

ARTICLE

AHI1, a pivotal neurodevelopmental gene, and *C6orf217* are associated with susceptibility to schizophrenia

Daniela Amann-Zalcenstein¹, Nili Avidan¹, Kyra Kanyas², Richard P Ebstein³, Yoav Kohn², Adnan Hamdan⁴, Edna Ben-Asher¹, Osnat Karni², Muhammed Mujahed⁵, Ronnen H Segman², Wolfgang Maier⁶, Fabio Macchiardi⁷, Jacques S Beckmann^{1,8}, Doron Lancet¹ and Bernard Lerer^{*,2}

¹Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel; ²Biological Psychiatry Laboratory, Department of Psychiatry, Hadassah-Hebrew University Medical Center, Jerusalem, Israel; ³Department of Psychology, Hebrew University and Sara Herzog Memorial Hospital, Jerusalem, Israel; ⁴Regional Mental Health Center, Taibe, Israel; ⁵Dr Kemal Hospital, Bethlehem, Palestinian Authority, Palestine; ⁶Department of Psychiatry, University of Bonn, Germany; ⁷Department of Biology, University of Milan, Milan, Italy; ⁸Dépt de Génétique Médicale, CHUV-UNIL, Lausanne, Switzerland

Schizophrenia, a severe neuropsychiatric disorder, is believed to involve multiple genetic factors. A significant body of evidence supports a pivotal role for abnormalities of brain development in the disorder. Linkage signals for schizophrenia map to human chromosome 6q. To obtain a finer localization, we genotyped 180 single nucleotide polymorphisms (SNPs) in a young, inbred Arab-Israeli family sample with a limited number of founders. The SNPs were mostly within a ~7 Mb region around the strong linkage peak at 136.2 Mb that we had previously mapped. The most significant genetic association with schizophrenia for single SNPs and haplotypes was within a 500 kb genomic region of high linkage disequilibrium (LD) at 135.85 Mb. In a different, outbred, nuclear family sample that was not appropriate for linkage analysis, under-transmitted haplotypes incorporating the same SNPs (but not the individual SNPs) were significantly associated with schizophrenia. The implicated genomic region harbors the *Abelson Helper Integration Site 1 (AHI1)* gene, which showed the strongest association signal, and an adjacent, primate-specific gene, *C6orf217*. Mutations in human *AHI1* underlie the autosomal recessive Joubert Syndrome with brain malformation and mental retardation. Previous comparative genomic analysis has suggested accelerated evolution of *AHI1* in the human lineage. *C6orf217* has multiple splice isoforms and is expressed in brain but does not seem to encode a functional protein. The two genes appear in opposite orientations and their regulatory upstream regions overlap, which might affect their expression. Both, *AHI1* and *C6orf217* appear to be highly relevant candidate genes for schizophrenia. *European Journal of Human Genetics* (2006) 14, 1111–1119. doi:10.1038/sj.ejhg.5201675; published online 14 June 2006

Keywords: schizophrenia; association; *AHI1*; *C6orf217*; 6q23; SNPs

*Correspondence: Professor B Lerer, Biological Psychiatry Laboratory, Department of Psychiatry, Hadassah-Hebrew University Medical Center, Ein Karem, Jerusalem 91120, Israel.
Tel: +972 2 6777185; Fax: +972 2 6439294;
E-mail: lerer@cc.huji.ac.il
Received 28 February 2006; revised 2 May 2006; accepted 2 May 2006;
published online 14 June 2006

Introduction

Schizophrenia (MIM 181500) is a severe neuropsychiatric disorder with overt onset in late adolescence or early adulthood. Its clinical manifestations include delusions, hallucinations and thought disorder (positive symptoms),

lack of motivation and poor social and occupational adjustment (negative symptoms) as well as cognitive dysfunction (primarily impairment of executive functions). With a lifetime risk of ~1%, schizophrenia is the leading cause of chronic psychiatric hospitalization and represents a major public health and economic burden. While the etiology of schizophrenia is not known, a substantial body of evidence supports a pivotal role for abnormalities of brain development in utero and postnatally.^{1,2}

It has long been hypothesized that genetic factors play a significant role in schizophrenia, with strong support from family, twin and adoption studies.^{3–5} Although linkage studies aimed at mapping schizophrenia susceptibility loci have not been fully consistent, a confluence of findings for certain chromosomal regions is gradually emerging^{6,7} with support from meta-analyses.^{8,9} Further efforts have focused on identifying susceptibility genes, mostly in regions previously implicated by linkage studies.^{10–12} There have been several findings that have been replicated in some but not all studies, most prominently *DTNBP1*, *Neuregulin1*, *G72* and *DAAO*, *RGS4*, *COMT*^{11,12} and *DISC1*.^{10,13} While these genes have potential pathophysiological relevance to schizophrenia and in some cases a putative role in brain development, pathogenic variants have not been identified.

Recently, considerable interest has focused on the long arm of chromosome 6 where several studies have mapped putative schizophrenia susceptibility loci^{14–20} (for a review see Kohn and Lerer²¹). Furthermore, several groups have reported evidence for linkage of bipolar disorder on chromosome 6q^{22–24} with additional support from a recent meta-analysis.²⁵ Given the extensive genomic distance spanned by these reports, it is feasible that the incriminated chromosome 6q region harbors more than one gene implicated in the pathogenesis of schizophrenia, bipolar disorder and possibly other neuropsychiatric phenotypes.²¹

Based on an autosomal scan of Arab-Israeli families, we previously reported linkage of schizophrenia to chromosome 6q23.²⁰ After fine mapping,²⁶ the peak nonparametric LOD score (NPL) was 4.98 ($P=0.0000058$) at D6S1626 (136.3 Mb) giving rise to a putative susceptibility region (NPL-1) of 3.90 Mb. In order to identify susceptibility genes within the region, we embarked on extensive genotyping of single nucleotide polymorphisms (SNPs) within and flanking 20 putative candidate genes selected for being expressed in brain and having potential pathophysiological relevance to schizophrenia, as well as in intergenic regions. We report the identification of a 500 kb genomic region that is significantly associated with schizophrenia even after correction for multiple testing. This region harbors two genes, the neurodevelopmentally implicated *Abelson Helper Integration Site 1* gene (*AH11*), and an adjacent, primate-specific transcript, *C6orf217*.

Subjects and methods

Family samples and diagnostic methods

The index family-based association sample for this study (TKT) consisted of 53 nuclear families of Arab-Israeli origin (190 individuals, 85 affected), 21 of whom were included in our genome scan,^{20,26} 32 were recruited subsequently. 34 were 'triad' families (affected proband plus both parents), 10 families had two affected offspring and nine had 3–5 affected. Families from the genome scan sample with more than one branch were represented in the family-based association sample by a single family unit. In two cases these family units were triads. The Hadassah-Hebrew University Medical Center Helsinki Committee (Internal Review Board) approved the project and all subjects gave written informed consent. Best-estimate consensus psychiatric diagnoses,²⁷ according to the Research Diagnostic Criteria (RDC)²⁸ and the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV),²⁹ were based on interviews with the Schedule for Affective Disorders and Schizophrenia-Lifetime Version (SADS-L),³⁰ the Family History Research Diagnostic Criteria (FH-RDC)³¹ and medical records. All diagnostic evaluations were completed without knowledge of the genotyping data. A broad diagnostic category was employed and this encompassed 65 subjects with schizophrenia, 17 with schizoaffective disorder and three with unspecified functional psychosis (RDC) (further details of these procedures are provided in 'Detailed description of family sample and diagnostic methods' (online only)).

The replication sample (BT) consisted of 209 nuclear families with a total of 678 individuals, 258 of whom were affected; 177 were triad families, 23 had two affected offspring and nine more than two (3–5). Owing to the preponderance of triad families this sample is not suitable for linkage analysis. The affected individuals were diagnosed with schizophrenia on the basis of interview with the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID,³²) according to DSM-IV.²⁹ This sample is also of Arab origin, but it is outbred and was recruited from different geographical regions of Israel and the Palestinian Authority. Reliability of diagnoses across the two samples was established by a diagnostic reliability study (described in full in 'Detailed description of family sample and diagnostic methods' (online only)).

Genotyping

Genotyping was performed without knowledge of clinical diagnosis. Information about SNPs was obtained from public databases (Ensembl release 29.35b, NCBI build 35) as well as from the Celera database. The Sequenom MassARRAY platform (Sequenom, San Diego, CA, USA) was used for high throughput genotyping. We used the protocol for high multiplex homogeneous MassEXTEND (hME) reactions (Sequenom, San Diego, CA, application notes). Genotyping assays were designed as multiplex

reactions using SpectroDESIGNER software version 2.0.7 (Sequenom, San Diego CA) after checking that the SNPs do not reside in repetitive elements. The acquired genotypes were checked for deviations from Mendelian inheritance using the program, PedManager (<http://www.broad.mit.edu/ftp/distribution/software/pedmanager/>). Ambiguous markers were removed from the analysis in any specific family, if re-examination of the raw data did not resolve the ambiguity. The rate of genotyping errors was below 2% for all markers in the entire sample.

Statistical analysis

PBAT version 3.0^{33,34} (<http://www.biostat.harvard.edu/~fbat/pbat.htm>), which incorporates an extended and improved transmission disequilibrium test (TDT),³⁵ was used for association analysis. PBAT statistics were calculated under the null hypothesis of linkage and no association using the sandwich option (sw) for robust estimation of the variance, conditioning on traits and parental genotypes. The null hypothesis of no linkage and no association was used for the replication sample. Haplotype analysis was restricted to adjacent SNPs. The mode of inheritance of schizophrenia is complex; therefore we chose the additive mode. We restricted the minimal number of informative families to 10 and set the minimal haplotype frequency cutoff to 0.05. Stringent Bonferroni correction for multiple testing was used separately for single SNP and haplotype analysis. Haploview Version 3.2 (<http://www.broad.mit.edu/mpg/haploview/>) was used to calculate intermarker linkage disequilibrium (LD) between all SNP pairs within a 1 Mb interval and to generate a graphical view of the LD pattern across the entire genomic region. Haplotype blocks were defined using the confidence interval algorithm.³⁶

In silico prediction of transcription factor binding sites

The putative promoter region of genes *AHI1* and *C6orf217* was scanned for the presence of 337 human positional weight matrices (PWMs) downloaded from TRANSFAC.

The promoter scan was carried out via a Perl script that utilizes the TFBS module <http://forkhead.cgb.ki.se/TFBS/>.³⁷ This module implements the PWM search algorithm described in.³⁸ The search cutoff used was 92%.

Results

We genotyped 180 SNPs within a 13.9 Mb linkage interval in 53 families from an Arab Israeli sample that shows linkage to schizophrenia on chromosome 6q23^{20,26} (Figure 1a). The region harbors 69 known genes, the majority showing brain expression (Figure 1b). Genotyped SNP density was highest in a 1 Mb area around the linkage peak at 136 Mb with a total of 58 SNPs at an average inter-SNP distance of 17.0 Kb (Figure 1c); the average density of

the remaining SNPs was 66.6 kb. A graphical display of the LD pattern giving rise to 23 LD blocks is shown in Figure 1d. The underlying D' and r^2 values for SNP by SNP comparison across the entire region are available from the authors on request.

Single SNP analysis revealed the highest association signal within a ~0.5 Mb genomic region (135.6–136.1 Mb). This association signal included the *AHI1* gene at 135.7 Mb, extending into the distal intergenic region between *AHI1* and the *Phosphodiesterase 7B* gene (*PDE7B*) (Table 1, Figure 2a, Full genotyping results in Supplementary Table S1 (online)). It also included a primate-specific gene of unknown function, *C6orf217*, for which there is EST-cluster evidence (UniGene cluster HS.510098). This gene is directly adjacent to *AHI1* and in high LD with it (Figure 3). After conservative Bonferroni correction for multiple testing (180 tests, $P=0.00028$), two out of the six SNPs within *AHI1*, remained significant. Within *C6orf217* four SNPs withstood correction. The last significant SNP in this region (rs1475069, $P=0.0000025$) is located ~48 kb distal to the last annotated exon of *C6orf217* in the intergenic region between it and *PDE7B*. In addition to the cluster of associated SNPs within the high LD region there were four SNPs in other genomic locations that withstood Bonferroni correction (Figure 2a, Supplementary Table S1 (online)).

C6orf217 is primate-specific gene consisting of 10 exons and it has several alternatively splice isoforms. The predicted protein length depends on the splice isoform with a maximum of 135 amino acids with no similarity to any other known protein.³⁹ Its largest open reading frame resides across exons one to three, while all the other exons seem to belong to the 3' untranslated region (UTR). *C6orf217* is expressed in brain, eye, kidney, testis, tongue, pancreas and lung during development as well as in the adult.

C6orf217 and *AHI1* transcripts are genomically situated head to head, with only 55 bp distance between their 5' ends, in all probability putting constraints on their mutual transcription. The two genes share two potential transcription factor binding sites (TFBS) for CREB (–30 bp for *AHI1* and –12 bp for *C6orf217*) and ATF at –36 *C6orf217* and at –20 for *AHI1*). Moreover, potential TFBSs for *AHI1* gene are located within *C6orf217*, for example, RFX1 (–64 bp), STAT1 (–272 bp), NF-AT (–385 bp), AP1 (–682 bp) and AML-1a (–992 bp), as well as TFs of *C6orf217* that are located at the *AHI1* gene, for example, T3R (–89), Nrf-1 (–139), SRY (–163) and E2F (–195).

Deviation from Hardy–Weinberg equilibrium was observed in affected offspring only, with six out of a total of nine deviating SNPs residing in the *AHI1–C6orf217* interval (Supplementary Figure S1 (online)). This puts additional emphasis on this genomic region, since the observation is distinct from the observed allelic association and reflects enrichment for specific genotypes in the

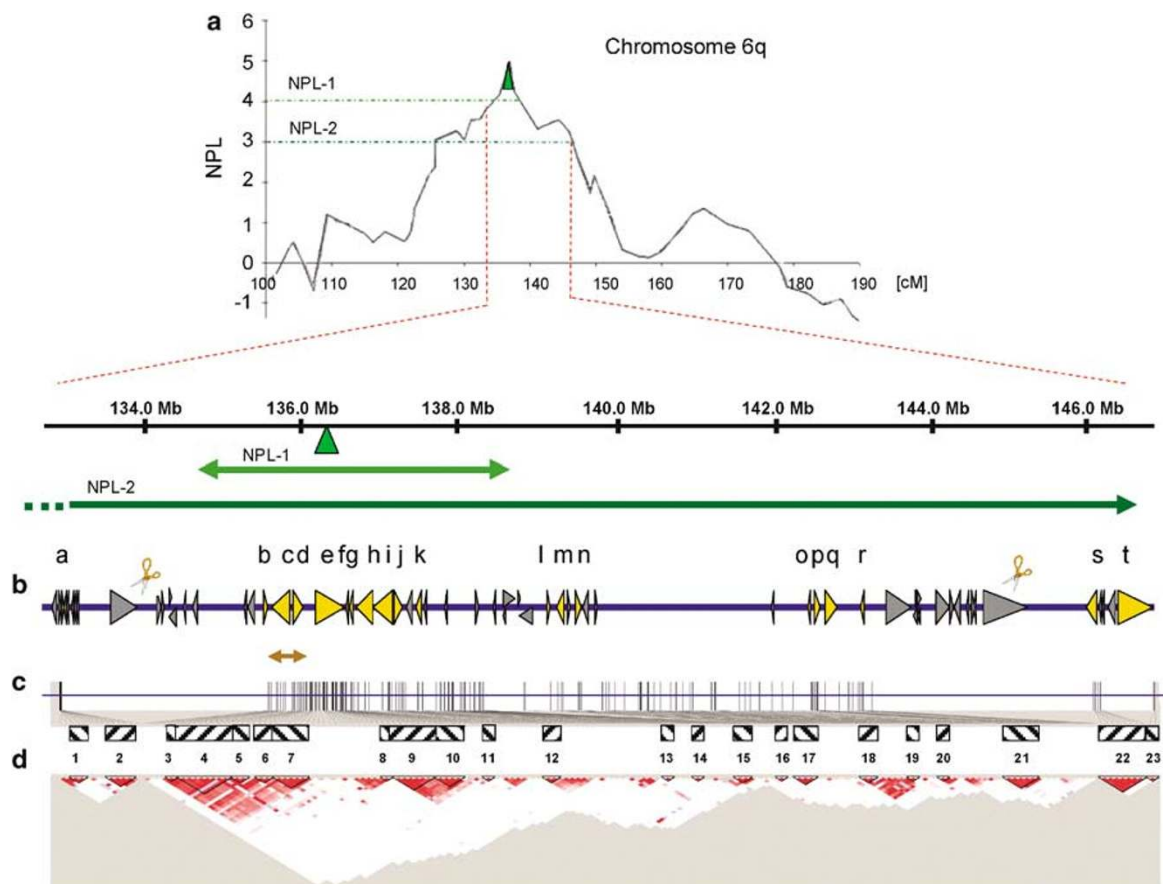


Figure 1 Genomic region within the NPL-2 confidence interval of a linkage peak on chromosome 6q23 for fine mapping of a schizophrenia susceptibility locus in an Arab-Israeli sample. (a) Multipoint, nonparametric linkage analysis of microsatellite markers on chromosome 6q under a broad diagnostic model showing a maximum NPL of 4.98 ($P=0.00000058$) at 136.97 cM.²⁶ The NPL-1 (3.9 Mb) and NPL-2 (~20 Mb) confidence intervals are indicated by broken lines in shades of green. The linkage peak is represented by a green triangle. (b) The NPL-1 and NPL-2 confidence intervals are indicated by arrows in shades of green. The position of the linkage peak is depicted by a green triangle at 136.3 Mb. Distribution of known genes spanning a ~14 Mb genomic region underneath the linkage peak is shown. Gene polarity is indicated by the orientation of the triangles. Yellow and grey triangles represent, respectively, genes covered by genotyped SNPs or devoid of genotyped SNPs. The scissors represent known recombination hotspots. The brown double-headed arrow represents the 500 kb high LD region around the *AHI1* gene. Genes covered by genotyped SNPs are: a: *TRAR4*, b: *MYB*, c: *AHI1*, d: *C6orf217*, e: *PDE7B*, f: *FAM54A*, g: *BCL2A*, h: *MAP7*, i: *MAP3K5*, j: *PEX7*, k: *IL22RA2*, l: *TNFAIP3*, m: *C6orf63*, n: *REPS1*, o: *HECA*, p: *NMBR*, q: *C6orf55*, r: *GPR126*, s: *EPM2A*, t: *GRM1* (c). Distribution of SNPs within the 14 Mb genomic region under the linkage peak. Thick black lines indicate dense clusters of SNPs. The SNP density is highest in a ~1 Mb region from 135.5 to 136.5 Mb with an average inter-SNP distance of 17.0 kb, followed by the ~7 Mb genomic region from 136.5 to 143.5 Mb with an average inter-SNP distance of 66.6 kb. The remaining SNPs are distributed across 3 candidate genes, namely *TRAR4* (a), *EPM2A* (s) and *GRM1* (t). Positions of the 23 haplotype blocks relative to the LD plot are shown underneath, as well as the LD plot generated using HAPLOVIEW software with pairwise SNP comparison for SNPs less than 1 Mb apart. Red indicates regions of high LD and white represents regions of low LD.

patients that might indicate the loss of a protective allele, increasing the risk of *C6orf217*. There is no indication that the increased level of consanguinity in this population affects the Hardy–Weinberg equilibrium across this genomic region, since the parents exhibit a normal distribution of genotypes throughout the entire interval. Furthermore, the calculations (using PBAT software) correct for linkage and multiple offspring within one family, which otherwise would inflate the results.

To increase the information provided by single SNP analysis, we performed three-SNP ‘sliding window’ haplo-

type analyses using all 180 SNPs. A total of 612 individual haplotypes were analyzed (on average three haplotypes per sliding window) putting the Bonferroni cutoff P -value for significant haplotype association at 0.000082. Outstanding haplotype association signals were observed within the *AHI1* gene (of the order of $P=0.0000000017$) with haplotypes encompassing rs9321501, followed by strong association of the *C6orf217* locus (Figure 2b, Supplementary Table S2 (online)). Haplotype association exceeded the association of single SNPs in this region in degree of significance implying a convergence towards the putative

Table 1 Single SNP association results for the high-density region underneath the linkage peak (20 SNPs) on chromosome 6q, from 135.6 to 136.2 Mb in the TKT sample

#	Ref Seq #	Position (bp)	Gene	#_info_fam	SNP	MAF	MA	OTA	P-value
14	rs3819409	135567089	MYB	19	G/T	0.40	T	T	0.304
15	rs1013891	135580332	MYB	25	G/A	0.33	A	A	0.6517
16	rs878175	135601850		19	T/C	0.18	C	C	0.6065
17	rs6931735	135666504	AHI1	24	A/G	0.25	G	A	0.0038
18	rs6912933	135669227	AHI1	25	A/G	0.20	G	A	0.0003
19	rs9321501	135683110	AHI1	24	A/C	0.22	C	A	0.0001
20	rs2746429	135714977	AHI1	22	T/C	0.22	C	T	0.0036
21	rs2614258	135718895	AHI1	23	G/A	0.17	A	G	0.0003
22	rs2614264	135752387	AHI1	16	A/G	0.25	G	A	0.1341
23	rs11154801	135781048	AHI1	21	C/A	0.18	A	C	7.06 × 10⁻⁷*
24	rs7750586	135869366	C6orf217	20	A/G	0.17	G	A	0.0002*
25	rs9647635	135882749	C6orf217	18	A/C	0.19	C	A	1.90 × 10⁻⁵*
26	rs12203875	135902107	C6orf217	27	A/G	0.31	G	A	0.0016
27	rs10484771	135939628	C6orf217	11	C/T	0.16	T	T	0.3003*
28	rs4896156	135986556	C6orf217	25	A/G	0.29	G	A	0.0011*
29	rs9376112	135969693	C6orf217	26	A/C	0.26	C	A	0.0173
30	rs2143681	135974698	C6orf217	25	G/A	0.27	A	G	0.0419
31	rs6925684	135978571	C6orf217	27	G/A	0.29	A	G	0.0142
32	rs6902485	135999485	C6orf217	14	G/A	0.10	A	A	0.5175
33	rs6935033	136022330	C6orf217	26	G/A	0.38	A	G	0.0004
34	rs7739635	136039471	C6orf217	20	C/T	0.37	T	C	5.90 × 10⁻⁶*
35	rs9494332	136050301	C6orf217	26	G/A	0.32	A	G	2.98 × 10⁻⁵*
36	rs1475069	136097927		24	A/C	0.14	C	A	2.51 × 10⁻⁶*
37	rs911507	136101116		30	G/A	0.38	A	G	0.0005
38	rs12211505	136116514		25	T/C	0.29	C	T	0.003
39	rs2208574	136117310		29	G/A	0.24	A	G	0.0015
40	rs9399161	136122157		29	A/G	0.29	G	A	0.0119
41	rs6925090	136158170		29	T/C	0.29	C	T	0.0063
42	rs947583	136175352		24	T/C	0.28	C	T	0.0063
43	rs1342236	136206493		17	G/A	0.11	A	G	0.7777
44	rs2298118	136217327	PDE7B	13	T/A	0.16	A	T	0.4847

P-values were obtained using PBAT. Bonferroni cutoff P-value for 180 tests is 0.00028. SNPs which pass the correction for multiple testing are highlighted in bold print. The asterisk (*) next to the P-value indicates SNPs, which show deviation from Hardy–Weinberg-equilibrium only in affected individuals.

Abbreviations: MAF = minor allele frequency, MA = minor allele, OTA = overtransmitted allele.

disease locus. Two individual haplotypes in other regions remained significant after correction for multiple testing. One is in an intergenic region at 138.1 Mb and the other almost 1 Mb away at 139.0 Mb with contributing SNPs residing in two different genes namely *C6orf63* and *REPS1*. For neither of these loci is there cumulative evidence from single SNP and haplotype analysis.

We genotyped eight out of 11 SNPs that survived stringent Bonferroni correction, in a second sample consisting of 209 outbred nuclear families of Arab origin, each with at least one offspring affected with schizophrenia (BT sample). Three SNPs (rs7739635, rs9494332, rs642162) were omitted for assay design reasons. None of the eight individual SNPs showed significant association with schizophrenia (Supplementary Table S3 (online)). However, sliding window analysis revealed haplotypes consisting of SNPs located within *AHI1* and *C6orf217* that were strongly associated with schizophrenia in the BT sample (threshold after Bonferroni correction for 25 haplotypes tested, $P=0.002$) (Figure 4, Supplementary

Table S4 (online)). The same SNPs compose the associated haplotypes in the TKT and BT samples but no shared haplotype was observed. Only the C allele of SNP rs9321501 located in *AHI1* is shared by both samples. The frequencies of common individual haplotypes in the two samples are comparable, although more rare haplotypes were observed in the larger and more outbred BT sample (Supplementary Table S4 (online)).

Discussion

We analyzed 53 families from a sample that shows linkage to schizophrenia on chromosome 6q23, by genotyping 180 SNPs in a genomic region of 13.9 Mb. A 0.5 Mb region of high LD in the vicinity of the original linkage peak at 136.3 Mb was significantly associated with the illness. The results of single SNP and haplotype association analyses focused on the *Abelson Helper Integration Site 1* (*AHI1*) gene and an adjacent, primate-specific gene of unknown function, *C6orf217*.

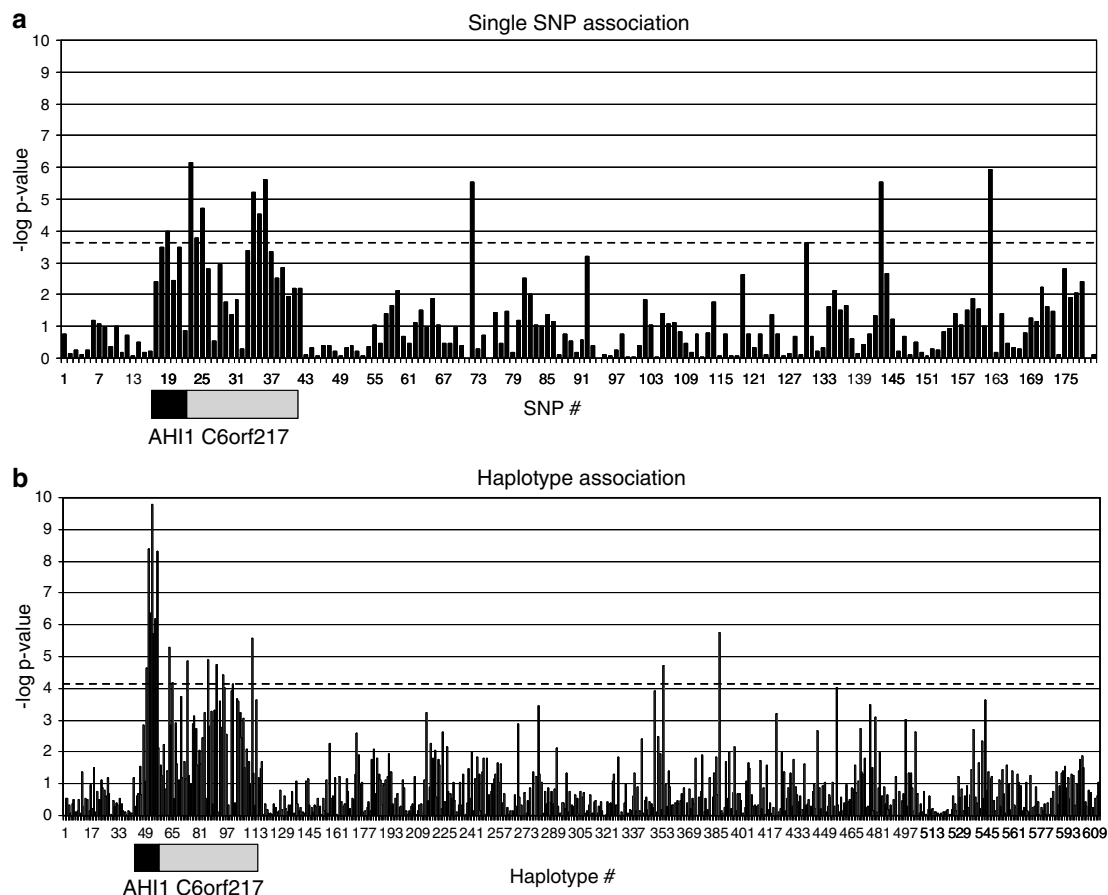


Figure 2 Association results within the NPL-2 confidence interval of a linkage peak on chromosome 6q23 of a schizophrenia-susceptibility-locus in an Arab-Israeli sample. (a) The SNP # corresponds to the numbering in Supplementary Table S1 (online). The broken line represents the Bonferroni cutoff for multiple testing ($P = 0.00028$). The significant SNP cluster from SNP #17 to #41 spans from *MYB* to *C6orf217*. The black box indicates the SNPs which reside within *AHI1* and the grey box indicates the SNPs residing in *C6orf217*. (b) Haplotype analysis across the 13.9 Mb genomic region on chromosome 6q using 3-SNP sliding windows. The haplotype # corresponds to the numbering in Supplementary Table S2 (online). The broken red line represents the Bonferroni cutoff for multiple testing ($P = 0.000082$). The black box indicates the haplotypes that reside within *AHI1*; the grey box indicates the SNPs residing in *C6orf217*. The most significant haplotype cluster around haplotype #53 is located within the *AHI1* gene. With only two common haplotypes in this region the most significant association is observed with the frequent and over-transmitted AAT haplotype (frequency: 0.67) consisting of SNPs rs6912933, rs9321501 and rs2746429 ($P = 0.0000000017$) and a similar, but less significant trend, with the complementary under-transmitted GCC haplotype (frequency: 0.22, $P = 0.0000019$).

The study of a complex disorder such as schizophrenia is facilitated by the homogeneous nature and possible founder effect in our well-characterized family sample.²⁰ The Arab-Israeli population is an ethnically homogeneous group that has a high birthrate, an unusually high level of consanguinity ($\sim 25\%$ first cousin marriages) and a low rate of intermarriage with other population groups.^{40,41} Our primary, TKT sample was derived from three towns that were founded 200–250 years ago by a limited number of families.⁴² In subsequent years there was immigration into the towns but the major population increase has been due to a high birthrate and low infant mortality in the past 75 years. Therefore, we assume a founder effect in this population, whereby most of the affected individuals will share the same chromo-

somal region identical by descent in the vicinity of the disease locus.

To validate our findings we examined a second family sample of Arab origin. This sample (BT) was outbred and showed no prior linkage to chromosome 6q23. Limited replication of the association in this sample provides additional support for the involvement of *AHI1* in the genetic etiology of schizophrenia but cannot be regarded as confirmatory. It is noteworthy that the BT sample consisted of sporadic rather than familial cases and that the phenotype in the replication sample was more narrowly defined and encompassed only patients with schizophrenia according to DSM-IV. The fact that no shared haplotype was identified in two Arab samples for which a common founder mutation is postulated, could be due to several

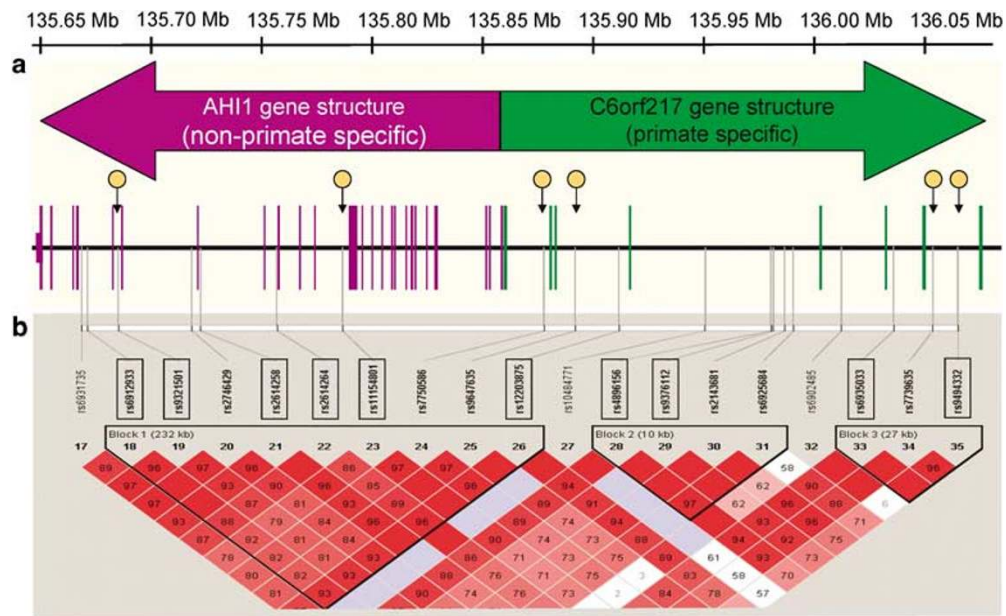


Figure 3 Genomic location and orientation of *AH1* and *C6orf217*. (a) The purple and green arrows indicate *AH1* and *C6orf217* orientation, respectively. Gene structure of *AH1* and *C6orf217* are depicted in purple and green vertical lines respectively, indicating exon positions. Yellow circles and their connecting arrows point to the location of significant SNPs (after Bonferroni correction). Grey lines indicate the positions of all SNPs genotyped in this genomic region. (b) Haploview representation of SNPs and their LD structure. Grey lines indicate the positions of the SNPs. Blocks were defined using the confidence interval method (integrated in Haploview). Haplotype tagging SNPs are surrounded by rectangles.

Cohort	SNPs					S	E(S)	P-value	Haplotype freq
	rs9321501	rs11154801	rs7750586	rs9647635	rs1475069				
TKT	C (0.22)	A (0.18)				13	25.4	0.00000	0.20
BT	C (0.36)	C (0.70)				20	32.1	0.00046	0.16
TKT	C (0.22)	A (0.18)	G (0.17)			12	20.1	0.00046	0.17
BT	C (0.36)	C (0.70)	A (0.73)			15	26.1	0.00032	0.13
TKT	C (0.22)	A (0.18)	G (0.17)	C (0.19)		12	20.1	0.00046	0.18
BT	C (0.36)	C (0.70)	A (0.73)	A (0.72)		15	25.6	0.00047	0.10
TKT	C (0.22)	A (0.18)	G (0.17)	C (0.19)	C (0.14)	5	13.5	0.00016	0.11
BT	C (0.36)	C (0.70)	A (0.73)	A (0.72)	A (0.76)	14	21.8	0.00415	0.10

Figure 4 Associated under-transmitted, but not common haplotypes around the *AH1* gene on chromosome 6q23 in the index (TKT) and replication (BT) samples. Major and minor alleles are represented by grey and white boxes, respectively. The associated nucleotides with their allele frequency for the given haplotypes are shown together inside the boxes. The numbers of observed versus expected transmissions of the haplotypes are referred to in columns S and E(S), respectively.

reasons. First, the SNPs defining the ‘at-risk-haplotype’ may not have been identified and genotyped yet; the SNP density in the replication sample does not allow any firm conclusion on this matter. Second, genetic and allelic heterogeneity within this ethnically homogeneous population cannot be excluded; such situations have been encountered for several monogenic traits^{43,44} and might also apply to complex multifactorial traits. Third,

an ancestral recombination event between the SNPs, rs9321501 and rs11154801, might have led the causative SNP to be found on complementary haplotypes in the two cohorts. As allele frequencies are comparable in the two ethnically matched samples, the existence of a common haplotypes is regarded as likely. Considering the under-transmitted haplotype, apart from the first SNP in the haplotype, rs9321501, complementary alleles are involved

in the two cohorts hampering identification of such an 'at-risk-haplotype'.

Human *AH11* is a very attractive biological candidate for a schizophrenia susceptibility gene. First, the gene is highly expressed in human fetal brain tissue and in the cerebellum and cerebral cortex of adult human brain.⁴⁵ Second, comparative genetic analysis suggests accelerated evolution of *AH11* in the human lineage, particularly the N-terminal region of the gene; this was interpreted by Ferland *et al*⁴⁵ to be a consequence of positive selection. Third, *AH11* contains several motifs that have been shown to be present in signaling molecules and in molecules mediating protein–protein interactions. Based on the study of murine *ah11* it has been suggested that *AH11* may be a docking site or scaffold protein recruiting other signaling molecules and modulating and integrating their action.⁴⁶

Recently, null-type and missense mutations in human *AH11* have been shown to cause the autosomal recessive brain disorder, Joubert Syndrome (JS).^{45,47} The characteristic features of JS include agenesis of the cerebellar vermis, ataxia, hypotonia, oculomotor apraxia and mental retardation. Additional phenotypic manifestations are diverse and may include cerebral polymicrogyria and involvement of systems other than the CNS.⁴⁷ It is of interest that *AH11* is the only JS locus that is associated with cortical abnormalities.⁴⁷ Less pronounced structural mutations in *AH11* or variants that influence expression levels could affect brain development in utero or postnatally in more subtle ways, thus contributing to susceptibility to schizophrenia. Brain imaging studies of schizophrenia, including patients in their first episode of psychosis, have consistently identified increased volume of the lateral ventricles, slightly decreased overall brain size, and decreases of gray matter, white matter and regional volumes in hippocampus, thalamus and frontal lobes.² Early post-mortem findings report neuronal disarray and abnormal migration of subplate neurons in the neocortical white matter suggesting prenatal neurodevelopmental abnormalities. More recent findings have emphasized abnormalities in neuronal size, arborization, and synaptic organization.¹¹ A gene such as *AH11* which is intimately involved in brain development, could be implicated in several of the brain abnormalities observed in schizophrenia.

The observed association of SNPs and haplotypes with schizophrenia extends from *AH11* into *C6orf217*. The association is more prominent in the TKT sample than in the BT replication sample. As *C6orf217* encodes a small protein with no known motif it may well function as a regulatory gene either at its mRNA or protein level. We may speculate that due to the close proximity of the two genes *AH11* expression level may be dependent on *C6orf217* gene expression.

Until now no causative, coding variants have been reported for any schizophrenia susceptibility gene. This has led to the assumption that changes in gene function

may be due to variation in gene regulation rather than changes associated with the protein *per se*. A more detailed study of *AH11* and *C6orf217* and their flanking regions is required to further elucidate their regulation and to facilitate the search for regulatory variants in addition to further fine mapping and 'classical' screening of coding regions.

Acknowledgements

The research was supported in part by the Israel Ministry of Science (Korean–Israeli Collaboration in Human Genome Research), the Benozio Center for Neurological Disease, the Israel Science Foundation (Israel National Academy of Sciences) and Hadassah Medical Organization. Genotyping was performed at the Crown Human Genome Center, Weizmann Institute of Science supported by the Israel Ministry of Science. The Biological Psychiatry Laboratory, Hadassah-Hebrew University Medical Center has been supported by the Harry Stern Family Foundation. The authors are indebted to the late Mrs Blanche Stern Prince and to Dr Morton Prince for their encouragement and support. The authors thank Tsviya Oelander and Michal Lapidot for their valuable advice.

References

- Weinberger DR: Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 1987; **44**: 660–669.
- Rapoport JL, Addington AM, Frangou S, Psych MR: The neurodevelopmental model of schizophrenia: update 2005. *Mol Psychiatry* 2005; **10**: 434–449.
- Gottesman I: *Schizophrenia Genesis: The Origin of Madness*. New York: WH Freeman, 1991.
- Tsuang MT, Stone WS, Faraone SV: Schizophrenia: a review of genetic studies. *Harv Rev Psychiatry* 1999; **7**: 185–207.
- Lichtermand D, Karbe E, Maier W: The genetic epidemiology of schizophrenia and of schizophrenia spectrum disorders. *Eur Arch Psychiatry Clin Neurosci* 2000; **250**: 304–310.
- Riley BP, McGuffin P: Linkage and associated studies of schizophrenia. *Am J Med Genet* 2000; **97**: 23–44.
- Baron M: Genetics of schizophrenia and the new millennium: progress and pitfalls. *Am J Hum Genet* 2001; **68**: 299–312.
- Badner JA, Gershon ES: Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. *Mol Psychiatry* 2002; **7**: 405–411.
- Lewis CM, Levinson DE, Wise LH *et al*: Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. *Am J Hum Genet* 2003; **73**: 34–48.
- Millar JK, Wilson-Annan JC, Anderson S *et al*: Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum Mol Genet* 2000; **9**: 1415–1423.
- Harrison PJ, Owen MJ: Genes for schizophrenia? Recent findings and their pathophysiological implications. *Lancet* 2003; **361**: 417–419.
- Shirts BH, Nimgaonkar V: The genes for schizophrenia: finally a breakthrough? *Curr Psychiatry Rep* 2004; **6**: 303–312.
- Hennah W, Varilo T, Kestila M *et al*: Haplotype transmission analysis provides evidence of association for DISC1 to schizophrenia and suggests sex-dependent effects. *Hum Mol Genet* 2003; **12**: 3151–3159.
- Cao Q, Martinez M, Zhang J *et al*: Suggestive evidence for a schizophrenia susceptibility locus on chromosome 6q and a confirmation in an independent series of pedigrees. *Genomics* 1997; **43**: 1–8.

- 15 Kaufmann CA, Suarez B, Malaspina D *et al*: NIMH Genetics Initiative Millenium Schizophrenia Consortium: linkage analysis of African-American pedigrees. *Am J Med Genet* 1998; **81**: 282–289.
- 16 Martinez M, Goldin LR, Cao Q *et al*: Follow-up study on a susceptibility locus for schizophrenia on chromosome 6q. *Am J Med Genet* 1999; **88**: 337–343.
- 17 Bailer U, Leisch F, Meszaros K *et al*: Genome scan for susceptibility loci for schizophrenia. *Neuropsychobiology* 2000; **42**: 175–182.
- 18 Levinson DF, Holmans P, Straub RE *et al*: Multicenter linkage study of schizophrenia candidate regions on chromosomes 5q, 6q, 10p, and 13q: schizophrenia linkage collaborative group III. *Am J Hum Genet* 2000; **67**: 652–663.
- 19 Lindholm E, Ekholm B, Shaw S *et al*: A schizophrenia-susceptibility locus at 6q25, in one of the world's largest reported pedigrees. *Am J Hum Genet* 2001; **69**: 96–105.
- 20 Lerer B, Segman RH, Hamdan A *et al*: Genome scan of Arab Israeli families maps a schizophrenia susceptibility gene to chromosome 6q23 and supports a locus at chromosome 10q24. *Mol Psychiatry* 2003; **8**: 488–498.
- 21 Kohn Y, Lerer B: Excitement and confusion on chromosome 6q: the challenges of neuropsychiatric genetics in microcosm. *Mol Psychiatry* 2005; **10**: 1062–1073.
- 22 Dick DM, Foroud T, Flury L *et al*: Genomewide linkage analyses of bipolar disorder: a new sample of 250 pedigrees from the National Institute of Mental Health Genetics Initiative. *Am J Hum Genet* 2003; **73**: 107–114.
- 23 McClinnis MG, Dick DM, Willour VL *et al*: Genome-wide scan and conditional analysis in bipolar disorder: evidence for genomic interaction in the National Institute of Mental Health genetics initiative bipolar pedigrees. *Biol Psychiatry* 2003; **54**: 1265–1273.
- 24 Middleton FA, Pato MT, Gentile KL *et al*: Genomewide linkage analysis of bipolar disorder by use of a high-density single-nucleotide-polymorphism (SNP) genotyping assay: a comparison with microsatellite marker assays and finding of significant linkage to chromosome 6q22. *Am J Hum Genet* 2004; **74**: 886–897.
- 25 McQueen MB, Devlin B, Faraone SV *et al*: Combined analysis from eleven linkage studies of bipolar disorder provides strong evidence of susceptibility loci on chromosomes 6q and 8q. *Am J Hum Genet* 2005; **77**: 582–595.
- 26 Levi A, Kohn Y, Kanyas K *et al*: Fine mapping of a schizophrenia susceptibility locus at chromosome 6q23: increased evidence for linkage and reduced linkage interval. *Eur J Hum Genet* 2005; **13**: 763–771.
- 27 Baron M, Endicott J, Lerer B *et al*: A pedigree series for mapping disease genes in bipolar affective disorder: sampling, assessment, and analytic considerations. *Psychiatr Genet* 1994; **4**: 43–55.
- 28 Spitzer RL, Endicott J, Robins E: Research diagnostic criteria: rationale and reliability. *Arch Gen Psychiatry* 1978; **35**: 773–782.
- 29 American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, Washington, DC: American Psychiatric Association, 1994.
- 30 Spitzer R, Endicott J: *The Schedule for Affective Disorders and Schizophrenia, Lifetime Version*. New York: New York State Psychiatric Institute, 1977.
- 31 Andreasen NC, Endicott J, Spitzer RL, Winokur G: The family history method using diagnostic criteria. Reliability and validity. *Arch Gen Psychiatry* 1977; **34**: 1229–1235.
- 32 First MB, Donovan S, Frances A: Nosology of chronic mood disorders. *Psychiatr Clin North Am* 1996; **19**: 29–39.
- 33 Lange C, DeMeo D, Silverman EK, Weiss ST, Laird NM: PBAT: tools for family-based association studies. *Am J Hum Genet* 2004; **74**: 367–369.
- 34 Steen KV, Lange C: PBAT: A comprehensive software package for genome-wide association analysis of complex family-based studies. *Hum Genomics* 2005; **2**: 67–69.
- 35 Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993; **52**: 506–516.
- 36 Gabriel SB, Schaffner SF, Nguyen H *et al*: The structure of haplotype blocks in the human genome. *Science* 2002; **296**: 2225–2229.
- 37 Lenhard B, Wasserman WW: TFBS: Computational framework for transcription factor binding site analysis. *Bioinformatics* 2002; **18**: 1135–1136.
- 38 Fickett JW: Quantitative discrimination of MEF2 sites. *Mol Cell Biol* 1996; **16**: 437–441.
- 39 Close J, Game L, Clark B, Bergounioux J, Gerovassili A, Thein SL: Genome annotation of a 1.5 Mb region of human chromosome 6q23 encompassing a quantitative trait locus for fetal hemoglobin expression in adults. *BMC Genomics* 2004; **5**: 33.
- 40 Jaber L, Bailey-Wilson JE, Haj-Yehia M, Hernandez J, Shohat M: Consanguineous matings in an Israeli-Arab community. *Arch Pediatr Adolesc Med* 1994; **148**: 412–415.
- 41 Jaber L, Halpern GJ, Shohat T: Trends in the frequencies of consanguineous marriages in the Israeli Arab community. *Clin Genet* 2000; **58**: 106–110.
- 42 Akel M, Masarwi UJ: *Taibe Bani Saab: Past Present and Future*. Israel: Alrama Publications, 1989.
- 43 Zlotogora J, Zaizov R, Klibansky C, Matoth Y, Bach G, Cohen T: Genetic heterogeneity in Gaucher disease. *J Med Genet* 1986; **23**: 319–322.
- 44 Beckmann JS: The Reunion paradox and the digenic model. *Am J Hum Genet* 1996; **59**: 1400–1402.
- 45 Ferland RJ, Eyaid W, Collura RV *et al*: Abnormal cerebellar development and axonal decussation due to mutations in AHI1 in Joubert syndrome. *Nat Genet* 2004; **36**: 1008–1013.
- 46 Jiang X, Hanna Z, Kaouass M, Girard L, Jolicoeur P: Ahi-1, a novel gene encoding a modular protein with WD40-repeat and SH3 domains, is targeted by the Ahi-1 and Mis-2 provirus integrations. *J Virol* 2002; **76**: 9046–9059.
- 47 Dixon-Salazar T, Silhavy JL, Marsh SE *et al*: Mutations in the AHI1 gene, encoding joubertin, cause Joubert syndrome with cortical polymicrogyria. *Am J Hum Genet* 2004; **75**: 979–987.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)