View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Crossref

Association Study of the Dystrobrevin-Binding Gene With Schizophrenia in Australian and Indian Samples

Elizabeth G. Holliday,[1,4] Herlina Y. Handoko,[1,3] Michael R. James,[3] John J. McGrath,[1,4] Deborah A. Nertney,[1] Sujit Tirupati,[5,6] Rangaswamy Thara,[5] Douglas F. Levinson,[7] Nicholas K. Hayward,[3] Bryan J. Mowry,[1,4] and Dale R. Nyholt[2]

[1] Queensland Centre for Mental Health Research, The Park, Centre for Mental Health, Wacol, Queensland, Australia
[2] Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Queensland, Australia
[3] Human Genetics Laboratory, Queensland Institute of Medical Research, Brisbane, Queensland, Australia
[4] Department of Psychiatry, University of Queensland, St Lucia, Queensland, Australia
[5] Schizophrenia Research Foundation India, West Anna Nagar Extension, Chennai, India
[6] Psychiatric Rehabilitation Service, Hunter New England Mental, Health James Fletcher Hospital, Newcastle, Australia
[7] Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Palo Alto, California, United States of America

Numerous studies have reported association between variants in the dystrobrevin binding protein 1 (dysbindin) gene (*DTNBP1*) and schizophrenia. However, the pattern of results is complex and to date, no specific risk marker or haplotype has been consistently identified. The number of single nucleotide polymorphisms (SNPs) tested in these studies has ranged from 5 to 20. We attempted to replicate previous findings by testing 16 SNPs in samples of 41 Australian pedigrees, 194 Australian cases and 180 controls, and 197 Indian pedigrees. No globally significant evidence for association was observed in any sample, despite power calculations indicating sufficient power to replicate several previous findings. Possible explanations for our results include sample differences in background linkage disequilibrium and/or risk allele effect size, the presence of multiple risk alleles upon different haplotypes, or the presence of a single risk allele upon multiple haplotypes. Some previous associations may also represent false positives. Examination of Caucasian HapMap phase II genotype data spanning the *DTNBP1* region indicates upwards of 40 SNPs are required to satisfactorily assess all nonredundant variation within *DTNBP1* and its potential regulatory regions for association with schizophrenia. More comprehensive studies in multiple samples will be required to determine whether specific *DTNBP1* variants function as risk factors for schizophrenia.

Schizophrenia (SZ [Mendelian Inheritance in Man, MIM 181500]) is a severe, debilitating disorder characterized by delusional beliefs, hallucinations and deficits in emotional and social behavior. It has a lifetime population prevalence of ~1%. Although the etiology of SZ is not well understood, family, twin and adoption studies suggest a strong, albeit complex, genetic component and linkage analyses have identified numerous promising susceptibility loci. One of the most consistently replicated is locus SCZD3 (MIM 600511) on chromosome 6p22.3 which falls into the fifth-ranked bin in a prominent meta-analysis (Lewis et al., 2003). Following their initial linkage finding in this region using 270 Irish high-density pedigrees, Straub et al. (2002) performed systematic linkage disequilibrium (LD) mapping of the region and detected significant evidence for association (p = .008–.0001) between SZ and several single nucleotide polymorphisms (SNPs) and haplotypes spanning the 140 kb gene encoding the dystrobrevin-binding protein (*DTNBP1*), also known as 'dysbindin' (MIM 607145).

Subsequent to the report by Straub et al. (2002), numerous studies have also detected association of SZ with *DTNBP1* variants. Schwab and colleagues (2003) tested six of the most positive SNPs from Straub's study in two family samples of predominantly German descent and detected significant evidence of association in both the individual samples and the combined sample (p = .00001 for a three-locus haplotype). Van den Oord and colleagues (2003) studied all 12 previously tested SNPs and an additional two *DTNBP1* SNPs in 268 high-density SZ families and detected significant association of individual markers (p = .019–.003) and an eight-marker haplotype (p = .002). Van den Bogaert and colleagues (2003) detected significant evidence for association of a five-marker haplotype with SZ in a Swedish case-control sample (p = .0098), but not in German or Polish samples. The evidence in the Swedish sample became more significant when analysis considered only cases

Received 16 February, 2006; accepted 16 May, 2006.

Address for correspondence: Elizabeth Holliday, Queensland Centre for Mental Health Research, The Park, Centre for Mental Health, Wacol QLD 4076, Australia. E-mail: lizh@qcmhr.uq.edu.au

Downloaded from https://www.cambridge.org/core. Open University Library, on 16 Jan 2020 at 00:38:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1375/twin.9.4.531

with a family history of SZ (p = .00009). Evidence implicating *DTNBP1* in SZ development has also been provided by Kirov and colleagues (2004) in 488 Bulgarian parent–offspring trios, Funke and colleagues (2004) in 258 European-American cases and 467 controls and 51 Hispanic cases and 32 controls, and Williams and colleagues (2004), who detected strong evidence for association with novel risk and protective haplotypes in 708 Welsh cases and 711 controls and, importantly, replicated their specific findings in an independent sample of 219 Irish cases and 231 controls. Evidence for involvement of *DTNBP1* in SZ has also been reported for non-European samples, by Tang and colleagues (2003) who observed transmission disequilibrium in 233 Han Chinese trios, Numakawa and colleagues (2004), who detected numerous associated haplotypes in 670 Japanese cases and 588 controls and Li and colleagues (2005), who detected significant overtransmission of SNPs and haplotypes to cases in 638 Han Chinese families.

Taken together, these data provide compelling evidence that *DTNBP1* is a susceptibility gene for SZ. However, some caution is also warranted. A number of studies have been unable to detect association of *DTNBP1* variants with SZ: Van den Bogaert and colleagues (2003) found no evidence for association in their case-control samples of German (418 cases, 285 controls) and Polish (294 cases, 113 controls) origin; neither did Funke and colleagues (2004), in an African-American subset of their case-control sample (215 cases, 74 controls), or Li and colleagues (2005), in their large Scottish sample (580 cases, 620 controls). Furthermore, although many studies have reported significant evidence for association, the pattern of results is complex. The associated alleles and haplotypes vary widely, and none has been consistently identified across studies. In several cases, the association results are in direct conflict. For example, while the studies of Straub and colleagues (2002) and Schwab and colleagues (2003) each detected association of disease with haplotypes containing the same markers, in the Schwab study the overtransmitted haplotype contained common alleles of the associated markers, and in the Straub study it contained the rare alleles. Numerous inconsistencies also exist between the results of other studies. Confirmation of a role for *DTNBP1* in SZ etiology remains to be established, and will require the identification of specific variants that alter disease susceptibility. In the present study, we have attempted further replication in two Australian samples and an Indian sample.

Materials and Methods

This study utilized an Australian Caucasian sample of 194 cases and 180 controls (McGrath et al., 2002), an Australian Caucasian family sample of 41 multiplex pedigrees (Levinson et al., 1998; Mowry et al., 2000), and a sample of 197 Indian Brahmin affected sibling pair (ASP) pedigrees (n = 85) and triads (n = 112) collected in Southern India. This study was approved by relevant institutional ethics committees, and all subjects gave written informed consent. Affected subjects satisfied a revised third edition *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III-R; American Psychiatric Association, 1987; Australian samples) or a fourth edition *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV; American Psychiatric Association, 1994; Indian sample) diagnosis of SZ or schizoaffective disorder (SA) based on structured diagnostic instruments, family informant reports and medical records. The DSM-III-R and DSM-IV are very similar with respect to SZ and SA diagnostic criteria. Diagnoses were assigned based on all available data by two independent psychiatrists, using a consensus Best Estimate Final Diagnosis. For ASP families, the inclusion criteria were that probands had a diagnosis of SZ and that at least one additional sibling was affected with either SZ or SA. Blood samples were obtained for affected siblings and available parents; if either parent was not available, unaffected siblings were also invited to participate. Subjects were excluded if their psychosis was secondary to substance abuse, if they had a known neurological disorder or severe mental retardation, or if they were unable to give informed consent.

We selected 16 SNPs from previous studies, including a number that have provided significant positive results. One SNP (rs2619538) was identified via mutation screening by Williams and colleagues (2004). Markers extended from the 5´ UTR to the 3´ UTR of *DTNBP1* and spanned a genomic region of 142 kb. Full sequence and other linked information can be found through the dbSNP Electronic-Database (Sherry et al., 1999) by using the NCBI SNP identification (rs) numbers.

Analogous to van den Bogaert et al. (2003), we calculated the power of our samples to replicate several previous findings using the genetic power calculator (Purcell et al., 2003). Based on the odds ratios reported by Schwab et al. (2003) for associated SNPs P1325 and P1635, and assuming a multiplicative inheritance model, the combined Australian Caucasian sample had a power of .89 to replicate their finding for P1325 and .86 for P1635 at a nominal significance level of .05. The Indian sample had a power of .84 to replicate the finding for P1325 and .80 for P1635. Based on the genotype relative risk (GRR) reported by van den Oord and colleagues (2003) for their associated high-risk haplotype, the Australian case-control sample had power of .97, the Australian family sample had power of 0.56 and the Indian sample had power of .99 to replicate their finding at a nominal significance level of .05, assuming a multiplicative mode of inheritance.

Genotyping was performed via a primer extension reaction and MALDI-TOF mass spectrometry (MassARRAY, Sequenom Inc., San Diego) as previously described (Bray et al., 2001; Sun et al., 2000). To identify markers with a large number of genotyping errors, the distributions of founder (pedigree samples) or control (case-control samples) genotypes were examined

Downloaded from https://www.cambridge.org/core. Open University Library, on 16 Jan 2020 at 00:38:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1375/twin.9.4.531

for deviation from Hardy-Weinberg equilibrium (HWE) using a χ^2 test with asymptotic inference, comparing genotype counts with those predicted by allelic frequencies. For pedigrees in which parental genotypes were available, error checking was also performed by examining Mendelian inheritance, using Pedmanager (M. P. Reeve and M. J. Daly, http://www.broad.mit.edu/ftp/distribution/software/pedmanager/).

Pairwise LD ($r^2 = \Delta^2$) between SNPs was calculated from founder or control alleles using LDMAX software (Abecasis & Cookson, 2000). The effective number of statistically independent SNPs (M_{eff}) was estimated using the SNPSpD World Wide Web interface (www.genepi.qimr.edu.au/general/daleN/SNPSpD), which estimates marker nonindependence using a matrix of pairwise haplotype correlation coefficients (Δ; Nyholt, 2004). The location and structure of haplotype blocks was assessed using Haploview (Barret et al., 2005), implementing the method of Gabriel et al. (2002).

Association analyses of individual SNPs and multilocus haplotypes were performed using the TDTPHASE (nuclear families) and COCAPHASE (case-control) programs (Dudbridge, 2003). TDTPHASE can perform both the transmission disequilibrium test (TDT) for phased genotype data and the haplotype-based haplotype relative risk test (HHRR) for unphased genotype data in nuclear families. When phase is known, conditional logistic regression is used to model the probabilities of marker allele transmission to affected offspring conditional on parental genotypes. When phase is unknown, unconditional logistic regression is applied to the combined likelihood of all parents and offspring, with transmitted haplotypes classed as 'cases' and untransmitted haplotypes as 'controls'. This approach is equivalent to the HHRR of Terwilliger and Ott (1992). COCAPHASE performs tests of association in case-control data using standard unconditional logistic regression. For all tests, maximum likelihood estimates of case and control haplotype frequencies are obtained using the expectation-maximization (EM) algorithm. Association analyses were performed for individual SNPs and sliding-window haplotypes constructed from adjacent markers, containing from 2 to 16 markers. Haplotypes with frequency less than 1% were dropped from analysis, as the frequencies of such haplotypes cannot be reliably estimated by the EM algorithm (Fallin & Schork, 2000). For the set of tests incorporating a given number of SNPs, global p values corrected for the multiple markers or haplotypes tested were calculated using a permutation test (10,000 replicates) in which the 'transmitted' and 'untransmitted' labels (TDTPHASE) or 'case' and 'control' labels (COCAPHASE) were reassigned. When pedigrees contained multiple affected siblings, the same randomised transmission status was used for all siblings in a sibship (*robustperm* option), thus providing a valid test of association in the presence of linkage.

Results

The Australian pedigree sample contained a total of 249 genotyped individuals, 114 of whom were affected. Of the 41 pedigrees, 28 contained two generations, 12

Table 1

Descriptive Statistics for SNPs Used in This Study

SNP No.	SNP	Position (bp)[a]	Gene region	Alleles	Australian families frequency[b]	Australian case-control frequency[c]	Indian families frequency[b]
1	P1583 (rs909706)	4338	Intron-1	G/A	.435	.332	.489
2	P1795 (rs1997679)	6304	Intron-1	G/A	.265	.345	.240
3	P1792 (rs1474605)	6997	Intron-1	A/G	.185	.219	.254
4	P1578 (rs1018381)	8139	Intron-1	C/T	.080	.098	.146
5	P1763 (rs2619522)	11,560	Intron-1	T/G	.177	.214	.237
6	P1765 (rs2619528)	15,380	Intron-3	G/A	.181	.216	.249
7	P1325 (rs1011313)	31,777	Intron-4	G/A	.106	.093	.158
8	P1635 (rs3213207)	37,107	Intron-4	A/G	.099	.124	.110
9	P1655 (rs2619539)	44,354	Intron-5	C/G	.487	.448	.595
10	rs3829893	49,572	Intron-5	G/A	.202	.168	.154
11	P1287 (rs760666)	76,088	Intron-7	G/A	.203	.201	.096
12	rs734129	76,198	Intron-7	A/G	.088	.095	.148
13	P1333 (rs742105)	92,135	Intron-7	G/A	.487	.446	.584
14	P1328 (rs742106)	140,729	Intron-9	G/A	.374	.384	.541
15	rs1047631	142,108	3′ downstream	A/G	.146	.137	.158

Note: [a] Position in base pairs, relative to the 5′ end (translation initiation site) of DTNBP1

[b] Frequency of the second allele in pedigree founders

[c] Frequency of the second allele in controls.

Downloaded from https://www.cambridge.org/core. Open University Library, on 16 Jan 2020 at 00:38:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1375/twin.9.4.531

contained three generations and 1 contained four generations. The Indian sample contained a total of 668 genotyped individuals, 329 of whom were affected.

Table 1 shows descriptive statistics for each SNP used in the current study. The third column shows the position of each marker in base pairs, relative to the 5′ end of *DTNBP1*, which contains the translation initiation site. The fourth column indicates the gene region within which individual SNPs are located. Allele frequencies were calculated in controls (case-control sample) or pedigree founders (family samples). For three of the SNPs (P1655, P1333, P1328), the major and minor alleles were swapped for Australian and Indian samples. In addition, comparison of SNP allele frequencies revealed significant differences between the Australian and Indian samples for a number of the SNPs (Australian founders vs. Indian founders: rs2619538 [$p = 4.5 \times 10^{-7}$], P1287 [$p = 6.2 \times 10^{-5}$], P1328 [$p = 5.1 \times 10^{-5}$]; Australian controls vs. Indian founders: rs2619538 [$p = 1.8 \times 10^{-18}$], P1583 [$p = 4.7 \times 10^{-7}$], P1795 [$p = .00019$], P1655 [$p = 2.9 \times 10^{-6}$], P1287 [$p = 1.8 \times 10^{-6}$], P1333 [$p = 1.1 \times 10^{-5}$], P1328 [$p = 5.3 \times 10^{-7}$]). No significant differences were observed between the Australian founder and Australian control allele frequencies. For these reasons, the Indian sample was analyzed separately in all association tests. The genotype distributions of all SNPs were in HWE in unrelated controls (case-control sample) and pedigree founders except rs2619538, which deviated significantly from HWE in all three samples (Australian pedigrees: $p = .0014$; Australian case-control: $p = 1.44 \times 10^{-7}$; Indian pedigrees: $p = 1.9 \times 10^{-20}$). This marker was consequently dropped from all further analyses. Subsequent investigations by our laboratory could not identify the reasons for this deviation. While genotyping error is a potential explanation, Li et al. (2005) also reported significant deviation of this marker's genotype frequencies from HWE and suggested the possible existence of a different SNP polymorphism under one of the primers used to detect rs2619538, which could easily bias allelic and genotypic frequency estimates.

The matrix of pairwise LD values, calculated using founder or control alleles, and expressed as the squared correlation coefficient for a 2 × 2 table (r^2) is shown in Figure 1. Values of r^2 are shown for the combined Australian sample (above diagonal) and the Indian sample (below diagonal). The LD structure is broadly similar for the two samples. For adjacent SNPs, high pairwise LD ($r^2 > .8$) was detected only between P1763 and P1765. High LD was also observed between four nonadjacent SNP pairs: P1792 and P1763; P1792 and P1765; P1578 and rs734129 and P1655 and P1333. Mean pairwise r^2 for adjacent SNPs was .2 for the combined Australian sample and .23 for the Indian sample. Using the approach of Nyholt (2004), the 15 SNPs equated to approximately 10.1 effectively independent SNPs in both the combined Australian and the Indian samples. Haplotype blocks for the Australian and Indian samples are shown in Figure 2. In the combined Australian sample, three haplotype blocks were identified, while in the Indian sample, five distinct blocks were observed.

Table 2 shows association analysis results for individual SNPs. For the combined Australian case-control and pedigree sample, SNPs were tested for association with SZ using a χ^2 statistic calculated from a 2 × 2 contingency table which combined the case-control case allele counts with familial HHRR transmitted allele counts and the case-control control allele counts with familial HHRR untransmitted allele counts. No SNP showed significant association with SZ ($p < .05$) in any of the samples. Virtually indistinguishable results were

Map	SNP #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
4338	1		.26	.12	.06	.13	.13	.18	.05	.58	.12	.14	.06	.65	.24	.08
6304	2	.26		.02	.21	.03	.02	.05	.05	.37	.10	.56	.21	.40	.03	.06
6997	3	.26	.19		.39	.95	.92	.03	.49	.01	.05	.07	.39	.01	.01	.23
8139	4	.15	.49	.5		.42	.39	.01	0	.08	.02	.03	.99	.08	0	.01
11560	5	.25	.17	.88	.42		.96	.02	.42	.01	.05	.07	.42	.01	.01	.24
15380	6	.23	.2	.88	.49	.84		.03	.42	.01	.05	.07	.39	.01	.01	.26
31777	7	.19	.04	.06	.03	.05	.06		0	.13	.01	.03	.01	.11	.01	.01
37107	8	.11	.02	.35	0	.36	.29	.02		.14	.03	.03	0	.14	0	.52
44354	9	.62	.32	.02	.13	.01	.01	.13	.06		.16	.21	.08	.91	.23	.04
49572	10	.16	.04	.05	.03	.05	.05	.03	.02	.25		.06	.02	.19	.09	.03
76088	11	.11	.35	.02	.02	.01	.03	.02	.01	.16	.02		.03	.22	.04	.03
76198	12	.14	.5	.51	1	.43	.5	.03	0	.14	.03	.02		.08	0	.01
92135	13	.65	.4	.04	.21	.03	.03	.13	.06	.89	.22	.14	.22		.21	.04
140729	14	.34	.03	.09	.01	.08	.06	0	.12	.22	.05	.04	.01	.17		.01
142108	15	.13	.05	.2	.02	.22	.19	.01	.62	.01	.03	.02	.02	.01	.16	

Figure 1

Linkage disequilibrium (LD: r^2) between SNP pairs for the combined Australian (above diagonal) and the Indian (below diagonal) samples.

Downloaded from https://www.cambridge.org/core. Open University Library, on 16 Jan 2020 at 00:38:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1375/twin.9.4.531

Table 2
Association Analysis Results for Individual SNPs

		Australian families		Australian case-control		Australian combined		Indian families	
SNP	Alleles	T/NT[a]	*p*	T/NT[a]	*p*	T/NT[a]	*p*	T/NT[a]	*p*
P1583 (rs909706)	G/A	70/68	.799	232/259	.507	302/327	.630	193/204	.430
P1795 (rs1997679)	G/A	104/102	.777	246/254	.405	350/356	.364	314/315	.930
P1792 (rs1474605)	A/G	118/117	.866	294/303	.223	412/420	.237	310/310	1.000
P1578 (rs1018381)	C/T	134/132	.606	324/350	.707	458/482	.549	355/355	1.000
P1763 (rs2619522)	T/G	117/118	.860	296/305	.213	413/423	.291	308/299	.418
P1765 (rs2619528)	G/A	115/112	.595	295/304	.218	410/416	.171	289/292	.789
P1325 (rs1011313)	G/A	124/127	.578	327/352	.958	451/479	.773	340/341	.919
P1635 (rs3213207)	A/G	134/134	1.000	326/338	.191	460/472	.245	357/359	.800
P1655 (rs2619539)	C/G	73/69	.633	203/214	.734	276/283	.606	226/244	.186
rs3829893	G/A	116/113	.618	292/323	.445	408/436	.689	350/343	.471
P1287 (rs760666)	G/A	115/110	.430	281/298	.683	396/408	.444	375/379	.608
rs734129	A/G	138/138	1.000	327/351	.862	465/489	.855	360/356	.661
P1333 (rs742105)	G/A	65/62	.715	197/215	.850	262/277	.991	220/235	.270
P1328 (rs742106)	G/A	90/87	.697	236/239	.261	326/326	.238	200/217	.211
rs1047631	A/G	130/131	.857	316/335	.558	446/466	.678	338/334	.676

Note: [a]The number of transmitted/nontransmitted alleles of the first allele in column 2.

obtained using phase known (TDT test) or phase unknown (HHRR test) data for pedigree samples. Results shown for family samples were obtained using the HHRR test.

We also performed association analyses for sliding-window haplotypes, testing all possible haplotypes encompassing 2 to 15 adjacent markers. This analysis naturally included assessment of identified haplotype blocks. A number of haplotypes showed association with disease at a nominally significant level ($p < .05$) in Australian pedigrees, and one haplotype was nominally associated with disease in Indian pedigrees (Figure 2), but these associations did not remain significant following correction for testing multiple SNPs and sliding-window haplotypes.

A number of previous studies have detected significant associations of *DTNBP1* variants with SZ in families containing multiple affected individuals (Schwab et al., 2003; Straub et al., 2002; Tang et al., 2003; van den Bogaert et al., 2003; van den Oord et al., 2003). Van den Bogaert et al. (2003) observed stronger evidence for association in family-history positive cases selected from their Swedish case-control cohort, and suggested that *DTNBP1* may be particularly involved in the development of SZ in cases with a familial loading of disease. To test for the presence of this effect in the Australian case-control sample, we performed analyses only in cases with a family history of SZ ($n = 54$). Nominally significant association ($p < .05$) was detected for a number of SNPs and haplotypes (data not shown). The strongest evidence for association was detected for a two-marker haplotype containing P1795 and P1792 ($p = .007$). However, no results remained significant following correction for testing all markers and sliding-window haplotypes.

Discussion

The current study was unable to replicate previously reported association of *DTNBP1* variants with SZ. No SNP showed even nominally significant association ($p < .05$) in any of the three samples or the combined Australian sample, despite our samples being comparable in size to those from previous positive studies and power calculations indicating that both the Indian and combined Australian samples had better than 80% power to replicate the results for P1325 and P1635 reported by Schwab et al. (2003). The strongest evidence for association was observed in the subset of Australian cases with a family history of SZ, even though this sample was less than one third the size of the unselected Australian cases. This result may reflect higher heritability and thus, relative contribution of all genetic factors to the phenotype in this subgroup, due to familial loading of the disease. In this sample, we were able to replicate association of P1325 and P1635 at a nominal p value of .05. However, these results failed to meet the criteria for global significance following correction for multiple testing. Furthermore, it is difficult to reconcile these results given the absence of associated SNPs in our pedigree (ASP) samples, which by their very nature have a family history of SZ.

Our study may have been unable to detect significant association of SZ with *DTNBP1* variants because of insufficient LD between marker and trait loci. Our power analyses assumed that the allele tested was the actual susceptibility allele, but in any association test,

Downloaded from https://www.cambridge.org/core. Open University Library, on 16 Jan 2020 at 00:38:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1375/twin.9.4.531

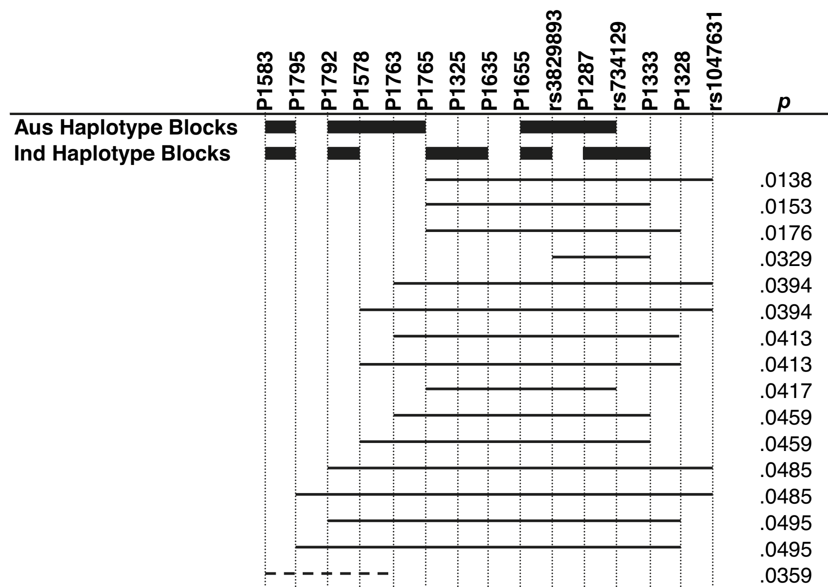

Figure 2

Haplotype block structure in Australian and Indian samples and *DTNBP1* haplotypes showing nominally significant association to SZ ($p < .05$).

Note: The upper two rows show haplotype block structure in combined Australian (Aus) and Indian (Ind) samples. The remaining rows show *DTNBP1* haplotypes with nominally significant association to SZ ($p < .05$) in combined Australian (unbroken lines) and Indian (broken line) samples.

the a priori probability of this being the case is low. In most cases, it is hoped that one or more SNPs are in LD with the true risk variant(s), and the power of the test is then highly dependent upon the extent of LD (Muller-Myhsok & Abel, 1997). As LD between the marker and trait locus decreases below 1, the sample size required to detect the association with approximately equivalent power increases by a factor of $1/r^2$ (Pritchard & Przeworski, 2001). It has been suggested that values of r^2, if greater than about .33, may indicate sufficient LD to be useful for association mapping (Ardlie et al., 2002). In both our combined Australian and Indian samples, r^2 values between flanking markers exceed .33 for just 3 of the 14 intermarker intervals, spanning a mere 8.4 kb of the 140 kb *DTNBP1* gene. This pattern of LD may have substantially reduced our power to capture allelic variation across most of the *DTNBP1* gene. Interestingly, previous studies are likely to have been more impacted by this same limitation. That is, only one study (Williams et al., 2004) used a higher SNP density than ours and although the study by Williams and colleagues genotyped 20 SNPs, only 9 were individually genotyped and thus available for haplotype analyses (the other 11 SNPs were genotyped only in a pooled case-control analysis). Other studies' detection of significant association with even fewer SNPs may result partly from sample differences in LD structure. While LD patterns between our study and many others were broadly similar, subtle differences were also evident. For example, Straub et al. (2002) reported an r^2 of .3 between P1765 and P1325, while the corresponding values in our samples were .03 (Australian sample) and .06 (Indian sample).

Other possible explanations for not replicating previous findings are the presence of locus and/or allelic heterogeneity at the *DTNBP1* locus. With *locus* heterogeneity, only a proportion of subjects in any sample harbour mutations at a given locus. For example, by chance the proportion of subjects in our samples with SZ resulting from allelic variation at *DTNBP1* may have been too small to produce significant frequency differences between case and control groups for the given samples sizes. With *allelic* heterogeneity, multiple risk-conferring mutations exist within a susceptibility gene, each having arisen upon its own haplotype background. Slager and colleagues (2000) have shown that the presence of multiple susceptibility alleles at a test locus can substantially reduce the power of association mapping, even when the allele tested is the actual susceptibility allele. The extent of allelic heterogeneity in SZ is unknown, but given the apparent absence of major genes for the disorder, the selective disadvantage associated with individual mutations is likely to be weak and allelic heterogeneity may thus be relatively common. The presence of allelic heterogeneity at *DTNBP1*, combined with the low a priori probability of any study directly testing a given susceptibility allele, may explain both the inconsistencies of previous results, and our inability to replicate any of them.

A further explanation for our negative findings is that a single *DTNBP1* susceptibility allele is inherited on a variety of different haplotypes. Samples may differ in the prevalence of individual haplotypes, producing diverse results across different samples. Our samples may have contained multiple associated haplotypes, reducing the apparent contribution of individual haplotypes. Alternatively, our samples may have contained a

Downloaded from https://www.cambridge.org/core. Open University Library, on 16 Jan 2020 at 00:38:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1375/twin.9.4.531

higher prevalence of rare haplotypes, precluding their detection with available sample sizes.

It is also possible that the previously reported effect sizes used for our power calculations were not appropriate for the Australian and Indian samples. The effect size of a risk allele or haplotype may vary between individuals or populations according to different frequencies of environmental factors or other genetic variants with which the susceptibility variant interacts. Depending upon the mechanism by which *DTNBP1* variants influence disease risk, such individual variation may be quite pronounced. Systematic mutation detection has failed to identify nonsynonymous alleles associated with disease risk (Williams et al., 2004), suggesting that disease susceptibility may be associated with altered *DTNBP1* mRNA expression levels. Supporting this hypothesis is the identification of *cis*-regulatory polymorphisms that influence *DTNBP1* expression in the human brain (Bray et al., 2003; Bray et al., 2005), combined with evidence of reduced *DTNBP1* mRNA and protein expression in the dorsolateral prefrontal cortex (Weickert et al., 2004) and hippocampus (Talbot et al., 2004) of individuals with SZ. If *cis*-acting polymorphisms are the true SZ susceptibility alleles, considerable variability of their effects may be expected across individuals and populations, as the functional effects of *cis*-regulatory polymorphisms are regulated by diverse factors including variation within other genomic sequences, the concentration and distribution of exogenous factors, and national origin (Rockman & Wray, 2002). Our inability to detect association with *DTNBP1* variants may thus reflect differences between our samples and previously studied samples with respect to ethnicity, specific environmental exposures and interacting genetic effects, which could act to reduce the effect size associated with a causal variant(s) in our samples.

In conclusion, our failure to detect significant association of *DTNBP1* variants with SZ may reflect the combined effects of low intermarker LD, sample differences in LD, the presence of small-effect variants, individual and population differences in effect size(s), locus heterogeneity, allelic heterogeneity and/or the presence of a single susceptibility allele upon a diversity of haplotypes. It is also possible that some previous results were false positives. While a large body of evidence supports the existence of SZ susceptibility variants at the *DTNBP1* locus, confirmation of a role for *DTNBP1* in SZ etiology will ultimately require the identification of specific susceptibility variants. Future studies should achieve higher power and efficiency using SNPs (tagSNPs) which 'tag' sets of correlated SNPs identified using available high density genotype data. For example, utilizing Caucasian HapMap phase II (www.hapmap.org) genotype data spanning the *DTNBP1* locus, from 20 kb upstream (rs9383069) to 10 kb downstream (rs2235260) of the 5′ and 3′ UTR respectively, an r^2 threshold of .8 and SNPs with a minor allele frequency (MAF) of at least .05, Haploview (Barret et al., 2005) identifies 33 nonredundant tagSNPs. When tagSNPs were forced to include the SNPs typed in the current study which were also available in HapMap data (13/15 SNPs), 38 SNPs were required, indicating that an additional 23 to 25 SNPs are required to satisfactorily assess all nonredundant variation within *DTNBP1* and its

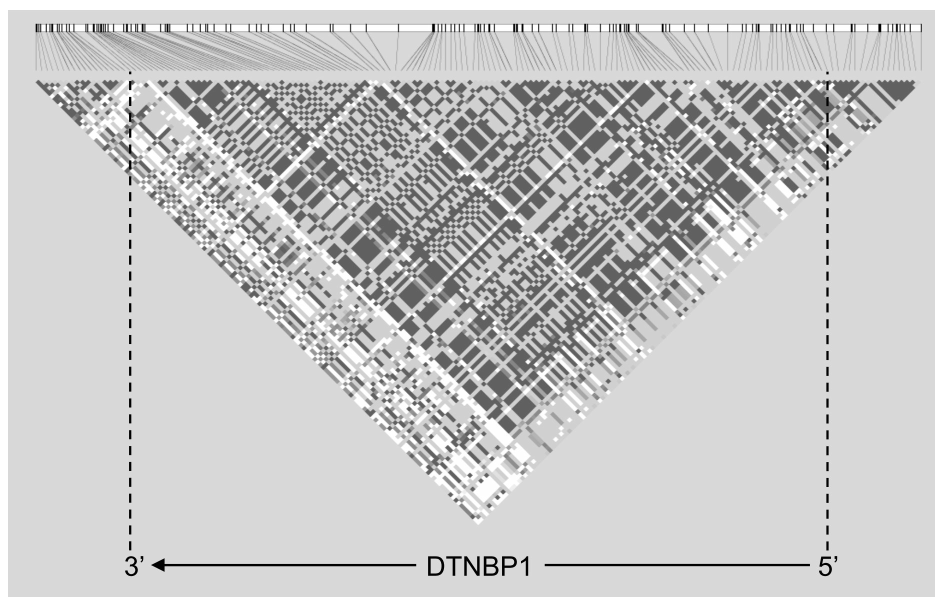

Figure 3

Haploview plot showing pairwise LD for HapMap phase II SNPs spanning *DTNBP1*, genotyped in the CEPH population.

Note: White (nonsignificant incomplete LD): LOD < 2, D′ < 1; Light grey (nonsignificant complete LD): LOD < 2, D′ = 1; Dark Grey (significant incomplete LD): LOD ≥ 2, D′ < 1; Black (significant complete LD): LOD ≥ 2, D′ = 1.

Downloaded from https://www.cambridge.org/core. Open University Library, on 16 Jan 2020 at 00:38:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1375/twin.9.4.531

potential regulatory regions for association with SZ. Indeed, examination of the LD plot (Figure 3) reveals a highly variable and complex pattern of LD with significant LD existing between markers spanning the entire region. The use of such comprehensive SNP sets, will substantially improve the power and consistency of future studies aiming to identify *DTNBP1* variants associated with SZ.

Acknowledgments

The authors gratefully acknowledge the participation of individuals and family members. We also acknowledge Heather Smith, Cheryl Filippich and Marilyn Walters for excellent technical assistance in the preparation of samples. Supported by the Australian National Health and Medical Research Council grants 910234, 941087, 971095 (BJM) and 143027 (BJM, RT); The Commonwealth of Australia Department of Health and Aged Care Grant RADGAC MH 145 Project 98/144 (JJM); Queensland Department of Health Grant 98/DHQ1012G (BJM); NIMH Grants R01-MH045097, KO2-01207 and K24-MH64197 (DFL).

References

Abecasis, G. R., & Cookson, W. O. (2000). GOLD - Graphical overview of linkage disequilibrium. *Bioinformatics, 16*, 182–183.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., Rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Ardlie, K. G., Kruglyak, L., & Seielstad, M. (2002). Patterns of linkage disequilibrium in the human genome. *Nature reviews. Genetics, 3*, 299–309.

Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics, 21*, 263–265.

Bray, M. S., Boerwinkle, E., & Doris, P. A. (2001). High-throughput multiplex SNP genotyping with MALDI-TOF mass spectrometry: Practice, problems and promise. *Human Mutation, 17*, 296–304.

Bray, N. J., Buckland, P. R., Owen, M. J., & O'Donovan, M. C. (2003). Cis-acting variation in the expression of a high proportion of genes in human brain. *Human Genetics, 113*, 149–153.

Bray, N. J., Preece, A., Williams, N. M., Moskvina, V., Buckland, P. R., Owen, M. J., & O'Donovan, M. C. (2005). Haplotypes at the dystrobrevin binding protein 1 (*DTNBP1*) gene locus mediate risk for schizophrenia through reduced DTNBP1 expression. *Human Molecular Genetics, 14*, 1947–1954.

Dudbridge, F. (2003). Pedigree disequilibrium tests for multilocus haplotypes. *Genetic Epidemiology, 25*, 115–121.

Fallin, D., & Schork, N. J. (2000). Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-maximization algorithm for unphased diploid genotype data. *American Journal of Human Genetics, 67*, 947–959.

Funke, B., Finn, C. T., Plocik, A. M., Lake, S., DeRosse, P., Kane, J. M., Kucherlapati, R., & Malhotra, A. K. (2004). Association of the *DTNBP1* locus with schizophrenia in a U.S. population. *American Journal of Human Genetics, 75*, 891–898.

Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., Liu-Cordero, S. N., Rotimi, C., Adeyemo, A., Cooper, R., Ward, R., Lander, E. S., Daly, M. J., & Altshuler, D. (2002). The structure of haplotype blocks in the human genome. *Science, 296*, 2225–2229.

Kirov, G., Ivanov, D., Williams, N. M., Preece, A., Nikolov, I., Milev, R., Koleva, S., Dimitrova, A., Toncheva, D., O'Donovan, M. C., & Owen, M. J. (2004). Strong evidence for association between the dystrobrevin binding protein 1 gene (*DTNBP1*) and schizophrenia in 488 parent-offspring trios from Bulgaria. *Biological Psychiatry, 55*, 971–975.

Levinson, D. F., Mahtani, M. M., Nancarrow, D. J., Brown, D. M., Kruglyak, L., Kirby, A., Hayward, N. K., Crowe, R. R., Andreasen, N. C., Black, D. W., Silverman, J. M., Endicott, J., Sharpe, L., Mohs, R. C., Siever, L. J., Walters, M. K., Lennon, D. P., Jones, H. L., Nertney, D. A., Daly, M. J., Gladis, M., & Mowry, B. J. (1998). Genome scan of schizophrenia. *American Journal of Psychiatry, 155*, 741–750.

Lewis, C. M., Levinson, D. F., Wise, L. H., DeLisi, L. E., Straub, R. E., Hovatta, I., Williams, N. M., Schwab, S. G., Pulver, A. E., Faraone, S. V., Brzustowicz, L. M., Kaufmann, C. A., Garver, D. L., Gurling, H. M., Lindholm, E., Coon, H., Moises, H. W., Byerley, W., Shaw, S. H., Mesen, A., Sherrington, R., O'Neill, F. A., Walsh, D., Kendler, K. S., Ekelund, J., Paunio, T., Lonnqvist, J., Peltonen, L., O'Donovan, M. C., Owen, M. J., Wildenauer, D. B., Maier, W., Nestadt, G., Blouin, J. L., Antonarakis, S. E., Mowry, B. J., Silverman, J. M., Crowe, R. R., Cloninger, C. R., Tsuang, M. T., Malaspina, D., Harkavy-Friedman, J. M., Svrakic, D. M., Bassett, A. S., Holcomb, J., Kalsi, G., McQuillin, A., Brynjolfson, J., Sigmundsson, T., Petursson, H., Jazin, E., Zoega, T., & Helgason, T. (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. *American Journal of Human Genetics, 73*, 34–48.

Li, T., Zhang, F., Liu, X., Sun, X., Sham, P. C., Crombie, C., Ma, X., Wang, Q., Meng, H., Deng, W., Yates, P., Hu, X., Walker, N., Murray, R. M., St Clair, D., & Collier, D. A. (2005). Identifying potential risk haplotypes for schizophrenia at the *DTNBP1* locus in Han Chinese and Scottish populations. *Molecular Psychiatry, 10*, 1037–1044.

McGrath, J., El-Saadi, O., Grim, V., Cardy, S., Chapple, B., Chant, D., Lieberman, D., & Mowry, B. (2002). Minor

Downloaded from https://www.cambridge.org/core. Open University Library, on 16 Jan 2020 at 00:38:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1375/twin.9.4.531

physical anomalies and quantitative measures of the head and face in patients with psychosis. *Archives of General Psychiatry, 59*, 458–464.

Mowry, B. J., Ewen, K. R., Nancarrow, D. J., Lennon, D. P., Nertney, D. A., Jones, H. L., O'Brien, M. S., Thornley, C. E., Walters, M. K., Crowe, R. R., Silverman, J. M., Endicott, J., Sharpe, L., Hayward, N. K., Gladis, M. M., Foote, S. J., & Levinson, D. F. (2000). Second stage of a genome scan of schizophrenia: Study of five positive regions in an expanded sample. *American Journal of Medical Genetics, 96*, 864–869.

Muller-Myhsok, B., & Abel, L. (1997). Genetic analysis of complex diseases. *Science, 275*, 1328–1329; author reply 1329–1330.

Numakawa, T., Yagasaki, Y., Ishimoto, T., Okada, T., Suzuki, T., Iwata, N., Ozaki, N., Taguchi, T., Tatsumi, M., Kamijima, K., Straub, R. E., Weinberger, D. R., Kunugi, H., & Hashimoto, R. (2004). Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia. *Human Molecular Genetics, 13*, 2699–2708.

Nyholt, D. R. (2004). A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *American Journal of Human Genetics, 74*, 765–769.

Pritchard, J. K., & Przeworski, M. (2001). Linkage disequilibrium in humans: Models and data. *American Journal of Human Genetics, 69*, 1–14.

Purcell, S., Cherny, S. S., & Sham, P. C. (2003). Genetic power calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics, 19*, 149–150.

Rockman, M. V., & Wray, G. A. (2002). Abundant raw material for cis-regulatory evolution in humans. *Molecular Biology and Evolution, 19*, 1991–2004.

Schwab, S. G., Knapp, M., Mondabon, S., Hallmayer, J., Borrmann-Hassenbach, M., Albus, M., Lerer, B., Rietschel, M., Trixler, M., Maier, W., & Wildenauer, D. B. (2003). Support for association of schizophrenia with genetic variation in the 6p22.3 gene, dysbindin, in sib-pair families with linkage and in an additional sample of triad families. *American Journal of Human Genetics, 72*, 185–190.

Sherry, S. T., Ward, M., & Sirotkin, K. (1999). dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Research, 9*, 677–679.

Slager, S. L., Huang, J., & Vieland, V. J. (2000). Effect of allelic heterogeneity on the power of the transmission disequilibrium test. *Genetic Epidemiology, 18*, 143–156.

Straub, R. E., Jiang, Y., MacLean, C. J., Ma, Y., Webb, B. T., Myakishev, M. V., Harris-Kerr, C., Wormley, B., Sadek, H., Kadambi, B., Cesare, A. J., Gibberman, A., Wang, X., O'Neill, F. A., Walsh, D., & Kendler, K. S. (2002). Genetic variation in the 6p22.3 gene *DTNBP1*, the human ortholog of the mouse Dysbindin Gene, is associated with schizophrenia. *American Journal of Human Genetics, 71*, 337–348.

Sun, X., Ding, H., Hung, K., & Guo, B. (2000). A new MALDI-TOF based mini-sequencing assay for genotyping of SNPS. *Nucleic Acids Research, 28*, E68.

Talbot, K., Eidem, W. L., Tinsley, C. L., Benson, M. A., Thompson, E. W., Smith, R. J., Hahn, C. G., Siegel, S. J., Trojanowski, J. Q., Gur, R. E., Blake, D. J., & Arnold, S. E. (2004). Dysbindin-1 is reduced in intrinsic, glutamatergic terminals of the hippocampal formation in schizophrenia. *The Journal of Clinical Investigation, 113*, 1353–1363.

Tang, J. X., Zhou, J., Fan, J. B., Li, X. W., Shi, Y. Y., Gu, N. F., Feng, G. Y., Xing, Y. L., Shi, J. G., & He, L. (2003). Family-based association study of *DTNBP1* in 6p22.3 and schizophrenia. *Molecular Psychiatry, 8*, 717–718.

Terwilliger, J. D., & Ott, J. (1992). A haplotype-based 'haplotype relative risk' approach to detecting allelic associations. *Human Heredity, 42*, 337–346.

van den Bogaert, A., Schumacher, J., Schulze, T. G., Otte, A. C., Ohlraun, S., Kovalenko, S., Becker, T., Freudenberg, J., Jonsson, E. G., Mattila-Evenden, M., Sedvall, G. C., Czerski, P. M., Kapelski, P., Hauser, J., Maier, W., Rietschel, M., Propping, P., Nothen, M. M., & Cichon, S. (2003). The *DTNBP1* (dysbindin) gene contributes to schizophrenia, depending on family history of the disease. *American Journal of Human Genetics, 73*, 1438–1443.

van den Oord, E. J., Sullivan, P. F., Jiang, Y., Walsh, D., O'Neill, F. A., Kendler, K. S., & Riley, B. P. (2003). Identification of a high-risk haplotype for the dystrobrevin binding protein 1 (*DTNBP1*) gene in the Irish study of high-density schizophrenia families. *Molecular Psychiatry, 8,* 499–510.

Weickert, C. S., Straub, R. E., McClintock, B. W., Matsumoto, M., Hashimoto, R., Hyde, T. M., Herman, M. M., Weinberger, D. R., & Kleinman, J. E. (2004). Human dysbindin (*DTNBP1*) gene expression in normal brain and in schizophrenic prefrontal cortex and midbrain. *Archives of General Psychiatry, 61*, 544–555.

Williams, N. M., Preece, A., Morris, D. W., Spurlock, G., Bray, N. J., Stephens, M., Norton, N., Williams, H., Clement, M., Dwyer, S., Curran, C., Wilkinson, J., Moskvina, V., Waddington, J. L., Gill, M., Corvin, A. P., Zammit, S., Kirov, G., Owen, M. J., & O'Donovan, M. C. (2004). Identification in 2 independent samples of a novel schizophrenia risk haplotype of the dystrobrevin binding protein gene (*DTNBP1*). *Archives of General Psychiatry, 61*, 336–344.

Downloaded from https://www.cambridge.org/core. Open University Library, on 16 Jan 2020 at 00:38:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1375/twin.9.4.531