Am. J. Hum. Genet. 77:468–476, 2005

Report

Genetic Evidence for a Distinct Subtype of Schizophrenia Characterized by Pervasive Cognitive Deficit

Joachim F. Hallmayer[1,2,3] Luba Kalaydjieva,[4,5] Johanna Badcock,[2,3] Milan Dragović,[2,3] Sarah Howell,[2,3] Patricia T. Michie,[2,3,7] Daniel Rock,[2,3] David Vile,[2,3] Rachael Williams,[2,3] Elizabeth H. Corder,[8] Kate Hollingsworth,[6] and Assen Jablensky[2,3]

[1]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Palo Alto, CA; [2]Centre for Clinical Research in Neuropsychiatry, [3]School of Psychiatry and Clinical Neuroscience, [4]Western Australian Institute for Medical Research, and [5]Centre for Medical Research, The University of Western Australia, and [6]Neurodegenerative Disorders Centre, Queen Elizabeth II Medical Centre, Perth, Australia; [7]School of Behavioural Sciences, The University of Newcastle, Callaghan, Australia; and [8]Center for Demographic Studies, Duke University, Durham, NC

A novel phenotyping strategy in schizophrenia, targeting different neurocognitive domains, neurobehavioral features, and selected personality traits, has allowed us to identify a homogeneous familial subtype of the disease, characterized by pervasive neurocognitive deficit. Our genome scan data indicate that this subtype, which accounts for up to 50% of our sample, has a distinct genetic basis and explains linkage to chromosome 6p24 reported previously. If representative of other populations, the ratio of schizophrenia subtypes observed in our families could have a profound impact on sample heterogeneity and on the power of genetic studies to detect linkage and association. Our proposed abbreviated battery of tests should facilitate phenotype characterization for future genetic analyses and allow a focus on a crisply defined schizophrenia subtype, thus promoting a more informed search for susceptibility genes.

The argument about whether schizophrenia (MIM 181500) is a single disease or a collection of pathogenetically distinct subtypes goes back to the inception of the diagnostic concept at the turn of the 20th century. E. Kraepelin (1909) viewed *dementia praecox* as a cognitive disorder, sometimes accompanied by delusions, hallucinations, and excitement, but essentially characterized by "weakening of the mainsprings of volition," "lowered mental efficiency," "unsteadiness of attention," "inability to sift, arrange and correct ideas, and to accomplish mental grouping of ideas" (pp. 7–74). Coining the term "schizophrenia" to replace "*dementia praecox*," E. Bleuler (1920) emphasized that it "is not a disease in the strict sense, but appears to be a group of diseases…Therefore we should speak of schizophrenias in the plural" (p. 373).

Received April 4, 2005; accepted for publication June 16, 2005; electronically published July 12, 2005.

Address for correspondence and reprints: Dr. Assen Jablensky, School of Psychiatry and Clinical Neuroscience, University of Western Australia, MRF Building, 50 Murray Street, Perth WA 6000, Australia. E-mail: assen@cyllene.uwa.edu.au

0002-9297/2005/7703-0013$15.00

The inherent heterogeneity of the original concept has been obfuscated in modern diagnostic classifications (DSM-IV and ICD-10), which are designed to meet the needs of patient management, not fundamental research; which give primacy to subjective symptoms; and which may not target phenotypes anchored in the biology of the illness (Heinrichs 2004). Limited understanding of phenotypic heterogeneity is a common challenge in genetic studies of complex disorders and is a major contributor to the slow progress of such studies. The search for susceptibility genes in schizophrenia is particularly aggravated by the dual predicament of likely etiological diversity and a potentially fallible phenotype based on the diagnostic classifications of psychopathological phenomena.

In this study, we adopted, from the outset, the original concept (Bleuler 1920) of schizophrenia as an amalgam of several underlying, etiologically distinct disorders, and we argue that objective measures of brain dysfunction are likely to facilitate their delineation. Recent evidence has indicated that patients with schizophrenia exhibit abnormalities in multiple cognitive domains that predate the onset of the disorder (Kremen et al. 1998;

Bilder et al. 2000), persist across changes in the clinical state (Hoff et al. 1999), are not attributable to antipsychotic medications (Torrey 2002), occur in nonpsychotic relatives (Sitskoorn et al. 2004), and are specific to schizophrenia, as compared with other psychotic disorders (Altshuler et al. 2004), which thus meets the criteria for an endophenotype (Gottesman and Gould 2003). We reasoned that such abnormalities reflect patterns of deficit across multiple cognitive functions and domains; therefore, searching for distinct illness subtypes on the basis of a single or a few such traits is unlikely to produce coherent results, because of small individual effect sizes and the extraordinary variability that characterizes the cognitive performance of patients with schizophrenia (Heinrichs 2004). We adopted a novel phenotyping strategy, aiming to identify composite profiles of cognitive performance. We employed a battery of tests (Hallmayer et al. 2003), targeting different neurocognitive domains, neurobehavioral features, and selected personality traits.

The study included a total of 531 Western Australian subjects of European descent: 388 members of 112 families affected by schizophrenia and 143 population controls. The families were ascertained through affected probands by monitoring of consecutive admissions to a psychiatric hospital. Probands and first-degree relatives underwent a diagnostic assessment based on a structured interview with the use of Schedules for Clinical Assessment in Neuropsychiatry (SCAN), version 2.0 (Wing et al. 1990), a review of case records, and a structured developmental history obtained from a key family member. Full pedigree descriptions and family histories were collected using the National Institute for Mental Health Family Interview for Genetic Studies. Research diagnoses were established by consensus between two senior clinicians (blinded to family relatedness), who reviewed independently all diagnostic information, including the videotape of the SCAN interview, and assigned ICD-10 and DSM-IV lifetime diagnoses. The 388 family members (of whom 145 were parents and 131 were siblings of probands) included 138 affected individuals with schizophrenia or schizophrenia spectrum disorders (SZ/SZS), which include schizoaffective disorder, schizotypal personality disorder, and other nonaffective psychoses. The 143 control subjects (73 females and 70 males), recruited from a list of Red Cross blood donors or by random sampling from local telephone directories, were screened for psychopathology and were excluded if they or any of their first-degree relatives had been diagnosed with SZ/SZS or bipolar affective disorder. Written informed consent was obtained from all participants. The study complied with the ethics guidelines of the institutions involved.

All 531 participants were administered a battery of tests that assessed performance across seven domains of neurocognitive function, for which evidence of heritability (Cannon et al. 2000) and acceptable effect sizes of test measures (Heinrichs and Zakzanis 1998) have been reported. The domains and corresponding tests were as follows.

General cognitive ability: The National Adult Reading Test (NART), which estimates prior or premorbid verbal IQ, and the Shipley Institute of Living Scale, which estimates current IQ (by conversion to a Wechsler Adult Intelligence Scale R [WAIS-R] score).

Sustained attention: Two versions of the visual Continuous Performance Task—the degraded-stimulus version (CPT-DS), which involves an increased demand on visual encoding, and the identical-pairs version (CPT-IP), which selectively engages working memory.

Executive function: The FAS version of the Controlled Oral Word Association Task, which tests verbal fluency.

Verbal memory: The Rey Auditory Verbal Learning Test, which measures immediate and delayed recall of word lists, retention after distraction, and errors (i.e., intrusions of nonlist words).

Speed of information processing: The Inspection Time (IT) task, which provides a measure of perceptual encoding and speed of processing, unconfounded by motor-reaction time.

Neurobehavioral features: A structured examination that evaluated a range of soft neurological signs; in addition, The Edinburgh Handedness Inventory of hand/foot/eye preferences.

Personality factors: The Schizotypal Personality Questionnaire (SPQ) (Raine 1991) and the Temperament and Character Inventory (TCI) (Cloninger and Svrakić 1994), both of which were completed by all participants.

We used grade of membership (GoM) analysis (Woodbury et al. 1978; Manton et al. 1994) to analyze the test results, with control individuals providing the baseline data. GoM is a form of latent structure analysis, directed at defining a parsimonious number of latent groups or patterns of responses (representing, e.g., biological processes or phenotypes) from complex data sets, and allowing individuals to resemble each group to varying degrees (rather than classifying them into mutually exclusive clusters, as done in standard latent class analysis). These properties enable GoM to account for individual heterogeneity under conditions of high dimensionality better than alternative methods do, as is the case with the multiple continuous cognitive traits we measured for schizophrenia in a modestly sized family sample. Two sets of parameters are estimated by maximum likelihood (Woodbury et al. 1978; Manton et al. 1994). One set represents the latent pure-type groups by the probabilities (λ_{kjl}) that a subject matching the *k*th group will exhibit the *i*th response for the *j*th variable. The other set represents the resemblance of individuals to the

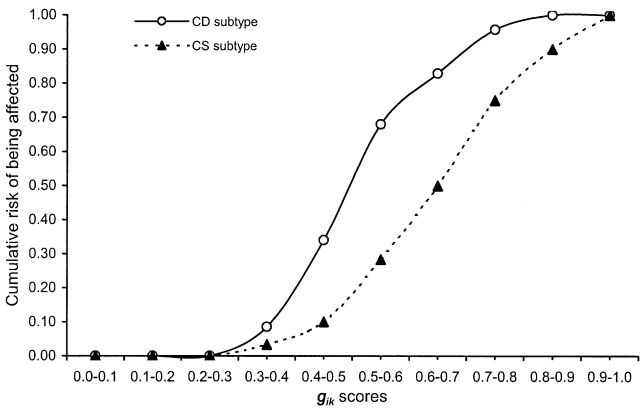

Figure 1 Association of the CD and CS traits with the clinical phenotype (SZ/SZS) at different cut-off values of the GoM (g_{ik}) scores. The cumulative curves are based on all 76 subjects (47 affected and 29 clinically unaffected) assigned to the CD subtype and all 86 subjects (60 affected and 26 clinically unaffected) assigned to the CS subtype, within the sample of 112 families. The CD trait shows a stronger association with clinical affection status than does the CS trait, over the entire g_{ik} range above the cut-off at 0.323. Of the affected subjects assigned to the CD trait, 68% had g_{ik} scores in the range 0.3–0.6 and 96% had g_{ik} scores of 0.3–0.8. In comparison, 28% of the affected subjects assigned to the CS trait had g_{ik} scores of 0.3–0.6 and 75% and had g_{ik} scores of 0.3–0.8. The highest observed g_{ik} score was 0.884 for the CD trait and 1.000 for the CS trait.

groups by GoM weights (g_{ik}), which range from 0 to 1, are constrained to sum to 1 over the groups for each subject, and measure the degree to which the *i*th subject resembles the *k*th pure type. The probability that the *i*th subject has the *l*th level of the *j*th variable is defined by a binary variable (i.e., $y_{ijl} = 0$ or 1). As a taxometric procedure, GoM differs from other clustering or factor analytical methods in that the g_{ik} scores retain full information on individuals and account for all individual heterogeneity in the data. By assigning a single g_{ik} score to an individual's GoM in each multivariate pure type, the analysis avoids the problem of multiple testing and the concomitant loss of power due to an increasing number of degrees of freedom. GoM is independent of distribution assumptions and is relatively robust to missing data. In genetic analysis, it combines multivariate continuous trait variables with different effect sizes or heritabilities, resulting in increased power (Kaabi and Elston 2003).

We used the neurocognitive test results, neurobehavioral features, and personality traits as internal variables to estimate the model parameters. Disease status, position within the pedigree, and demographic information were used descriptively as external variables to provide estimates of their frequencies for each pure type, conditioned on the model parameters. The data in this study were analyzed using the beta 1.01 version of the DSIGoM software (Decision Systems).

Optimal partitioning of the data (by a maximum-likelihood criterion) was obtained with four pure types. Two of these types, referred to as "cognitive deficit" (CD) and "cognitively spared" (CS), displayed markedly contrasting profiles (table 1), with probabilities of 80.6% (CD) and 70.0% (CS) of being expressed in subjects with SZ/SZS and in a proportion of their unaffected relatives. The CD pure type was characterized by a high probability of poor performance on the majority of cognitive tasks, an increased prevalence of nonlocalizing ("soft") neurological signs, and non–right-handedness (table 1). The CS pure type exhibited high scores for psychometric schizotypy and for traits associated with psychosis proneness. The remaining two pure types (cognitively intact and cognitively preserved) had little chance (5% and 7%, respectively) of expression in affected subjects and were represented mainly by well-functioning siblings, unaffected parents (some with mild, age-related cognitive nonoptimality), and controls.

We used the individual-level GoM coefficients (i.e., g_{ik} scores, which, for each subject, ranged from 0 to 1 for each pure type and summed to 1 over the four pure types) to classify the probands on the basis of the g_{ik} score indicating their greatest affinity to one of the pure types: predominantly CD (41 subjects), CS (51 subjects), or non-CD/CS (20 subjects with low scores for both CD and CS pure types). On the basis of the proband's predominant phenotype assignment, we classified the families into CD, CS, and non-CD/CS groups and examined the familial aggregation of the g_{ik} scores as continuous variables. Intrafamilial aggregation was significant (analysis of variance [ANOVA] $F = 1.67$; $P < .001$) in the families with the proband assigned to the CD group but not in the families with the proband assigned to the CS or non-CD/CS group (table 2). Next, we analyzed the association of the g_{ik} scores for the CD and CS traits, as continuous variables, with clinical illness (SZ/SZS) in family members. The results (fig. 1) indicated an association of the CD trait with SZ/SZS at all g_{ik} cut-off points ≥ 0.3, supporting cosegregation of the CD trait with the clinical phenotype. Finally, to estimate the relative risk ratio (λ_r), we compared the prevalence of the CD and CS types among the first-degree relatives of probands and among controls. For the CD type, the prevalence was 18.9% among relatives and 2.0% among controls, yielding $\lambda_r = 9.5$. By contrast, the CS prevalence was 16.8% among relatives and 14.0% among controls, yielding $\lambda_r = 1.2$ for this type.

The identification of generalized pervasive CD as an endophenotype in some, but not all, subjects SZ/SZS and their families suggested that it might characterize a genetically distinct schizophrenia subtype. To test this hypothesis, we conducted a whole-genome scan, with 380 microsatellites from the LMS1-HD5 marker set (Applied Biosystems), providing an average intermarker distance

Table 1

Pure-Type Profiles Identified by GoM Analysis (112 Families; 388 Individuals)

Variables[a]	*H*	λ_k Probabilities for Pure Type[b]			
		CD	CS	Cognitively Intact	Cognitively Preserved[c]
Neurocognitive domains:					
General ability:					
Estimated premorbid IQ (≤95)	.73	***84.7***	.0	.0	.0
Current IQ (≤91)	.87	***84.6***	.0	.0	.0
Verbal memory:					
Immediate word recall (≤21)	.92	***100.0***	.0	.0	.0
Delayed word recall (≤5)	.87	***71.5***	.0	.0	.0
Lexical retrieval:					
Verbal fluency, words (≤26)	.64	***85.2***	.0	.0	.0
Sustained attention:					
CPT-DS (dL ≤ 4.3)	.56	***91.8***	.0	.0	.0
CPT-IP (dL ≤ 2.2)	.59	***85.0***	.0	.0	.0
Speed of information processing:					
Inspection time (≥36 s)	.35	***66.5***	40.5	.0	72.3
Neurobehavioral features:					
Neurological soft signs (≥3)	.30	***100.0***	10.9	12.1	14.0
Non–right-handed (LQ ≤ 45)	.12	22.4	3.4	8.3	8.3
Personality factors:					
Schizotypal symptoms[d] (≥23)	.95	.0	***100.0***	.0	.0
Harm avoidance[e] (≥19)	.50	.0	***90.1***	.0	.0
Self-transcendence[f] (≥19)	.62	28.4	***79.8***	.0	.0
Self-directedness[g] (≥37)	.66	.0	.0	81.7	65.9
External variables:					
Proband	…	***72.8***	***67.8***	.0	.0
First-degree relative	…	27.2	32.9	100.0	100.0
Affected (SZ/SZS)	…	80.6	70.0	4.8	6.8
Other diagnosis[h]	…	9.5	23.3	45.6	28.0
Female	…	37.5	22.1	79.4	37.2
Age ≥45 years	…	45.9	10.0	19.7	**96.3**

NOTE.—*H* is an information-content coefficient, estimating the contribution of each variable to the final maximum-likelihood solution ($H \geq 0.10$ indicates significant contribution). Important data patterns are shown in bold italics.

[a] Measures of neurocognitive domains, neurobehavioral features, and personality factors were used as internal variables defining the pure types (the cut-off values for each test were the upper and lower bounds of the extreme quantile of the distribution of test scores). The external variables of the pure types are similarly described by λ_k probabilities but were not used in the identification of the pure types. dL = signal-detection index; LQ = lateralization quotient (by Edinburgh Handedness Inventory).

[b] Pure types are extreme profiles, partitioning the data in accordance with a maximum-likelihood criterion, and are described by percent probabilities (λ_k) that an individual expressing completely a given pure type will exhibit the characteristic.

[c] Age-related.

[d] Ideas of reference, odd beliefs, suspiciousness, unusual perceptions, constricted affect, social anxiety, and eccentric behavior (by SPQ).

[e] Anticipatory worry and fear of uncertainty (by TCI).

[f] Unusual spiritual experiences and beliefs (by TCI).

[g] Purposefulness and resourcefulness (by TCI).

[h] Current or lifetime-ever affective, anxiety, other neurotic, personality, and substance use disorders.

of 9.81 cM (see table 3 for a list of markers and genetic distances). For the linkage analysis, we used an integrated genetic map (produced by David Duffy at the Queensland Institute for Medical Research; see David Duffy's QIMR Homepage), which contains interpolated genetic map positions estimated via locally weighted linear regression from the physical map (National Center for Biotechnology Information [NCBI] build 34.3) and the published deCODE and Marshfield maps. The genome scan was performed using 93 of the families (34

Table 2

Characteristics of Families and Family Members Assigned to Cognitive Phenotype Groups

Variable	Finding in Phenotype Group[a]		
	CD	CS	Non-CD/CS[b]
Family characteristics[c]:			
No. of families/probands	41 (34)	51 (42)	20 (17)
Total no. of members	147 (127)	174 (155)	67 (60)
No. affected with SZ/SZS	54 (45)	59 (50)	25 (22)
No. affected who have CD phenotype	47 (38)	0 (0)	0 (0)
No. unaffected who have CD phenotype[d]	15 (11)	11 (10)	3 (2)
Total no. who have CD phenotype	62 (49)	11 (10)	3 (2)
Familial aggregation of g_{ik} scores[e]:			
Square of g_{ik}, between families[f]	.067	.085	.088
Square of g_{ik}, within families[f]	.040	.085	.073
F (significance)	1.666 (P = .000)	.996 (NS)	1.202 (NS)
g_{ik} For the CD phenotype[f]	.554 ± .143	.243 ± .192	.087 ± .108
g_{ik} For the CS phenotype[f]	.159 ± .154	.701 ± .168	.059 ± .097
Neurocognitive performance[g]:			
Estimated premorbid IQ[f]	92.6 ± 9.4	101.4 ± 9.8	105.6 ± 8.9
Current IQ[f]	84.8 ± 12.2	97.4 ± 11.3	107.4 ± 10.0
Immediate word recall (no. of words)[f]	17.6 ± 5.6	23.8 ± 5.2	27.2 ± 5.7
Delayed word recall (no. of words)[f]	4.4 ± 2.6	7.1 ± 2.8	9.0 ± 3.0
Verbal fluency (no. of words)[f]	25.0 ± 8.8	32.1 ± 9.6	38.2 ± 11.3
CPT-DS (dL)[h]	3.62 (1.4)	5.15 (1.2)	5.48 (1.2)
CPT-IP (dL)[h]	2.04 (1.3)	3.75 (1.7)	4.2 (1.4)
Inspection time (s)[f]	46.4 ± 23.6	35.8 ± 12.8	36.2 ± 12.3

Note.—After the identification of neurocognitive pure types, each subject was classified as phenotypically CD, CS, or non-CD/CS on the basis of the GoM (g_{ik}) score showing the greatest positive deviation from the mean g_{ik} for any of the four pure types (i.e., a subject's g_{ik} for each pure type was divided by the mean g_{ik} for that pure type).

[a] Families were indexed as CD, CS, or non-CD/CS in accordance with the proband's cognitive phenotype.

[b] Includes the cognitively intact and cognitively preserved pure types from table 1.

[c] The values in parentheses are the numbers included in the genetic analysis.

[d] Considered to be affected in the genetic analyses.

[e] Familial aggregation of the g_{ik} scores as continuous traits (examined by one-way ANOVA) was highly significant for the CD trait but failed to reach significance for the CS and non-CD/CS phenotypes. NS = not significant.

[f] Values are mean ± SD.

[g] Members displaying each group's index phenotype were compared across the family groups with regard to cognitive performance.

[h] dL = signal-detection index.

of which had been assigned to the CD subtype), with a total of 342 fully characterized members, including 140 classified as "affected"—117 with the diagnosis of SZ/SZS and 23 nonpsychotic relatives with the CD phenotype. Linkage analysis was performed with the program package GENEHUNTER-PLUS (Kong and Cox 1997). We used diagnosis and neurocognitive profile as a bivariate phenotype. For the nonparametric linkage analysis, all individuals with a SZ/SZS diagnosis, as well as clinically unaffected relatives classified by GoM as either CD or CS subtype, were considered to be affected. Multipoint nonparametric LOD (NPL) scores were calculated with the Spairs statistic and a linear model (fig. 2, left panels). The data were assessed further using ordered subset analysis (OSA) (Hauser et al. 2004), in which the proband's g_{ik} score for subtype CD was used as a continuous trait covariate to rank the families in descending order from highest to lowest CD score (fig. 2, right panels). OSA identifies the rank-ordered subset of families that provide maximal evidence for linkage (and no a priori specification of the subset is required). The statistical significance of the change in the LOD score between the subset and the overall sample (ΔLOD) is evaluated by a permutation test as described elsewhere (Hauser et al. 2004).

The greatest increase in the LOD score detected by

Table 3

Markers Typed in the Genome Scan

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

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

Figure 2 NPL analysis and OSA of chromosomes 1–22

OSA (table 4) was in the 6p24 region, where a maximum LOD of 3.07 and ΔLOD of 1.546 ($P < .01$) were obtained at marker *D6S309* (20.25 cM). The result was accounted for by the families ranked from 1st to 47th by the proband's g_{ik} score for CD (range 0.884–0.306). In addition to the 34 families classified by GoM as subtype CD, the 47 OSA-positive families included 11 families assigned to the CS subtype and 2 families assigned to the non-CD/CS subtype (fig. 3*A*). Thus, the GoM endophenotype grouping done prior to analysis, for which a stringent cut-off value for CD was applied, proved to be significantly predictive of the observed linkage to 6p, correctly classifying 72.3% of the OSA-positive families and 100% of the OSA-negative families (86.0% correct classification overall).

To corroborate further the prediction that the CD subtype is genetically distinct, we performed parametric multipoint linkage analysis, assuming a CD-susceptibility allele of dominant inheritance. The gene frequency was set at 0.02, with penetrance values of 0.5 for heterozygote and homozygote disease-allele carriers and 0.001 for noncarriers. In these calculations, individuals with no SZ/SZS diagnosis and a CD score of 0 were considered to be unaffected. We genotyped 15 additional microsatellite markers across the 6p25-22 region (average distance 1.72 cM) and analyzed the entire chromosome 6 (with 52 markers) separately for the 47 OSA-positive families and the remaining 46 families. This analysis revealed clear differences between the two sets of families (fig. 3*B*). The combined maximum LOD score for the 47 families in the CD group increased to 3.32, again at marker *D6S309*. By contrast, for the remaining families, linkage was excluded for the entire region, and the LOD score at *D6S309* was −2.12.

Our linkage findings coincide precisely with the strongest signal on 6p25-p22 reported previously (Straub et al. 1995) in a large sample of high-density Irish families with schizophrenia. A follow-up study of the same sam-

Table 4

Results of OSA

Chromosome	cM	Nearest Marker	Maximum LOD[a]	ΔLOD	*P*	Linked Families (%)
1	251.7	*D1S2800*	.558	.558	.92	11
2	145.8	*D2S112*	2.458	1.226	.08	78
3	88.7	*D3S1285*	1.017	.858	.47	10
4	74.3	*D4S1592*	.983	.459	.64	45
5	78.4	*D5S647*	2.640	.000	.75	NA
6	20.25	***D6S309***	***3.073***	***1.546***	***.01***	51
7	54.9	*D7S484*	.921	.000	.92	NA
8	7.4	*D8S264*	1.615	.900	.10	29
9	17.3	*D9S286*	1.371	1.107	.16	33
10	103.8	*D10S1686*	1.631	1.319	.09	33
11	130.6	*D11S925*	1.897	1.160	.21	33
12	112.3	*D12S346*	1.631	1.448	.13	29
13	108.9	*D13S173*	.701	.684	.61	33
14	104.3	*D14S65*	1.211	.964	.25	12
15	46.0	*D15S978*	2.071	1.815	.13	37
16	30.8	*D16S3075*	1.562	.888	.12	37
17	136.5	*D17S928*	2.279	.440	.25	40
18	117.8	*D18S462*	1.554	1.422	.12	33
19	10.6	*D19S209*	1.797	.000	.73	100
20	33.8	*D20S186*	1.088	.946	.49	34
21	11.6	*D21S1256*	.799	.000	.86	NA
22	55.3	*D22S423*	.738	.698	.36	9

NOTE.—Maximum LOD scores and ΔLOD from baseline for each chromosome, calculated by OSA in family subgroups identified by ordering the family LOD scores from highest to lowest, on the basis of the GoM coefficients (g_{ik} scores) for the CD type. The baseline LOD score is defined as the LOD score summed over all families, and it is the difference between the maximum LOD and ΔLOD. NA = none of the data sets produced a maximum LOD score greater than the baseline. Key results are shown in bold italics.

[a] Denotes the maximum LOD score of all ordered subsets of families.

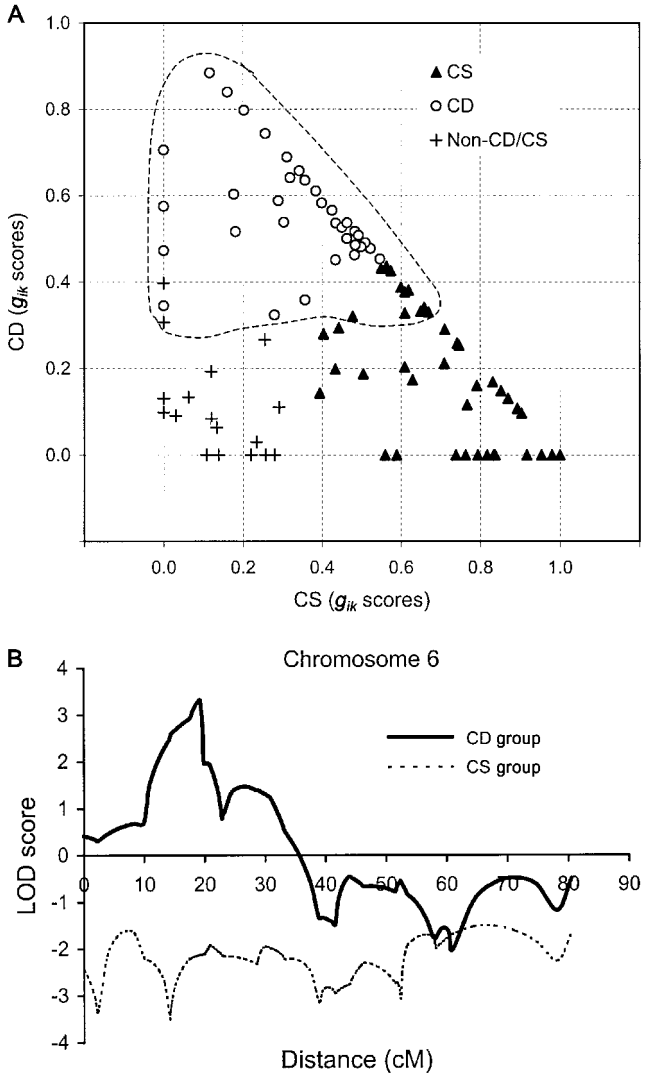

Figure 3 Genetic findings for chromosome 6p in 93 fully characterized families with schizophrenia. *A*, Classification of the families as linked or not linked to the 6p24 locus by OSA of the genome scan data, with use of the proband's g_{ik} score for the CD type as a continuous trait covariate. The positive linkage signal was accounted for by the 47 families (*encircled by the dashed line*) with the highest g_{ik} scores for the CD subtype. *B*, Linkage curves for chromosome 6p produced by parametric linkage analysis, conducted separately for the CD-subtype and CS-subtype families. Fine mapping of the 6p25-22 region was done with the addition of 15 microsatellite markers, resulting in an average intermarker distance of 1.72 cM. The CD-subtype families had a LOD score of 3.32 at marker *D6S309*. At that position, non–CD-subtype families (including the CS and non-CS subtypes) had a LOD score of −2.12. In the non–CD-subtype families, linkage was excluded for the entire 6p region.

ple suggested that the region might, in fact, contain two schizophrenia-susceptibility genes (Straub et al. 2002*b*). After the reported association of SNPs in the *DTNBP1* gene on 6p22 (Straub et al. 2002*a*), further studies have focused on replication and additional association analyses (reviewed by Owen et al. [2004]), leaving the 6p25-24 region unexplored. Our results, combined with previous findings (Straub et al. 1995, 2002*b*; Bailer et al. 2000; Hwu et al. 2000), lead us to suggest that it is the CD schizophrenia subtype that accounts for the linkage of schizophrenia to chromosome 6p25-24 and that the region contains a novel susceptibility gene of relatively strong effect.

A post hoc comparison between the OSA-positive families in the CD group and the OSA-negative families in the CS and non-CD/CS groups (fig. 4) indicated that SZ/SZS-affected subjects did not differ significantly with regard to age at onset or to duration or severity of clinical illness, which argues against the idea that CD and CS subtypes are different stages of a single disease process. However, the comparison did reveal differences in terms of diagnosed psychopathology; "core" schizophrenia was more common among the affected members of the OSA-positive families, with a ratio of core schizophrenia cases to schizophrenia spectrum disorder cases (SZ:SZS) of 2.9 in the OSA-positive families, compared with an SZ:SZS ratio of 2.1 in the OSA-negative families. Further, a greater percentage (12%) of first-degree relatives in the OSA-negative families had experienced episodes of bipolar or recurrent depressive disorder, compared with the percentage of relatives in the OSA-positive families (6%), suggesting heterogeneity and susceptibility to both schizophrenia and mood disorders in the CS-subtype families. As expected, the most salient differences between the two groups were related to cognitive function (fig. 4), with the CD-subtype families showing lower scores for general cognitive ability, memory, executive function, sustained attention, and speed of information processing. Post hoc logistic regression analysis (by Wald's backward stepwise method) identified verbal memory deficits (delayed word recall), poor sustained attention (by CPT-DS), and estimated current and previous IQ (by NART), as the four neurocognitive parameters (all significant at $P < .01$) that together classified correctly 91.8% of probands (combined sensitivity 91.2%; specificity 92.3%) as subtype CD or CS, suggesting that, in future genetic studies, efficient endophenotype ascertainment may be possible with an abbreviated neurocognitive-screening battery.

Having identified a cognitive subtype of schizophrenia linked to the 6p25-22 region, we acknowledge the limitation of the modest sample size in the present study as a constraint on our effort to have a more comprehensive characterization of a range of possible neurocognitive endophenotypes in this disorder. The apparent lack of familiality in the CS subtype is attributable to further heterogeneity and insufficient power to resolve familial aggregation, as a result of smaller effect sizes of the personality-trait measures defining this endophenotype (in contrast to relatively large effect sizes for the neurocog-

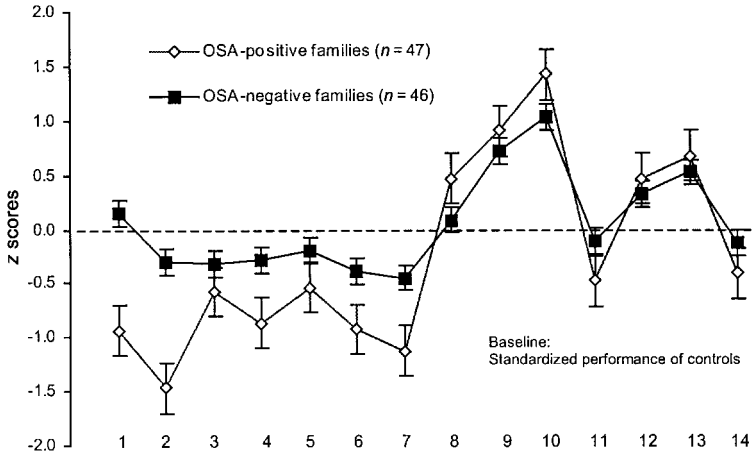

Figure 4 Post hoc phenotypic comparisons of the families included in the genome scan and OSA and averaged test performance profiles of the members (n = 342) of the 93 families included in the genome scan and OSA. The differences (by t test) between CD-subtype and non–CD-subtype (including the CS and non-CS subtypes) families were significant for previous or premorbid IQ, verbal memory, and sustained attention ($P < .001$); verbal fluency ($P < .01$); speed of information processing ($P < .05$); and non–right-handedness ($P < .05$). Neurocognitive domains: 1 = previous or premorbid IQ; 2 = current IQ; 3 = verbal memory, immediate recall; 4 = verbal memory, delayed recall; 5 = verbal fluency; 6 = sustained attention CPT-IP/dL; 7 = sustained attention CPT-DS/dL; 8 = slow information processing (by IT task). Neurobehavioral features: 9 = neurological soft signs; 10 = handedness (lateralization quotient). Personality traits: 11 = schizotypal symptoms; 12 = harm avoidance; 13 = self-transcendence; 14 = self-directedness.

nitive markers identifying the familial CD endophenotype). Future studies should investigate the cognitive architecture of the CS group in greater detail. This group is characterized by more complex symptomatology (systematized delusions, abnormalities of high-order reasoning, etc.) than is the CD group, reflecting a plausible distinction between the two groups in terms of pervasively reduced cognitive efficiency in CD-subtype individuals versus dysfunctional cognitive control in the CS-subtype individuals (a similar distinction has been proposed by Brebion et al. [2005]). Although powerful enough for reliable assessment of reduced cognitive efficiency, our present battery of tests may not be targeting the essential features of a putative CS endophenotype likely to display abnormalities in complex functions, such as self-monitoring, source attribution, and cognitive inhibition, as well as in their modulation by external (environmental) and internal (affective) background factors.

Notwithstanding such caveats, our data bring back to center stage the critical issue of phenotypic and etiological diversity in schizophrenia. The neurocognitive tests that we used to dissect the disease phenotype outlined a relatively homogeneous, familial, genetically distinct subtype closely corresponding to the *dementia praecox* described by Kraepelin (1909). In our sample, this subtype accounted for ~30% of families on the basis of the more stringent g_{ik} cut-off value in our GoM analysis and for up to ~50% of families on the basis of the OSA of the linkage data. If this sample is representative of other European populations, the observed prevalence of disease subtypes would have serious implications for the power of genetic studies—traditionally based on broad DSM-IV and ICD-10 diagnostic categories—to detect linkage and association. Parsing the phenotypic complexity of schizophrenia may open the way to a more informed search for specific pathogenetic pathways and underlying genetic mechanisms. Neurocognitive profiling with the proposed abbreviated battery, which comprises the most informative tests, should facilitate research, focusing on a more crisply defined, familial subtype of schizophrenia.

Acknowledgments

This work was supported by grants from the National Health and Medical Research Council, Australia (to A.J., J.F.H., and P.T.M.), and the North Metropolitan Health Services, Perth, Western Australia. The Neurodegenerative Disorders Centre genotyping facility, where fine mapping of the chromosome 6p was performed, is supported by GlaxoSmithKline. We thank V. Morgan for assistance with the preparation of the manuscript. The following individuals made specific contributions to earlier stages of this project: J. Box, T. W. Budd, J. E. Cooper, P. Davis, M. D'Ercole, J. Johnston, A. Kent, L. Kløve, R. Stienstra, J. Todd, H. Wichmann, D. Wood, and M. Woodbury. Staff of the Graylands Hospital and the mental health clinics within the North Metropolitan Health Services, Perth, provided patient referrals. We especially thank the patients, family members, and other volunteers who participated in this study.

Web Resources

The URLs for data presented herein are as follows:

Decision Systems, http://www.dsisoft.com/ (for DSIGoM software and analysis)
David Duffy's QIMR Homepage, http://www.qimr.edu.au/davidD/ (for integrated genetic map)
NCBI, http://www.ncbi.nlm.nih.gov
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for schizophrenia)

References

Altshuler LL, Ventura J, van Gorp WG, Green MF, Theberge DC, Mintz J (2004) Neurocognitive function in clinically stable men with bipolar I disorder or schizophrenia and normal control subjects. Biol Psychiatry 56:560–569

Bailer U, Leisch F, Meszaros K, Lenzinger E, Willinger U, Strobl R, Gebhardt C, Gerhard E, Fuchs K, Sieghart W, Kasper S, Hornik K, Aschauer NH (2000) Genome scan for susceptibility loci for schizophrenia. Neuropsychobiology 42:175–182

Bilder RM, Goldman RS, Robinson D, Reiter G, Bell L, Bates JA, Pappadopulos E, Wislon DF, Alvir JM, Woerner MG, Geisler S, Kane JM, Lieberman JA (2000) Neuropsychology of first-episode schizophrenia: initial characterization and clinical correlates. Am J Psychiatry 157:549–559

Bleuler E (1920) Lehrbuch der Psychiatrie. Springer, Berlin [reprinted English translation (1976) Textbook of psychiatry. Arno Press, New York]

Brebion G, Gorman JM, Malaspina D, Amador X (2005) A model of verbal memory impairments in schizophrenia: two systems and their associations with underlying cognitive processes and clinical symptoms. Psychol Med 35:133–142

Cannon TD, Huttunen MO, Lonnqvist J, Tuulio-Henriksson A, Pirkola T, Glahn D, Finkelstein J, Hietanen M, Kaprio J, Koskenvuo M (2000) The inheritance of neuropsychological dysfunction in twins discordant for schizophrenia. Am J Hum Genet 67:369–382

Cloninger CR, Svrakić D (1994) The Temperament and Character Inventory (TCI): a guide to its development and use. Center for Psychobiology of Personality, Washington University, St. Louis

Gottesman II, Gould TD (2003) The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry 160:636–645

Hallmayer JF, Jablensky A, Michie P, Woodbury M, Salmon B, Combrinck J, Wichmann H, Rock D, D'Ercole M, Howell S, Dragović M, Kent A (2003) Linkage analysis of candidate regions using a composite neurocognitive phenotype correlated with schizophrenia. Mol Psychiatry 8:511–523

Hauser ER, Watanabe RM, Duren WL, Bass MP, Langefeld CD, Boehnke M (2004) Ordered subset analysis in genetic linkage mapping of complex traits. Genet Epidemiol 27:53–63

Heinrichs RW (2004) Meta-analysis and the science of schizophrenia: variant evidence or evidence of variants? Neurosci Biobehav Rev 28:379–394

Heinrichs RW, Zakzanis KK (1998) Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. Neuropsychology 12:426–445

Hoff AL, Sakuma M, Wieneke M, Horon R, Kushner M, DeLisi LE (1999) Longitudinal neuropsychological follow-up study of patients with first-episode schizophrenia. Am J Psychiatry 156:1336–1341

Hwu HG, Lin M-W, Lee P-C, Lee SF-C, Ou-Wang W-C, Liu C-M (2000) Evaluation of linkage of markers on chromosome 6p with schizophrenia in Taiwanese families. Am J Med Genet 96:74–78

Kaabi B, Elston RC (2003) New multivariate test for linkage, with application to pleiotropy: fuzzy Haseman-Elston. Genet Epidemiol 24:253–264

Kraepelin E (1909) Psychiatrie. 8 Auflage. Barth, Leipzig [reprinted English translation (1971) Dementia praecox and paraphrenia. Krieger Publishing, Huntington, New York]

Kremen WS, Buka SL, Seidman LJ, Goldstein JM, Koren D, Tsuang MT (1998) IQ decline during childhood and adult psychotic symptoms in a community sample: a 19-year longitudinal study. Am J Psychiatry 155:672–677

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

Manton KG, Woodbury MA, Tolley DH (1994) Statistical applications using fuzzy sets. John Wiley, New York

Owen MJ, Williams NM, O'Donovan MC (2004) The molecular genetics of schizophrenia: new findings promise new insights. Mol Psychiatry 9:14–27

Raine A (1991) The SPQ: a scale for the assessment of schizotypal personality based on DSM-III-R criteria. Schizophr Bull 17:555–564

Sitskoorn MM, Aleman A, Ebisch SJ, Appels MC, Kahn RS (2004) Cognitive deficits in relatives of patients with schizophrenia: a meta-analysis. Schizophr Res 71:285–295

Straub RE, Jiang Y, MacLean CJ, Ma Y, Webb BT, Myakishev MV, Harris-Kerr C, Wormley B, Sadek H, Kadambi B, Cesare AJ, Gibberman A, O'Neill FA, Walsh D, Kendler KS (2002*a*) Genetic variation in the 6p22.3 gene *DTNBP1,* the human ortholog of the mouse dysbindin gene, is associated with schizophrenia. Am J Hum Genet 71:337–348

Straub RE, MacLean CJ, Ma Y-L, Webb BT, Myakishev MC, Harris-Kerr C, Wormley B, Sadek H, Kadambi B, O'Neill AF, Walsh D, Kendler KS (2002*b*) 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

Straub RE, MacLean CJ, O'Neil FA, Burke J, Murphy B, Duke F, Shinkwin R, Webb BR, Zhang J, Walsh D, Kendler KS (1995) A potential vulnerability locus for schizophrenia on chromosome 6p24-22: evidence for genetic heterogeneity. Nat Genet 11:287–293

Torrey EF (2002) Studies of individuals with schizophrenia never treated with antipsychotic medications: a review. Schizophr Res 58:101–115

Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Jablensky A, Regier D, Sartorius N (1990) SCAN: Schedules for Clinical Assessment in Neuropsychiatry. Arch Gen Psychiatry 47:589–593

Woodbury MA, Clive J, Garson A Jr (1978) Mathematical typology: a grade of membership technique for obtaining disease definition. Comput Biomed Res 11:277–298