17 | 1 | 0 | 13 | 1 | 0 |
| *CHRNA7* | 18 | 2 | 0 | 15 | 2 | 0 |
| *COMT* | 29 | 1 | 0 | 19 | 1 | 0 |
| *ARVCF* | 41 | 1 | 0 | 20 | 1 | 0 |
| Total | 648 | 30 | 3 | 433 | 21 | 2 |

[a] Shown are the numbers of all tested SNPs and of tag SNPs with pointwise empirical Armitage trend test values of p<0.05 and p<0.01. Note that 4.6% of tested and 4.8% of tag SNPs had p<0.05, and 0.5% in each set had p<0.01, i.e., chance expectation.

**FIGURE 2. Quantile-Quantile Plot of Observed Versus Expected p Values for Tag Single Nucleotide Polymorphisms (SNPs)[a]**




[a] The blue dots represent the relationship between the expected (x axis) and observed (y axis) p values for pointwise nominal Armitage trend tests for the 433 SNPs that represent tags (at $r^2$>0.8) for common SNPs in each gene. The solid line represents the null expectation. The observed distribution is within the 95% confidence interval of the null expectation, consistent with a lack of evidence in our sample for association with schizophrenia in the tested candidate genes. The lowest p values are slightly below the line (less significant than expected) but still within the confidence interval.

ing a high priority on diagnostic accuracy. Excluding schizoaffective disorder cases from the analysis did not change the results. Another potential difference in our study compared to many others in the field is that we used psychiatrically screened comparison subjects, which led to excluding from genotyping ~8% of the comparison subjects who would have otherwise been in the experiment but were excluded for not denying one or more of three psychosis screens: treatment, diagnosis, or presence of 1) schizophrenia or schizoaffective disorder, 2) auditory hallucinations or delusions, or 3) bipolar disorder or manic depression. The largest single item contributing to the 8% was endorsement of previous treatment or diagnosis of bipolar disorder or manic depression, which was endorsed by 3.6% of the comparison subjects of European ancestry. The screening of comparison subjects can increase power by reducing the number of affected subjects in the comparison portion of the sample. However, if the large majority of those excluded from the comparison group based on a suspicion that they might have been affected were in reality unaffected, it can be argued that such screening would represent an overall loss of power owing to a larger effect of a smaller sample size in the comparison subjects. We do not know which might be the case in our sample but adopted the more conservative approach of using screened comparison subjects.

What do we learn from these results? First, we cannot definitively rule out a role for any of these genes in schizophrenia. Many of the odds ratios for association are in a plausible range (1.10–1.23) for small susceptibility effects but below what would produce significant p values in this sample or in the smaller samples used in previous studies. The larger odds ratios in some previous reports could either be false positives or inflated estimates of the genetic effects, as is common in initial reports—the so-called "winner's curse" (29). Also, only the hypothesis of association with common SNPs has been tested in a reasonably systematic way, both here and in the previous studies of these genes. We will learn more from future studies using resequencing methods to detect rare SNPs and genome-wide SNP arrays to detect genomic deletions and insertions as well as large-scale analyses of gene-gene interactions.

Second, the results demonstrate the importance of large-scale, systematic tests of genomic hypotheses. Although these candidate genes represent the best findings of the first generation of positional approaches to schizophrenia, evidence for each of them has been modest and/or inconsistent. Many of the initial associations were identified by the screening of candidate regions with what would now be considered small samples and inadequate coverage of common SNPs as well as (in some cases) older genotyping technologies that yield more missing data and higher error rates than current methods. Genome-wide association studies of large samples provide more powerful and systematic tests of common SNPs and of insertion/deletion variants throughout the genome and have already produced robust replicable association findings for other complex genetic phenotypes (10, 30–33). Multiple genome-wide association studies of schizophrenia are under way, including our study of the Molecular Genetics of Schizophrenia Parts 1 and 2 and Schizophrenia Genetics Initiative samples, the first phase of which is part of the Genetic Association Information Network (34). One caveat is that large-scale SNP arrays do not optimally cover every gene, so focused studies such as this one will still be

needed for genes whose role in schizophrenia is supported by candidate gene, linkage, genome-wide association, or biological studies. More systematic approaches to studying rare DNA sequence variants should also soon be available.

Our results suggest that, taken together, common DNA variants in these 14 genes are unlikely to explain a large proportion of the genetic risk for schizophrenia in populations of European ancestry. More robust findings are likely to be discovered using genome-wide association methods and, as our knowledge of the biology of mental illness continues to improve, focused studies of genes based on more precise mechanistic hypotheses. Nevertheless, although larger samples could possibly detect small genetic effects that were missed in this experiment, our findings suggest it is unlikely that true associations exist at the population level for the alleles that have formed the basis for the large candidate gene literature for these 14 postulated schizophrenia candidate genes.

Received Oct. 7, 2007; revision received Nov. 30, 2007; accepted Dec. 3, 2007 (doi: 10.1176/appi.ajp.2007.07101573). From the Center for Psychiatric Genetics, Department of Psychiatry and Behavioral Sciences, Evanston Northwestern Healthcare and Feinberg School of Medicine, Northwestern University, Evanston, Ill.; the Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Palo Alto, Calif.; the Departments of Psychiatry and Genetics, Washington University, St. Louis; the Queensland Centre for Mental Health Research and the Queensland Institute for Medical Research, Brisbane, Queensland, Australia; the Department of Psychiatry, University of Colorado Health Sciences Center, Denver; the School of Nursing, Louisiana State University Health Sciences Center, New Orleans; the Atlanta Veterans Affairs Medical Center and Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta; the Mental Health Clinical Research Center and Department of Psychiatry, University of Iowa College of Medicine, Iowa City, Iowa; the Department of Psychiatry, Mount Sinai School of Medicine, New York; the Department of Psychiatry, University of California, San Francisco; and the Département de Génétique, INSERM-U563, Institut National de la Recherche et de la Santé Médicale, Toulouse, France. Address correspondence and reprint requests to Dr. Gejman, Evanston Northwestern Healthcare Research Institute, 1001 University Place, Evanston, IL 60201-3137; pgejman@northwestern.edu (e-mail).

Data and biomaterials from the NIMH Schizophrenia Genetics Initiative, the Molecular Genetics of Schizophrenia Part 1, and the Molecular Genetics of Schizophrenia Part 2 case samples and the Molecular Genetics of Schizophrenia Part 2 comparison sample were collected by a number of investigators and institutions as detailed in data supplement Table 1, under the following grants: NIMH Schizophrenia Genetics Initiative U01s: MH46276, MH46289, and MH46318; and Molecular Genetics of Schizophrenia Part 1 and Part 2 R01s: MH67257, MH59588, MH59571, MH59565, MH59587, MH60870, MH59566, MH59586, MH61675, and MH60879. Research Career Development Awards (to Drs. Duan and Sanders) at the Evanston Northwestern Healthcare Research Institute, Evanston, Ill., supported this work.

The authors thank the study participants for their assistance and the individuals at each participating institution for their contributions, especially Heather Smith, Cheryl Filippich, Duncan McLean, Margaret Baier, Erich Conrad, Sherri Spera Chalona, Edmond Bedjeti, Roberta Fishman, Jihad M. Abdallah, Ilya M. Karagodin, Katherine Radwanski, Sandra Shi, Taylor J. Reif, Rick Li, Mike Dennis, John Corcoran, and Douglas A. Fugman. Genotyping was carried out at Evanston Northwestern Healthcare Research Institute.

Dr. Amin has received funds from Pfizer, Organon, and the Foundation for NIH. Dr. Black has received research support from Shire and Forest, has been on the speakers bureau of Pfizer, and has received consulting honoraria from Forest and Jazz. The remaining authors report no competing interests.

Drs. Sanders and Duan made equal contributions to this article.

## References

1. 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: Heritability estimates for psychotic disorders: the Maudsley twin psychosis series. Arch Gen Psychiatry 1999; 56:162–168
2. Gottesman II, Shields J: Schizophrenia: The Epigenetic Puzzle. Cambridge, UK, Cambridge University Press, 1982
3. Kendler KS, McGuire M, Gruenberg AM, Spellman M, O'Hare A, Walsh D: The Roscommon family study, II: the risk of nonschizophrenic nonaffective psychoses in relatives. Arch Gen Psychiatry 1993; 50:645–652
4. Lindsay EA, Morris MA, Gos A, Nestadt G, Wolyniec PS, Lasseter VK, Shprintzen R, Antonarakis SE, Baldini A, Pulver AE: Schizophrenia and chromosomal deletions within 22q11.2. Am J Hum Genet 1995; 56:1502–1503
5. Millar JK, Wilson-Annan JC, Anderson S, Christie S, Taylor MS, Semple CA, Devon RS, Clair DM, Muir WJ, Blackwood DH, Porteous DJ: Disruption of two novel genes by a translocation cosegregating with schizophrenia. Hum Mol Genet 2000; 9:1415–1423
6. Owen MJ, Williams NM, O'Donovan MC: The molecular genetics of schizophrenia: new findings promise new insights. Mol Psychiatry 2004; 9:14–27
7. Sullivan PF: The genetics of schizophrenia. PLoS Med 2005; 2: e212
8. IHMC: The International HapMap Consortium: a haplotype map of the human genome. Nature 2005; 437:1299–1320
9. Todd JA: Statistical false positive or true disease pathway? Nat Genet 2006; 38:731–733
10. WTCCC: Wellcome Trust Case Control Consortium: genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447:661–678
11. Suarez BK, Duan J, Sanders AR, Hinrichs AL, Jin CH, Hou C, Buccola NG, Hale N, Weilbaecher AN, Nertney DA, Olincy A, Green S, Schaffer AW, Smith CJ, Hannah DE, Rice JP, Cox NJ, Martinez M, Mowry BJ, Amin F, Silverman JM, Black DW, Byerley WF, Crowe RR, Freedman R, Cloninger CR, Levinson DF, Gejman PV: 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. Am J Hum Genet 2006; 78:315–333
12. 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: Genome-wide search for schizophrenia susceptibility loci: the NIMH Genetics Initiative and Millennium Consortium. Am J Med Genet 1998; 81:275–281
13. Faraone SV, Matise T, Svrakic D, Pepple J, Malaspina D, Suarez B, Hampe C, Zambuto CT, Schmitt K, Meyer J, Markel P, Lee H, Harkavy Friedman J, Kaufmann C, Cloninger CR, Tsuang MT: Genome scan of European-American schizophrenia pedigrees: results of the NIMH Genetics Initiative and Millennium Consortium. Am J Med Genet 1998; 81:290–295
14. Cavalli-Sforza LL, Menozzi P, Piazza A: The History and Geography of Human Genes. Princeton, NJ, Princeton University, 1994
15. Nurnberger JI Jr, Blehar MC, Kaufmann CA, York-Cooler C, Simpson SG, Harkavy-Friedman J, Severe JB, Malaspina D, Reich T: Diagnostic Interview for Genetic Studies: rationale, unique features, and training: NIMH Genetics Initiative. Arch Gen Psychiatry 1994; 51:849–859

16. Gershon ES, DeLisi LE, Hamovit J, Nurnberger JI Jr, Maxwell ME, Schreiber J, Dauphinais D, Dingman CW, Guroff JJ: A controlled family study of chronic psychoses: schizophrenia and schizoaffective disorder. Arch Gen Psychiatry 1988; 45:328–336
17. Leckman JF, Sholomskas D, Thompson WD, Belanger A, Weissman MM: Best estimate of lifetime psychiatric diagnosis: a methodological study. Arch Gen Psychiatry 1982; 39:879–883
18. Kessler RC, Andrews G, Mroczek D, Ustun B, Wittchen HU: The World Health Organization Composite International Diagnostic Interview Short Form (CIDI-SF). Int J Methods Psychiatr Res 1998; 7:171–185
19. Eysenck SBG, Eysenck HJ, Barrett P: A revised version of the Psychoticism Scale. Pers Individ Dif 1985; 6:21–29
20. Smith MW, Patterson N, Lautenberger JA, Truelove AL, McDonald GJ, Waliszewska A, Kessing BD, Malasky MJ, Scafe C, Le E, De Jager PL, Mignault AA, Yi Z, De The G, Essex M, Sankale JL, Moore JH, Poku K, Phair JP, Goedert JJ, Vlahov D, Williams SM, Tishkoff SA, Winkler CA, De La Vega FM, Woodage T, Sninsky JJ, Hafler DA, Altshuler D, Gilbert DA, O'Brien SJ, Reich D: A high-density admixture map for disease gene discovery in African Americans. Am J Hum Genet 2004; 74:1001–1013
21. Reiner AP, Ziv E, Lind DL, Nievergelt CM, Schork NJ, Cummings SR, Phong A, Burchard EG, Harris TB, Psaty BM, Kwok PY: Population structure, admixture, and aging-related phenotypes in African American adults: the Cardiovascular Health Study. Am J Hum Genet 2005; 76:463–477
22. Lao O, van Duijn K, Kersbergen P, de Knijff P, Kayser M: Proportioning whole-genome single-nucleotide-polymorphism diversity for the identification of geographic population structure and genetic ancestry. Am J Hum Genet 2006; 78:680–690
23. Campbell CD, Ogburn EL, Lunetta KL, Lyon HN, Freedman ML, Groop LC, Altshuler D, Ardlie KG, Hirschhorn JN: Demonstrating stratification in a European American population. Nat Genet 2005; 37:868–872
24. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81:559–575
25. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D: Efficiency and power in genetic association studies. Nat Genet 2005; 37:1217–1223
26. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006; 38: 904–909
27. Dudbridge F: Pedigree disequilibrium tests for multilocus haplotypes. Genet Epidemiol 2003; 25:115–121
28. Marchini J, Howie B, Myers S, McVean G, Donnelly P: A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007; 39:906–913
29. Ioannidis JP, Patsopoulos NA, Evangelou E: Heterogeneity in meta-analyses of genome-wide association investigations. PLoS ONE 2007; 2:e841
30. Gudmundsson J, Sulem P, Manolescu A, Amundadottir LT, Gudbjartsson D, Helgason A, Rafnar T, Bergthorsson JT, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Xu J, Blondal T, Kostic J, Sun J, Ghosh S, Stacey SN, Mouy M, Saemundsdottir J, Backman VM, Kristjansson K, Tres A, Partin AW, Albers-Akkers MT, Godino-Ivan Marcos J, Walsh PC, Swinkels DW, Navarrete S, Isaacs SD, Aben KK, Graif T, Cashy J, Ruiz-Echarri M, Wiley KE, Suarez BK, Witjes JA, Frigge M, Ober C, Jonsson E, Einarsson GV, Mayordomo JI, Kiemeney LA, Isaacs WB, Catalona WJ, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K: Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet 2007; 39:631–637
31. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S: Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007; 316:1331–1336
32. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC: A common allele on chromosome 9 associated with coronary heart disease. Science 2007; 316:1488–1491
33. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, Healey CS, Bowman R, Meyer KB, Haiman CA, Kolonel LK, Henderson BE, Le Marchand L, Brennan P, Sangrajrang S, Gaborieau V, Odefrey F, Shen CY, Wu PE, Wang HC, Eccles D, Evans DG, Peto J, Fletcher O, Johnson N, Seal S, Stratton MR, Rahman N, Chenevix-Trench G, Bojesen SE, Nordestgaard BG, Axelsson CK, Garcia-Closas M, Brinton L, Chanock S, Lissowska J, Peplonska B, Nevanlinna H, Fagerholm R, Eerola H, Kang D, Yoo KY, Noh DY, Ahn SH, Hunter DJ, Hankinson SE, Cox DG, Hall P, Wedren S, Liu J, Low YL, Bogdanova N, Schurmann P, Dork T, Tollenaar RA, Jacobi CE, Devilee P, Klijn JG, Sigurdson AJ, Doody MM, Alexander BH, Zhang J, Cox A, Brock IW, MacPherson G, Reed MW, Couch FJ, Goode EL, Olson JE, Meijers-Heijboer H, van den Ouweland A, Uitterlinden A, Rivadeneira F, Milne RL, Ribas G, Gonzalez-Neira A, Benitez J, Hopper JL, McCredie M, Southey M, Giles GG, Schroen C, Justenhoven C, Brauch H, Hamann U, Ko YD, Spurdle AB, Beesley J, Chen X, Mannermaa A, Kosma VM, Kataja V, Hartikainen J, Day NE, Cox DR, Ponder BA: Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 2007; 447:1087–1093
34. GAIN Collaborative Research Group, Manolio TA, Rodriguez LL, Brooks L, Abecasis G, Collaborative Association Study of Psoriasis, Ballinger D, Daly M, Donnelly P, Faraone SV, International Multi-Center ADHD Genetics Project, Frazer K, Gabriel S, Gejman P, Molecular Genetics of Schizophrenia Collaboration, Guttmacher A, Harris EL, Insel T, Kelsoe JR, Bipolar Genome Study, Lander E, McCowin N, Mailman MD, Nabel E, Ostell J, Pugh E, Sherry S, Sullivan PF, Major Depression Stage 1 Genomewide Association in Population-Based Samples Study, Thompson JF, Warram J, Genetics of Kidneys in Diabetes (GoKinD) Study, Wholley D, Milos PM, Collins FS: New models of collaboration in genome-wide association studies: the Genetic Association Information Network. Nat Genet 2007; 39:1045–1051