RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics PART B

Identification of Neuroglycan C and Interacting Partners as Potential Susceptibility Genes for Schizophrenia in a Southern Chinese Population

**Hon-Cheong So,[1] Pui Y. Fong,[2] Ronald Y.L. Chen,[1] Tomy C.K. Hui,[1] Mandy Y.M. Ng,[2] Stacey S. Cherny,[1,2,4] William W.M. Mak,[2,3,5] Eric F.C. Cheung,[6] Raymond C.K. Chan,[1,7,8] Eric Y.H. Chen,[1] Tao Li,[9] and Pak C. Sham[1,2,4]*

[1]Department of Psychiatry, University of Hong Kong, Hong Kong SAR, China

[2]Genome Research Centre, University of Hong Kong, Hong Kong SAR, China

[3]Department of Biochemistry, University of Hong Kong, Hong Kong SAR, China

[4]The State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong SAR, China

[5]Hong Kong Biotechnology Education Resource Centre, Hong Kong SAR, China

[6]Castle Peak Hospital, Hong Kong SAR, China

[7]Neuropsychology and Applied Cognitive Neuroscience Laboratory, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[8]Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[9]Division of Psychological Medicine and SGDP Centre, Institute of Psychiatry, King's College, London, UK

Received 3 July 2008; Accepted 2 March 2009

Chromosome 3p was reported by previous studies as one of the regions showing strong evidence of linkage with schizophrenia. We performed a fine-mapping association study of a 6-Mb high-LD and gene-rich region on 3p in a Southern Chinese sample of 489 schizophrenia patients and 519 controls to search for susceptibility genes. In the initial screen, 4 SNPs out of the 144 tag SNPs genotyped were nominally significant ($P < 0.05$). One of the most significant SNPs (rs3732530, $P = 0.0048$) was a non-synonymous SNP in the neuroglycan C (*NGC*, also known as *CSPG5*) gene, which belongs to the neuregulin family. The gene prioritization program Endeavor ranked *NGC* 8th out of the 129 genes in the 6-Mb region and the highest among the genes within the same LD block. Further genotyping of *NGC* revealed 3 more SNPs to be nominally associated with schizophrenia. Three other genes (*NRG1*, *ErbB3*, *ErbB4*) involved in the neuregulin pathways were subsequently genotyped. Interaction analysis by multifactor dimensionality reduction (MDR) revealed a significant two-SNP interaction between *NGC* and *NRG1* ($P = 0.015$) and three-SNP interactions between *NRG1* and *ErbB4* ($P = 0.009$). The gene *NGC* is exclusively expressed in the brain. It is implicated in neurodevelopment in rats and was previously shown to promote neurite outgrowth. Methamphetamine, a drug that may induce psychotic symptoms, was reported to alter the expression of *NGC*. Taken together, these results suggest that *NGC* may be a novel candidate gene, and neuregulin signaling pathways may play an important role in schizophrenia. © 2009 Wiley-Liss, Inc.

Key words: neuregulin-1; *CSPG5*; *ErbB4*; genetic association; interaction

How to Cite this Article:
So H-C, Fong PY, Chen RYL, Hui TCK, Ng MYM, Cherny SS, Mak WWM, Cheung EFC, Chan RCK, Chen EYH, Li T, Sham PC. 2010. Identification of Neuroglycan C and Interacting Partners as Potential Susceptibility Genes for Schizophrenia in a Southern Chinese Population.
Am J Med Genet Part B 153B:103–113.

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: University of Hong Kong Strategic Research Theme on Genomics; Grant sponsor: Hong Kong Research Grants Council General Research Fund; Grant number: 757905.

Hon-Cheong So and Pui Y. Fong contributed equally to this work.

*Correspondence to:
Pak C. Sham, Department of Psychiatry, 10/F Laboratory Block, LKS Faculty of Medicine, University of Hong Kong, Pokfulam, Hong Kong SAR, China. E-mail: pcsham@hkucc.hku.hk

Published online 14 April 2009 in Wiley InterScience (www.interscience.wiley.com)
DOI 10.1002/ajmg.b.30961

INTRODUCTION

Schizophrenia (SZ [MIM 181500]) is a severe, common neurodevelopmental psychiatric disorder affecting nearly 1% of the human population [Sawa and Snyder, 2002]. Although the disorder has high heritability (around 80% as shown by twin studies) [Cardno and Gottesman, 2000], the overall evidence from family and linkage studies indicates complex inheritance with the involvement of multiple genetic loci. Linkage and association studies have so far identified a few susceptibility genes of modest effect, including neuregulin 1 (*NRG1*), dysbindin (*DTNBP1*), and D-amino-acid oxidase (*DAAO*) [Harrison and Weinberger, 2005].

Chromosome 3p is one region with evidence of linkage with SZ, first implicated by a genome-wide linkage scan [Pulver et al., 1995] which identified 3p and 8p to have the two highest LOD scores (2.34 and 3.00, respectively). Another linkage study on targeted chromosomal regions yielded a two-point LOD score of ~1.5 in support of linkage to 3p [Maziade et al., 2001]. The best evidence of linkage to 3p came from the meta-analysis performed by Lewis et al. [2003]. This meta-analysis included data from 20 schizophrenia genome scans and employed a rank-based approach to combine evidence from different scans. The second bin on chromosome 3p showed strong statistical significance for linkage, ranking 2nd in the unweighted analysis and 3rd in the weighted analysis (in which each study was weighted by the square root of its sample size).

In this study, we attempt to screen chromosome 3p for susceptibility genes for SZ in a Southern Chinese sample of 489 schizophrenia patients and 519 controls. We decided to concentrate on the 47–53 Mb region of 3p, which has high levels of linkage disequilibrium (LD) and gene density (see Fig. 1). The region was interrogated by tag SNPs selected using a clustering algorithm [Ao et al., 2005] which has been further improved by using both LD and biological information in the selection process [Sham et al., 2007]. Using these tools, we have conducted a high-density association study of this 47–53 Mb region of 3p to search for SZ susceptibility genes. Candidate genes that contain significant SNPs from this screen were then subjected to further exploration which included the genotyping of additional SNPs in these genes and their interacting partners.

MATERIALS AND METHODS

Case–Control Sample

Four hundred eighty-nine patients with schizophrenia diagnosed by DSM-IV were recruited from mental hospitals in Hong Kong. Five hundred nineteen controls without serious mental disorders were recruited from the community. Details of sample collection can be found elsewhere [Chen et al., 2001].

Assay Design and Genotyping

We employed the Sequenom MassARRAY™ system for assay design and genotyping, using SpectroDESIGNER software version 2.0.0.17. Genotyping was carried out by primer extension and MALDI-TOF mass spectrometry. For quality control, 20 samples were duplicated and 4 negative controls were included in each 384-well plate. A successful assay was defined as one with (i) $\geq$ 90% of genotyping calls; and (ii) control genotypes not significantly deviating from the Hardy–Weinberg equilibrium ($P > 0.01$).

Initial Association Screen

For tag SNP selection, we downloaded genotype data on 1693 SNPs in the 3p region typed on the HapMap Asian population (JPT and CHB), data release 2005–03_16a_phaseI (Build 34). The CLUSTAG [Ao et al., 2005] program, which employs an agglomerative clustering algorithm, was used for the selection of tag SNPs, with the R^2 threshold set at 0.8. If a SNP failed in the assay design, it was replaced by the second best SNP within the same cluster. From the 814 common SNPs (those of minor allele frequency >5%), we selected 218 tag SNPs and successfully designed genotyping assays for 214 of them. Of these, 162 SNPs genotyped passed the quality control criteria. The average genotype call rate was 96%. Eighteen

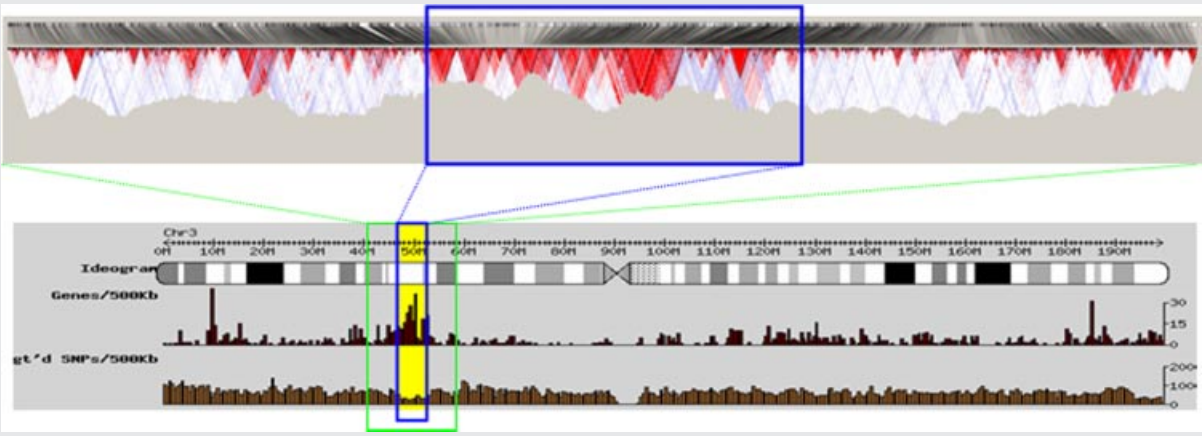

FIG. 1. The top figure is a LD plot of 41–59 Mb on chromosome 3. The plot shows a high LD region in the middle (47–53 Mb, boxed region). The figure below is the ideogram of chromosome 3 with plots of gene density and SNP density. The 6-Mb high LD region is highlighted and shows a high gene density. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

SNPs with minor allele frequency under 1% were discarded, leaving 144 markers for the association analyses. To ensure that the remaining markers still captured the variation in the 3p region adequately, we reassessed the coverage of these 144 tag SNPs using Haploview. The proportion of the genetic variation captured was calculated with reference to HapMap release 21. We employed aggressive tagging with 2- and 3-marker haplotype tags. The original 144 tag SNPs captured 76% of the alleles in the linkage region at r-squared $\geq$0.8. The average R^2 is 0.948.

Statistical Analyses for Association

The programs WHAP (ver 2.06) [Saxena et al., 2007] (http://pngu.mgh.harvard.edu/~purcell/whap/) and UNPHASED (ver 3.09) [Dudbridge, 2003] were used to screen all the 144 markers for both allelic and haplotype-based association. Genotype data were phased and haplotype frequencies were estimated using a standard E-M approach implemented in both programs. The programs performed a regression-based haplotype association test for all haplotypes. Moving window analyses were performed with single marker and with two to four consecutive markers as haplotypes. The false discovery rate (FDR) approach [Benjamini and Hochberg, 1995] was used to correct for multiple testing. Odds ratios and confidence intervals were calculated for the significant markers. LD statistics between markers were calculated and represented graphically by HaploView [Barrett et al., 2005].

Prioritization of Genes Using Bioinformatics Tools

To incorporate biological significance into the consideration of potential candidate genes, we used the software Endeavor [Aerts et al., 2006] to rank all the genes included in our initial screen. Endeavor prioritizes candidate genes underlying diseases or biological pathways by their similarity to a set of "training genes" which are specified by the user. Training genes are those with previous evidence for association with a particular disease or pathway. The program assumes that genes involved in the same disease or pathway are likely to share features with the known candidates. The program employs up to 12 different data sources, such as pathway membership in KEGG, annotations in Gene Ontology, sequence similarity and protein interactions to assess the properties of the training set. The final ranking is obtained by combining all data sources using order statistics. Endeavor is among the most extensively validated software designed for this purpose. For example, it has been validated by leave-one-out cross-validations using 703 disease and pathway genes, and by in vivo studies in zebrafish [Aerts et al., 2006]. In the present study, a training set was selected from the SZgene database (http://www.schizophreniaforum.org/res/sczgene/default.asp accessed 17th September 2007) and a review article on SZ genetics by Straub and Weinberger [2006]. Genes in the "Top Results" list from SZgene, all of which have at least one variant showing a nominally significant result in meta-analyses, were selected, as well as those listed in Table I in Straub and Weinberger's review (list of training genes is shown in Table I). To ensure the reliability of the rankings, we also repeated our analyses leaving one of the training genes out each time. This ensures that the results would not be seriously affected by individual training genes and protects against the risk that some training genes may later turn out to have no association with schizophrenia. We included all genes in the 47–53 Mb region on chromosome 3p as the testing set.

Weighted FDR Analysis

The rankings obtained from Endeavor were also used to weight the *P*-values obtained from the initial screen of the linkage region in a weighted FDR approach [Roeder et al., 2006]. The weight is based on the rankings of the 60 genes containing the genotyped markers. Different degrees of emphasis on biological information could be specified using different functions of the rankings; we used weights that were either directly proportional to ranking (linear weights) or to ranking squared (quadratic weights). Variants belonging to the same gene were given the same weighting. SNPs lying outside a gene were regarded as belonging to the nearest gene, provided that the gene was within 50 kb of the SNP.

Follow-Up Analyses of Identified Genes

For follow-up analyses, we used the WCLUSTAG [Sham et al., 2007] program for selection of additional tag SNPs within the most promising gene from the initial screen (which was neuroglycan C, see below). WCLUSTAG allows users to adjust the clustering threshold according to the location or presumed functional significance of the SNPs. In this study, the R^2 tagging threshold was set to be 1.0 for SNPs in coding regions, 0.8 for SNPs within 5′UTR, 3′UTR, 5′upstream and 3′downstream regions, and 0.3 for introns. We also genotyped possible interacting partners of neuroglycan C (*NGC*), namely *NRG1*, *ErbB3*, and *ErbB4*. *NRG1* is the first member in the neuregulin family and is one of the most established susceptibility genes for SZ [Harrison and Law, 2006]. *ErbB3* is a direct interacting partner of *NGC* while both *ErbB3* and *ErbB4* are binding receptors for *NRG1* (see Discussion Section).

Potential interactions among genetic markers in the same or different genes were assessed by multifactor dimensionality reduction (MDR) [Ritchie et al., 2001; Hahn et al., 2003], a model-free approach for detecting interactions between loci in association studies. MDR reduces the multilocus genotype data to one dimension by classifying each multilocus genotype as either high-risk or low-risk by the ratio of affected individuals to unaffected individuals with that genotype. We used the standard 10-fold cross validation to evaluate the accuracy of prediction of 2-SNP and 3-SNP combinations and permutation tests to calculate empirical *P*-values. In order to test whether the interaction is significant over the main effects and whether the results are similar under a parametric approach, we also carried out logistic regression for interacting SNP pairs identified by MDR. The SNPs were coded in 0, 1, and 2 under the assumption of an additive model. The overall significance of interaction was assessed by a likelihood ratio test, comparing the full model (with all main effects and interaction terms) with the reduced model (with only the main effects).

Bioinformatics Analyses of Candidate Genes

We looked for functional domains that could be relevant to the disease for SNPs that showed statistically significant association.

TABLE I. Training Set of Genes for Endeavor Analysis (Alternative Names for Genes Are Shown in Brackets)

	Gene name	Gene symbol
1	Neuregulin-1	*NRG1*
2	Disrupted in schizophrenia 1	*DISC1*
3	Regulator of G-protein signaling 4	*RGS4*
4	Proline oxidase, mitochondrial precursor (Proline dehydrogenase)	*PRODH*
5	Dysbindin (Dystrobrevin-binding protein 1)	*DTNBP1*
6	D-amino-acid oxidase	*DAO* (*DAAO*)
7	Catechol O-methyltransferase	*COMT*
8	D(3) dopamine receptor	*DRD3*
9	Neuronal acetylcholine receptor subunit alpha-7 precursor	*CHRNA7*
10	RAC-alpha serine/threonine-protein kinase (protein kinase B)	*AKT1* (*PKB*)
11	D(2) dopamine receptor (dopamine D2 receptor)	*DRD2*
12	Metabotropic glutamate receptor 3 precursor	*GRM3*
13	Serine/threonine-protein phosphatase 2B catalytic subunit gamma	*PPP3CC*
14	Glutamate decarboxylase 1	*GAD1*
15	Receptor tyro sine-protein kinase erbB-4 precursor	*ERBB4*
16	Fasciculation and elongation protein zeta 1	*FEZ1*
17	Thioredoxin domain-containing protein 5 precursor	*TXNDC5* (*MUTED*)
18	Orofacial cleft 1 candidate 1	*OFCC1* (*MRDS1*)
19	D-amino acid oxidase activator (protein G72)	*DAOA* (*G72*)
20	Glutamate [NMDA] receptor subunit epsilon 2 precursor (N-methyl D-aspartate receptor subtype 2B)	*GRIN2B* (*NMDAR2B*)
21	Gamma-aminobutyric-acid receptor subunit beta-2 precursor	*GABRB2*
22	Tryptophan 5-hydroxylase 1	*TPH1*
23	Sodium-dependent serotonin transporter (5HT transporter)	*SLC6A4*
24	Interleukin-1 beta precursor	*IL1B*
25	Dopamine receptor D4	*DRD4*
26	Methylenetetrahydrofolatereductase	*MTHFR*
27	Plexin-A2 precursor (Semaphorin receptor OCT)	*PLXNA2*
28	Apolipoprotein E precursor	*APOE*
29	D(1A) dopamine receptor	*DRD1*
30	Haptoglobin precursor	*HP*
31	Cellular tumor antigen p53 (tumor suppressor p53)	*TP53* (*P53*)

SNP sequences were downloaded from public databases. The sequences were submitted to Pfam [Finn et al., 2006], a public protein family database. Pfam has a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains and families. The Wise2 software package (http://www.ebi.ac.uk/Wise2/) was used to blast DNA and protein sequences. Multiple species alignment was obtained for the domains to which our significant markers belong. For the gene of interest, the domain structure of its protein was obtained from UniProt (Universal Protein Resource) (http://www.pir.uniprot.org/). Gene expression information was obtained from the Stanford Microarray Database (http://genome-www5.stanford.edu/).

RESULTS

Association Screen Identified *NGC* as a Promising Candidate Gene for Schizophrenia

In single marker association analyses, four SNPs were found to be nominally significant at $P < 0.05$ (Fig. 2 and Table II), although the global permutation *P*-value of the most significant of the 144 tests was only 0.3107. The best two markers rs3732530 ($P = 0.0048$) and rs4858867 ($P = 0.0040$) were situated in exon 2 in neuroglycan C (*NGC*, also known as chondroitin sulfate proteoglycan 5, or *CSPG5*) and intron 13 in microtubule associated protein 4 (*MAP4*), respectively. The R^2 between these 2 markers is 0.246. The false discovery rate (or *q*-value) for the top 2 markers was 0.343. *P*-values for all 144 markers in the initial screen are shown in Supplementary Table I.

We further screened all possible haplotypes with sizes of two, three, and four markers (Supplementary Table II). Haplotype analyses did not reveal any results with stronger significance than single locus analysis, with the best results being for haplotypes that clustered around the most significant SNPs in single-marker analyses. The best haplotype was formed by the markers rs3732530 and rs4858867 ($P = 0.0066$).

We examined the LD structure of the associated regions based on HapMap data and noted that most of the significant signals from single-marker and haplotype analyses were clustered around an LD block containing the genes *NGC*, *SMARCC1*, *DHX30*, and *MAP4*. Our results suggested associations to

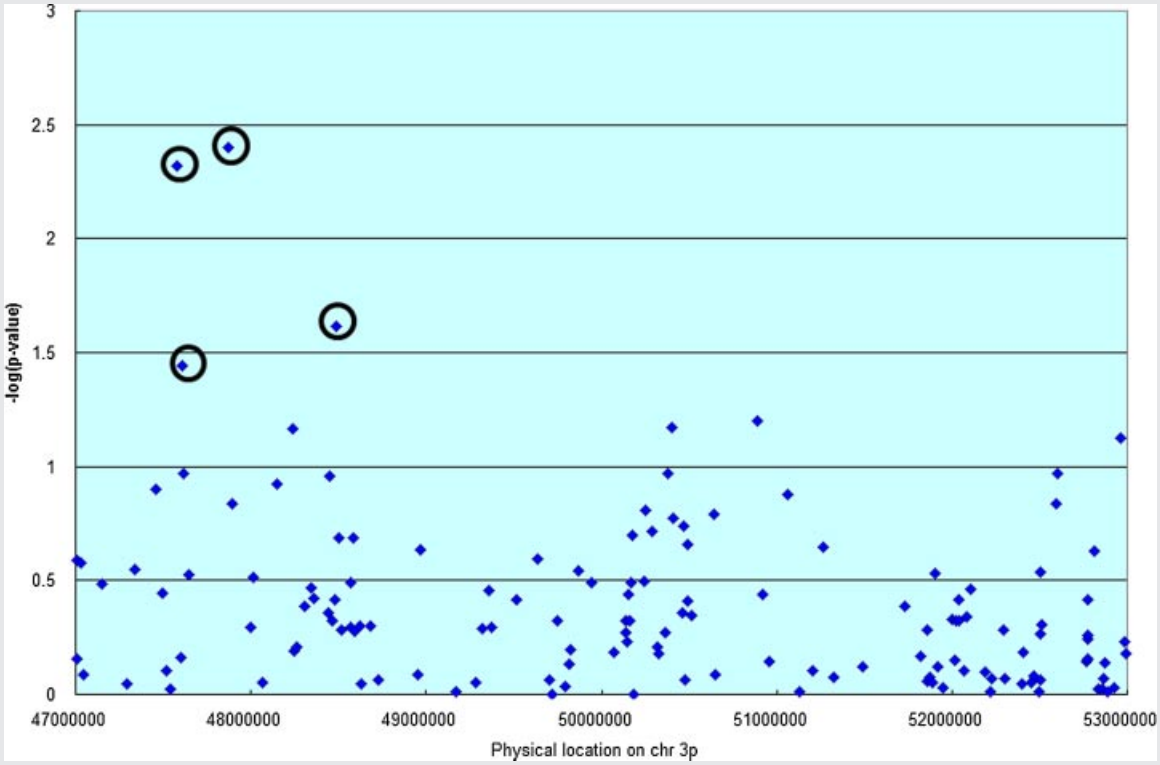

FIG. 2. Statistical significance of the 144 markers in the initial association screen. The four markers showing nominal significance ($P < 0.05$) are circled. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

this block, but could not distinguish the exact gene implicated in the disease due to high LD in the region. Hence we made use of the results from Endeavor to further prioritize genes in this region.

Using all 31 training genes, Endeavor ranked *NGC* 8th, *MAP4* 29th, *DHX30* 57th, and *SMARCC1* 67th out of the 129 genes in the 6MB region scanned (Table III). The *NGC* SNP rs3732530 had the lowest weighted *P*-value among the 144 genotyped SNPs. The weighted FDRs for rs3732530 were 0.411 and 0.302 under linear and quadratic weighting schemes, respectively. We repeated the analysis leaving each of the training genes out and calculated the average ranks. They were in general similar to the original ranks. The average rank of *NGC* was 9th and the actual leave-one-out rank ranged from 6th to 15th. The marker with the lowest weighted *P*-value using the average ranks was again rs3732530, having weighted FDRs of 0.413 and 0.303 respectively under linear and quadratic weighting schemes. We also calculated the weighted *P*-values in the least favorable scenario when *NGC* had the lowest leave-one-out rank (15th). Notably, the *NGC* SNP rs3732530 still achieved the lowest weighted *P*-value among the 144 genotyped markers.

As *NGC* was ranked by Endeavor as the most functionally relevant among the four genes in the implicated LD block and previous biological experiments also suggested a role of *NGC* in neurodevelopment (see Discussion Section), we believed that it is the most promising candidate in the region and further genotyped this gene and its interacting partners.

Follow-Up and MDR Analyses of NGC, NRG1, ErbB3, and ErbB4

In the follow-up phase, four out of seven markers tested within *NGC* showed nominally significant signals ($P < 0.05$). The best SNP in this gene (rs3732530) survived permutation correction for multiple testing within the gene (corrected $P = 0.024$). Two haplotypes were also nominally significant and the best haplotype was formed by markers rs3732530 and rs12489865 ($P = 0.0179$) (Table IV, Supplementary Tables III and IV).

For *ErbB4*, two markers (rs2371276 and rs13032249) were found to be significant with *P*-values 0.0063 and 0.0465, respectively. Four haplotypes also showed $P < 0.05$. Although the best *ErbB4* marker did not withstand permutation correction, the *q*-values for the best two markers (*NGC*_5 and *ErbB4*_11) in the second stage of association analysis (considering all 58 SNPs) were 0.183. None of the *NRG1* markers was significant when considered individually. However, three haplotypes showed nominal significance. The *P*-value for the most significant haplotype was 0.011, although this was not significant after permutation correction. No significant result was found in the three markers genotyped in *ErbB3* (Table IV).

In MDR analyses, five marker combinations were shown to be statistically significant by permutation tests which corrected for the inclusion of multiple SNPs. The results were significant at an FDR of 0.05. The best two-SNP model was composed of variants from *NGC* and *NRG1*, while the best three-SNP model included SNPs from

TABLE II. Initial Association Screen of 144 Markers (Only Significant Results Shown)

			Allele				
Marker	**Gene**		**A**	**C**	**LR chisq**	***P*-value**	**OR (95%CI)**
12	NGC	Cases	605 (0.646)	331 (0.354)	7.952	0.0048	1.30 (1.08–1.57)
	rs3732530	Controls	572 (0.584)	408 (0.416)			
			Allele				
Marker	**Gene**		**C**	**T**	**LR chisq**	***P*-value**	**OR (95%CI)**
14	SMARCC1	Cases	271 (0.281)	695 (0.720)	4.400	0.0359	1.23 (1.01–1.49)
	rs4274776	Controls	329 (0.324)	687 (0.676)			
			Allele				
Marker	**Gene**		**C**	**T**	**LR chisq**	***P*-value**	**OR (95%CI)**
17	MAP4	Cases	777 (0.873)	113 (0.127)	8.287	0.0040	1.46 (1.13–1.90)
	rs4858867	Controls	762 (0.825)	162 (0.175)			
			Allele				
Marker	**Gene**		**A**	**T**	**LR chisq**	***P*-value**	**OR (95%CI)**
33	SCOTIN	Cases	795 (0.848)	143 (0.153)	5.084	0.0241	1.31 (1.04–1.66)
	rs6442126	Controls	822 (0.809)	194 (0.191)			

LR chisq, likelihood ratio chi-square; *NGC*, neuroglycan C (also named *CSPG5*, chondroitin sulfate proteoglycan 5); *SMARCC1*, SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1; *MAP4*, microtubule associated protein 4.

ErbB4 and *NRG1*. Only the best 2-SNP model comprising *NGC* and *NRG1* SNPs was also significant when logistic regression was employed (Table V).

Bioinformatics Analysis for NGC

None of the significant *NGC* SNPs had flanking sequences that lie in any structural units, except one, rs3732530, which was found to be in the chondroitin sulfate attachment domain (accession number PF06566 in the Pfam database). This SNP is a non-synonymous SNP, in which the change of allele from T to G results in an amino acid change from Val to Gly. The sequence of this domain is highly conserved among mouse, rat, human, and chicken, with an additional leucine- and proline-rich sequence attached to the chondroitin sulfate chains in chicken. This domain together with the cytoplasmic domain form the major part of the NGC protein.

DISCUSSION

Using a fine-mapping strategy and a range of statistical tools such as WCLUSTAG, weighted FDR and MDR, we identified neuroglycan C (*NGC*), a membrane-spanning chondroitin sulfate proteoglycan, which is exclusively expressed in the brain, as a novel candidate gene for SZ. This finding is supported by both statistical and biological evidence.

In the association analyses, a non-synonymous SNP in *NGC* (rs3732530) showed a moderately high level of significance which survived permutation correction considering all the examined variants in the gene. Although a global permutation test for all 144 markers in the first screening stage did not show significance, the weighted FDR for the best two markers (in *NGC* and *MAP4*) was only 0.223, using quadratic weighting according to the rankings from Endeavor. It is noteworthy that *NGC* ranked 8th out of the 129 genes in the 6-MB region scanned and the highest among the genes lying in the same LD block.

Importantly, a large body of evidence from biological experiments also supports a role for *NGC* in the pathogenesis of SZ. *NGC* is exclusively expressed in the brain, and was discovered in the developing rat brain. *NGC* is first expressed at embryonic day 16 and the expression level continues to rise with a peak at the early postnatal period, coinciding with the period of active neurite formation and synaptogenesis in the rat brain. Expression of the gene decreases as the brain matures [Watanabe et al., 1995]. *NGC* is expressed in association with axons and dendrites in the CNS during development [Schumacher et al., 1997; Aono et al., 2000]. Nakanishi et al. [2006] showed that a recombinant ectodomain of *NGC* core protein enhances neurite outgrowth in rat neurons, supporting the idea that *NGC* is involved in neurodevelopment.

In addition, *CALEB* (chicken acidic leucine-rich EGF-like domain containing brain protein), a chick homolog of *NGC*, demonstrates biological functions associated with pathogenesis of schizophrenia. Expression of *CALEB* is restricted to the developing and adult nervous system. The component containing chondrotin sulfate was expressed in the embryonic nervous system and down-regulated in the adult. *CALEB* is involved in neurite formation; antibodies against *CALEB* result in impairment of neurite extension in a permissive environment [Schumacher et al., 1997]. Another

TABLE III. Endeavor Ranks of All Genes in the 6 Mb Region Scanned

Gene symbol	Overall rank	Average rank
GRM2	1	1.42
PLXNB1	2	1.84
SLC26A6	3	4.16
MST1	4	4.06
ITIH4	5	5.65
MST1R	6	5.74
GNAI2	7	6.39
NGC	8	8.90
BSN	9	9.13
IMPDH2	10	11.55
AMT	11	11.58
CAMKV	12	16.23
SEMA3B	13	14.45
LAMB2	14	12.61
COL7A1	15	15.35
MAPKAPK3	16	14.97
ITIH1	17	19.84
USP4	18	23.39
SETD2	19	20.90
STAB1	20	19.45
DAG1	21	17.90
GNL3	22	21.97
TLR9	23	22.23
PRKAR2A	24	22.84
SHISA5	25	26.45
CDC25A	26	28.48
BAP1	27	27.00
NBEAL2	28	30.77
MAP4	29	31.68
RASSF1	30	30.61
ACY1	31	32.16
QARS	32	34.52
RHOA	33	30.81
DUSP7	34	33.06
ZMYND10	35	37.48
HYAL1	36	38.97
DOCK3	37	38.00
SCAP	38	38.00
TUSC4	39	46.42
QRICH1	40	40.35
HYAL2	41	41.16
CACNA2D2	42	44.52
NISCH	43	46.55
ARIH2	44	45.42
PCBP4	45	40.23
RAD54L2	46	49.19
UQCRC1	47	41.65
SLC25A20	48	50.48
TREX1	49	49.65
USP19	50	62.19
PBRM1	51	53.26
SEMA3F	52	49.74
UBE1L	53	56.26
P4HTM	54	52.32
IHPK1	55	52.84
SLC38A3	56	54.71
DHX30	57	55.16

TABLE III. (*Continued*)

Gene symbol	Overall rank	Average rank
TEX264	58	62.32
TUSC2	59	59.19
CISH	60	60.77
HYAL3	61	56.13
ABHD14A	62	51.32
SFMBT1	64	65.48
HEMK1	65	67.32
SEMA3G	66	64.03
SMARCC1	67	64.00
ITIH3	68	69.68
GPX1	69	59.00
RBM5	70	75.35
IHPK2	71	71.13
NEK4	72	78.29
KIF9	73	78.00
PPM1M	74	80.84
GLYCTK	75	74.74
KLHL18	76	75.35
TNNC1	77	72.35
MON1A	78	75.68
GNAT1	79	76.90
PARP3	80	75.29
VPRBP	81	82.26
PTPN23	82	78.19
NT5DC2	83	81.77
RPL29	84	83.00
TMEM89	85	90.32
RNF123	86	85.10
CAMP	87	77.19
APEH	88	88.77
TMEM115	89	86.68
WDR51A	90	90.97
CCDC36	91	89.81
GLT8D1	92	86.35
RBM15B	93	93.19
RBM6	94	96.90
TRAIP	95	96.26
AMIGO3	96	96.16
C3orf71	97	97.90
WDR6	98	100.35
ABHD14B	99	101.32
FBXW12	100	93.35
NICN1	101	101.84
TCTA	102	104.10
PFKFB4	103	107.68
SPCS1	104	102.55
C3orf62	105	105.03
GPR62	106	93.61
PHF7	107	107.00
ALAS1	108	108.23
NME6	109	106.35
CCDC71	110	109.16
CYB561D2	111	112.87
RRP9	112	118.06
ZNF589	113	115.00
GMPPB	114	112.97
DALRD3	115	115.32
TMEM113	116	115.39

TABLE III. (*Continued*)

Gene symbol	Overall rank	Average rank
CCDC51	117	112.19
C3orf54	118	117.65
TMEM103	119	118.74
IQCF2	120	119.35
UCN2	121	118.97
DNAH1	122	119.45
KLHDC8B	123	123.35
IQCF1	124	123.90
C3orf60	125	124.61
C3orf45	126	125.10
CDH29	127	127.16
STGC3	128	127.81
C3orf18	129	125.23

study [Juttner et al., 2005] showed that *CALEB*-deficient mice exhibit impaired synapse function in early postnatal stages.

NGC has been implicated in the effects of methamphetamine and cocaine, two drugs that are known to produce psychotic symptoms such as paranoid delusions and hallucinations. Ishikawa et al. [2006] reported that repeated administration of methamphetamine significantly increased mRNA levels of *NGC* in numerous rat brain areas, such as frontal cortex, nucleus accumbens and hippocampus. Moreover, *NGC* protein levels in the nucleus accumbens increased in rats repeatedly treated by methamphetamine. Up-regulation in both mRNA and protein levels of *NGC* was also found in chronic cocaine-treated rats [Toda et al., 2002].

NGC possesses an *EGF*-like extracellular domain and is an active growth factor. It is considered to be the sixth member of the neuregulin family, as it acts as a direct ligand for ErbB3 tyrosine kinases that can transactivate *ErbB2* [Kinugasa et al., 2004]. Neuregulin 1 (*NRG1*), the first member of the neuregulin family, is a strong candidate gene for SZ. It is involved in a large variety of processes that may be relevant to the pathogenesis of SZ, such as neuronal migration, oligodendrocyte development, myelin formation, and neurotransmitter receptor (particularly NMDA receptor) function [Corfas et al., 2004; Harrison and Law, 2006]. Being another member of the neuregulin family, *NGC* might share some of the functional properties with *NRG1* and hence may be involved in the pathogenesis of SZ.

NRG1 has been reported to be associated with SZ in different populations. The initial study on *NRG1* by Stefansson et al. [2002] and many other studies [Yang et al., 2003; Corvin et al., 2004; Hall et al., 2004; Li et al., 2004, 2006; Tang et al., 2004; Zhao et al., 2004] identified SNPs and haplotypes near the 5′ end of the gene to be associated with SZ. On the other hand, the 3′ end region was also implicated in a few studies [Yang et al., 2003; Li et al., 2004; Petryshen et al., 2005], two involving Chinese samples [Yang

TABLE IV. All Markers Showing Nominal Significance ($P < 0.05$) in Single SNP (Allelic or Genotypic) or Haplotype Analyses in the Follow-Up Phase

		P-value				
Marker number	**rs number**	**Genotypic**	**Allelic**	**2-marker**	**3-marker**	**4-marker**
NGC						
NGC_3	rs10865948	0.090	**0.029**			
NGC_4	rs11716779	0.230	0.538	**0.040**		
NGC_5	rs3732530	**0.007**	**0.005**	**0.018**		
NGC_6	rs12489865	**0.022**	**0.017**			
NGC_7	rs3755637	0.091	**0.046**			
NRG1						
NRG1_5	rs1531746	0.664	0.801	**0.016**		
NRG1_6	rs10503904	0.594	0.658			
NRG1_9	rs6986716	0.243	0.094	**0.011**	**0.020**	
NRG1_10	rs10503915	0.055	0.286			
NRG1_11	rs7016691	0.928	0.711			
ErbB4						
ErbB4_10	rs2272024	0.994	0.982	**0.036**		
ErbB4_11	rs2371276	**0.002**	**0.006**	**0.033**		**0.040**
ErbB4_12	rs10205553	0.834	0.945			
ErbB4_13	rs6435659	0.432	0.986			
ErbB4_14	rs6435660	0.067	0.132			
ErbB4_17	rs6435665	0.241	0.118			**0.017**
ErbB4_18	rs6435670	0.941	0.757			
ErbB4_19	rs839530	0.551	0.278			
ErbB4_20	rs12478950	0.227	0.265			

P-values for haplotype analyses are shown next to the 1st SNP in the sliding window.
P-values that are nominally significant (<0.05) are in bold. For haplotype analyses, only nominally significant results are shown.

TABLE V. Average Prediction Accuracies of the Best Combinations of SNPs Identified by MDR and Logistic Regression

	SNP combination	rs number of SNPs	Testing accuracy	MDR p	Logistic regression *P*
2-SNP					
1st	*NGC_3, NRG1_12*	rs10865948, rs4733130	0.5711	**0.015**	**0.009**
2nd	*ErbB4_11, ErbB4_19*	rs2371276, rs839530	0.5642	**0.031**	0.125
3rd	*ErbB4_11, NGC_1*	rs2371276, rs7628631	0.5522	0.097	0.140
3-SNP					
1st	*ErbB4_6, ErbB4_13, NRG1_8*	rs10197225, rs6435659, rs2347504	0.5754	**0.009**	0.996
2nd	*ErbB4_3, ErbB4_11, NRG1_7*	rs3748962, rs2371276, rs1545961	0.5751	**0.009**	0.409
3rd	*ErbB4_6, ErbB4_13, NGC_6*	rs10197225, rs6435659, rs12489865	0.5617	**0.043**	0.295

P-values for MDR were derived from 1,000 permutations.
P-values for logistic regression were derived by comparing the full model (with all main effects and interaction terms) with the reduced model (with only the main effects). $P < 0.05$ are in bold.

et al., 2003; Li et al., 2004]. The 5′ and 3′ regions containing susceptibility variants were designated regions A and B by Thomson et al. [2007]. Interestingly, the risk haplotypes identified in this study are between regions A and B, a region of this large gene (~1.1 Mb) that most previous studies have not thoroughly screened, except Thomson et al. [2007] and Benzel et al. [2007]. Further comprehensive studies of *NRG1* may be warranted, especially in non-Caucasian groups.

NRG1-ErbB signaling regulates a variety of processes that may be involved in SZ [Corfas et al., 2004]. *ErbB3* and *ErbB4* are hence good functional candidate genes for SZ. *ErbB3* is a direct interacting partner of *NGC*. Three microarray studies [Hakak et al., 2001; Tkachev et al., 2003; Aston et al., 2004] have shown that *ErbB3*, as a myelination-related gene, has altered expression in the post-mortem brains of schizophrenic patients. Oligodendrocyte dysfunction and disturbances in myelination are receiving increasing attention as important mechanisms underlying SZ [Karoutzou et al., 2008]. *ErbB3* has also been examined previously in three association studies [Benzel et al., 2007; Kanazawa et al., 2007; Watanabe et al., 2007], but no susceptibility variants have been identified. We also failed to find any positive associations in *ErbB3* in this study.

ErbB4 is another biological candidate for SZ and association studies have been performed on this gene. Silberberg et al. [2006] identified three SNPs in *ErbB4* which showed marked differences in allele and genotype frequencies between schizophrenic patients and controls. Benzel et al. [2007] screened 109 SNPs in *ErbB4* and found 14 SNPs with *P*-values <0.05 in either genotypic or allelic tests. Our study is the first to investigate *ErbB4* polymorphisms in an Asian population. The two nominally significant SNPs and four-marker haplotype did not withstand permutation correction, but the FDR of the most significant *ErbB4* SNP was nevertheless 0.183, which can be considered suggestive, though not conclusive.

It should be noted that as the focus of the current study is on fine-mapping of chromosome 3p and discovery of *NGC* as a novel candidate gene, the genotyping coverage of the above interacting partners (*NRG1*, *ErbB3*, and *ErbB4*) may not be comprehensive enough. It is therefore possible that some susceptibility variants in these genes were not captured.

The finding in the MDR analyses of significant statistical interactions between SNPs at *NGC* and *NRG1* is intriguing as these two genes are both members of the neuregulin family and share *ErbB3* as a receptor. Thus it is plausible that they may play complementary or synergistic roles in the pathogenesis of schizophrenia, for example, in neurodevelopmental processes. Notably, the interaction between this pair of SNPs was also confirmed by logistic regression and the interaction was significant over the main effects. We also identified two sets of 3-locus interactions between *NRG1* and *ErbB4*, consistent with the results of functional and association studies showing these two genes interact. Statistical epistasis between the Icelandic haplotype of *NRG1* and a *ErbB4* variant was previously reported by Norton et al. [2006]. Interestingly, we found statistical interaction between *NGC* and *ErbB4* that was also modestly significant. Previous experiments showed that *NGC* directly binds to *ErbB3* but not to *ErbB4*. Although *NGC* and *ErbB4* may not be direct interacting partners, the two genes may be involved in complementary pathways related to brain development or other processes implicated in SZ. *NGC*, *NRG1*, and *ErbB4* may be part of a network regulating a variety of processes involved in SZ.

Although these three-locus interactions were biologically plausible and supported by MDR, they did not show significance when logistic regression was used. This may be due to the different nature of the two approaches. MDR is non-parametric and completely model-free. MDR requires no mode of inheritance to be specified as any multilocus genotype can be formed as the high-risk or low-risk groups. It has high power to detect interactions in the absence of main effects [Ritchie et al., 2003]. Despite being more flexible, the genotype patterns identified by MDR may be very complex and difficult to interpret. On the other hand, logistic regression is a form of generalized linear model and does not involve partitioning of genotype combinations into high-risk and low-risk groups. Logistic regression may run into problems when dealing with large numbers of SNPs as the number of possible interaction terms increases rapidly with each additional SNP. Inclusion of too many independent variables with respect to the number of outcome events increases type 1 and type 2 errors [Peduzzi et al., 1996]. An inheritance model is often assumed to limit the number of terms in the regression model (we assumed an additive model in this study). However, logistic regression has an advantage over MDR in that it allows evaluation of interaction controlling for main effects. In view of the substantial differences between the two methods, some discrepancies between the test results are expected. Nevertheless, one should also alert to the possibility of false positives, particularly because we have included all SNPs regardless

of whether they were independently associated with the disease. The interactions identified in this study should not be regarded as confirmatory and further association and functional studies are required to clarify the findings.

ErbB3 was not found in the interacting pairs in this study. However, its functional importance cannot be excluded. It is possible that the functional variant has not been captured (e.g., the variant may be a rare one) or there may be variants lying outside the gene which regulate its expression.

In summary, we have performed a high-density association screen of a high-LD, gene-rich region on chromosome 3p and discovered *NGC* as a novel candidate gene for SZ. A role for *NGC* in SZ is supported by previous biological findings. Our methodology demonstrates the use of tag SNPs, bioinformatics tools, weighted FDR and MDR in association studies to identify new susceptibility genes for complex diseases. Although our study provides statistical evidence for the involvement of *NGC* and other candidate genes in SZ, our results cannot be considered conclusive. Further replications, preferably with larger samples and in other populations, are desirable for validation of our findings.

ACKNOWLEDGMENTS

The project was supported by a University of Hong Kong small project grant, the University of Hong Kong Strategic Research Theme on Genomics, and the Hong Kong Research Grants Council General Research Fund 757905. We are also grateful to C.L. Kwok, F. Lieh-Mak, J.H. Zhao, D. Collier, R. Murray, H.K. Cheung, P.C. Lu, C.H. Yuen and C.K. Lin for their help in subject recruitment.

REFERENCES

Aerts S, Lambrechts D, Maity S, Van Loo P, Coessens B, De Smet F, Tranchevent LC, De Moor B, Marynen P, Hassan B, et al. 2006. Gene prioritization through genomic data fusion. Nat Biotechnol 24(5): 537–544.

Ao SI, Yip K, Ng M, Cheung D, Fong PY, Melhado I, Sham PC. 2005. CLUSTAG: Hierarchical clustering and graph methods for selecting tag SNPs. Bioinformatics 21(8):1735–1736.

Aono S, Keino H, Ono T, Yasuda Y, Tokita Y, Matsui F, Taniguchi M, Sonta S, Oohira A. 2000. Genomic organization and expression pattern of mouse neuroglycan C in the cerebellar development. J Biol Chem 275(1):337–342.

Aston C, Jiang L, Sokolov BP. 2004. Microarray analysis of postmortem temporal cortex from patients with schizophrenia. J Neurosci Res 77(6): 858–866.

Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21(2):263–265.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc 57: 289–300.

Benzel I, Bansal A, Browning BL, Galwey NW, Maycox PR, McGinnis R, Smart D, St Clair D, Yates P, Purvis I. 2007. Interactions among genes in the ErbB-Neuregulin signalling network are associated with increased susceptibility to schizophrenia. Behav Brain Funct 3:31.

Cardno AG, Gottesman II. 2000. Twin studies of schizophrenia: From bow-and-arrow concordances to star wars Mx and functional genomics. Am J Med Genet 97(1):12–17.

Chen RY, Sham P, Chen EY, Li T, Cheung EF, Hui TC, Kwok CL, Lieh-Mak F, Zhao JH, Collier D, et al. 2001. No association between T102C polymorphism of serotonin-2A receptor gene and clinical phenotypes of Chinese schizophrenic patients. Psychiatry Res 105(3): 175–185.

Corfas G, Roy K, Buxbaum JD. 2004. Neuregulin 1-erbB signaling and the molecular/cellular basis of schizophrenia. Nat Neurosci 7(6):575–580.

Corvin AP, Morris DW, McGhee K, Schwaiger S, Scully P, Quinn J, Meagher D, Clair DS, Waddington JL, Gill M. 2004. Confirmation and refinement of an 'at-risk' haplotype for schizophrenia suggests the EST cluster, Hs.97362, as a potential susceptibility gene at the Neuregulin-1 locus. Mol Psychiatry 9(2):208–213.

Dudbridge F. 2003. Pedigree disequilibrium tests for multilocus haplotypes. Genet Epidemiol 25(2):115–121.

Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, et al. 2006. Pfam: Clans, web tools and services. Nucleic Acids Res 34(Database issue):D247–D251.

Hahn LW, Ritchie MD, Moore JH. 2003. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. Bioinformatics 19(3):376–382.

Hakak Y, Walker JR, Li C, Wong WH, Davis KL, Buxbaum JD, Haroutunian V, Fienberg AA. 2001. Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. Proc Natl Acad Sci USA 98(8):4746–4751.

Hall D, Gogos JA, Karayiorgou M. 2004. The contribution of three strong candidate schizophrenia susceptibility genes in demographically distinct populations. Genes Brain Behav 3(4):240–248.

Harrison PJ, Law AJ. 2006. Neuregulin 1 and schizophrenia: Genetics, gene expression, and neurobiology. Biol Psychiatry 60(2):132–140.

Harrison PJ, Weinberger DR. 2005. Schizophrenia genes, gene expression, and neuropathology: On the matter of their convergence. Mol Psychiatry 10(1):40–68, image 5.

Ishikawa K, Nitta A, Mizoguchi H, Mohri A, Murai R, Miyamoto Y, Noda Y, Kitaichi K, Yamada K, Nabeshima T. 2006. Effects of single and repeated administration of methamphetamine or morphine on neuroglycan C gene expression in the rat brain. Int J Neuropsychopharmacol 9(4):407–415.

Juttner R, More MI, Das D, Babich A, Meier J, Henning M, Erdmann B, Mu Ller EC, Otto A, Grantyn R, et al. 2005. Impaired synapse function during postnatal development in the absence of CALEB, an EGF-like protein processed by neuronal activity. Neuron 46(2):233–245.

Kanazawa T, Glatt SJ, Tsutsumi A, Kikuyama H, Koh J, Yoneda H, Tsuang MT. 2007. Schizophrenia is not associated with the functional candidate gene ERB B3: Results from a case-control study. Am J Med Genet Part B 144B(1):113–116.

Karoutzou G, Emrich HM, Dietrich DE. 2008. The myelin-pathogenesis puzzle in schizophrenia: A literature review. Mol Psychiatry 13:245–260.

Kinugasa Y, Ishiguro H, Tokita Y, Oohira A, Ohmoto H, Higashiyama S. 2004. Neuroglycan C, a novel member of the neuregulin family. Biochem Biophys Res Commun 321(4):1045–1049.

Lewis CM, Levinson DF, Wise LH, DeLisi LE, Straub RE, Hovatta I, Williams NM, Schwab SG, Pulver AE, Faraone SV, et al. 2003. Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. Am J Hum Genet 73(1):34–48.

Li T, Stefansson H, Gudfinnsson E, Cai G, Liu X, Murray RM, Steinthorsdottir V, Januel D, Gudnadottir VG, Petursson H, et al. 2004. Identification of a novel neuregulin 1 at-risk haplotype in Han schizophrenia Chinese patients, but no association with the Icelandic/Scottish risk haplotype. Mol Psychiatry 9(7):698–704.

Li D, Collier DA, He L. 2006. Meta-analysis shows strong positive association of the neuregulin 1 (NRG1) gene with schizophrenia. Hum Mol Genet 15(12):1995–2002.

Maziade M, Roy MA, Rouillard E, Bissonnette L, Fournier JP, Roy A, Garneau Y, Montgrain N, Potvin A, Cliche D, et al. 2001. A search for specific and common susceptibility loci for schizophrenia and bipolar disorder: A linkage study in 13 target chromosomes. Mol Psychiatry 6(6):684–693.

Nakanishi K, Aono S, Hirano K, Kuroda Y, Ida M, Tokita Y, Matsui F, Oohira A. 2006. Identification of neurite outgrowth-promoting domains of neuroglycan C, a brain-specific chondroitin sulfate proteoglycan, and involvement of phosphatidylinositol 3-kinase and protein kinase C signaling pathways in neuritogenesis. J Biol Chem 281(34):24970–24978.

Norton N, Moskvina V, Morris DW, Bray NJ, Zammit S, Williams NM, Williams HJ, Preece AC, Dwyer S, Wilkinson JC, et al. 2006. Evidence that interaction between neuregulin 1 and its receptor erbB4 increases susceptibility to schizophrenia. Am J Med Genet Part B 141B(1):96–101.

Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. 1996. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 49(12):1373–1379.

Petryshen TL, Middleton FA, Kirby A, Aldinger KA, Purcell S, Tahl AR, Morley CP, McGann L, Gentile KL, Rockwell GN, et al. 2005. Support for involvement of neuregulin 1 in schizophrenia pathophysiology. Mol Psychiatry 10(4):366–374, 328.

Pulver AE, Lasseter VK, Kasch L, Wolyniec P, Nestadt G, Blouin JL, Kimberland M, Babb R, Vourlis S, Chen H, et al. 1995. Schizophrenia: A genome scan targets chromosomes 3p and 8p as potential sites of susceptibility genes. Am J Med Genet 60(3):252–260.

Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF, Moore JH. 2001. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. Am J Hum Genet 69(1):138–147.

Ritchie MD, Hahn LW, Moore JH. 2003. Power of multifactor dimensionality reduction for detecting gene-gene interactions in the presence of genotyping error, missing data, phenocopy, and genetic heterogeneity. Genet Epidemiol 24(2):150–157.

Roeder K, Bacanu SA, Wasserman L, Devlin B. 2006. Using linkage genome scans to improve power of association in genome scans. Am J Hum Genet 78(2):243–252.

Sawa A, Snyder SH. 2002. Schizophrenia: Diverse approaches to a complex disease. Science 296(5568):692–695.

Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, et al. 2007. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 316(5829):1331–1336.

Schumacher S, Volkmer H, Buck F, Otto A, Tarnok A, Roth S, Rathjen FG. 1997. Chicken acidic leucine-rich EGF-like domain containing brain protein (CALEB), a neural member of the EGF family of differentiation factors, is implicated in neurite formation. J Cell Biol 136(4):895–906.

Sham PC, Ao SI, Kwan JS, Kao P, Cheung F, Fong PY, Ng MK. 2007. Combining functional and linkage disequilibrium information in the selection of tag SNPs. Bioinformatics 23(1):129–131.

Silberberg G, Darvasi A, Pinkas-Kramarski R, Navon R. 2006. The involvement of ErbB4 with schizophrenia: Association and expression studies. Am J Med Genet Part B 141B(2):142–148.

Stefansson H, Sigurdsson E, Steinthorsdottir V, Bjornsdottir S, Sigmundsson T, Ghosh S, Brynjolfsson J, Gunnarsdottir S, Ivarsson O, Chou TT, et al. 2002. Neuregulin 1 and susceptibility to schizophrenia. Am J Hum Genet 71(4):877–892.

Straub RE, Weinberger DR. 2006. Schizophrenia genes—Famine to feast. Biol Psychiatry 60(2):81–83.

Tang JX, Chen WY, He G, Zhou J, Gu NF, Feng GY, He L. 2004. Polymorphisms within 5' end of the Neuregulin 1 gene are genetically associated with schizophrenia in the Chinese population. Mol Psychiatry 9(1):11–12.

Thomson PA, Christoforou A, Morris SW, Adie E, Pickard BS, Porteous DJ, Muir WJ, Blackwood DH, Evans KL. 2007. Association of Neuregulin 1 with schizophrenia and bipolar disorder in a second cohort from the Scottish population. Mol Psychiatry 12(1):94–104.

Tkachev D, Mimmack ML, Ryan MM, Wayland M, Freeman T, Jones PB, Starkey M, Webster MJ, Yolken RH, Bahn S. 2003. Oligodendrocyte dysfunction in schizophrenia and bipolar disorder. Lancet 362(9386):798–805.

Toda S, McGinty JF, Kalivas PW. 2002. Repeated cocaine administration alters the expression of genes in corticolimbic circuitry after a 3-week withdrawal: A DNA macroarray study. J Neurochem 82(5):1290–1299.

Watanabe E, Maeda N, Matsui F, Kushima Y, Noda M, Oohira A. 1995. Neuroglycan C, a novel membrane-spanning chondroitin sulfate proteoglycan that is restricted to the brain. J Biol Chem 270(45):26876–26882.

Watanabe Y, Fukui N, Nunokawa A, Muratake T, Kaneko N, Kitamura H, Someya T. 2007. No association between the ERBB3 gene and schizophrenia in a Japanese population. Neurosci Res 57(4):574–578.

Yang JZ, Si TM, Ruan Y, Ling YS, Han YH, Wang XL, Zhou M, Zhang HY, Kong QM, Liu C, et al. 2003. Association study of neuregulin 1 gene with schizophrenia. Mol Psychiatry 8(7):706–709.

Zhao X, Shi Y, Tang J, Tang R, Yu L, Gu N, Feng G, Zhu S, Liu H, Xing Y, et al. 2004. A case control and family based association study of the neuregulin1 gene and schizophrenia. J Med Genet 41(1):31–34.