Published in final edited form as:

Neurosci Lett. 2010 May 14; 475(2): 80–84. doi:10.1016/j.neulet.2010.03.047.

Rare *NRXN1* promoter variants in patients with schizophrenia

Abhishek K. Shah¹, Nina M Tioleco¹, Karen Nolan², Joseph Locker³, Katherine Groh¹, Catalina Villa¹, Pavla Stopkova⁴, Erika Pedrosa¹, and Herbert M. Lachman^{1,*}

¹Department of Psychiatry and Behavioral Sciences, Division of Basic Research, Albert Einstein College of Medicine, Bronx, New York

²Nathan S. Kline Institute for Psychiatric Research, Orangeburg, New York; Department of Psychiatry, New York University School of Medicine, New York

³Department of Pathology, Albert Einstein College of Medicine

⁴Prague Psychiatric Center, Prague, Czech Republic

Abstract

Copy number variants (CNVs) affecting the neurexin 1 (*NRXN1*) gene have been found in a subgroup of patients with schizophrenia (SZ). *NRXN1* expression is complex, with multiple alternative splice forms generated from two major transcripts; *NRXN1α* and *NRXN1β*. The majority of CNVs in SZ are deletions affecting the proximal *NRXN1α* exons and promoter region. Rare chromosomal events are useful in understanding the genetic basis of complex psychiatric disorders since affected genes become feasible targets to analyze for more subtle genetic alterations. As a first step towards this goal, we resequenced the *NRXN1α* promoter region in 170 patients with SZ and a similar number of controls. Two rare mutations were identified in the patient population. One previously unknown single nucleotide polymorphism (SNP) was found in controls. Bioinformatics analysis suggests that binding to several transcription factors may be affected by the minor alleles. The findings suggest that in addition to chromosomal alterations disrupting the *NRXN1α* promoter, rare point mutations in the region may also be involved in SZ pathogenesis.

INTRODUCTION

The three members of the neurexin gene family, *NRXN1*, *NRXN2* and *NRXN3*, code for proteins that together with its binding partners, the neuroligins, are involved in presynaptic and postsynaptic differentiation at GABAergic and glutamatergic synapses [6,10,16,22]. There is mounting evidence implicating both neurexin and neuroligin encoding genes as primary factors in neuropsychiatric disorders. In addition, for example, *NRXN1* and *NRXN3* have been implicated in alcoholism, and addiction to nicotine and illicit drugs [17,14,2,24]. Recently, several groups have identified patients with SZ who have deletions disrupting the *NRXN1α* promoter region and proximal (N-terminal encoding) exons [7,8,12,23,25,27,28]. In addition, several groups have reported deletions and translocations affecting *NRXN1* in autism spectrum disorders (ASD) [4,11,20]. The identification of structural variants affecting the *NRXN1α* promoter region suggests that altered expression of the long alpha isoform contributes to SZ

© 2010 Elsevier Ireland Ltd. All rights reserved.

*corresponding author HM Lachman, Department of Psychiatry and Behavioral Sciences, Division of Basic Research, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, New York 10461, Phone: 718-430-2428; Herb.Lachman@einstein.yu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

susceptibility. These observations also suggest that more subtle disruptions of gene expression, such as promoter mutations, might be involved in disease susceptibility in a subgroup of SZ patients. To test this hypothesis, we resequenced the *NRXN1* α promoter region in patients and controls.

METHODS

Patients

Patients with SZ (n=170) were recruited from the Nathan Kline Institute/Rockland Psychiatric Center (NKI/RPC). Clinical diagnosis was established according to DSM-IV criteria using the SCID (research version, N=122) or chart review of chronic inpatients with an admitting diagnosis of SZ (N=48). U.S. controls were Caucasian blood bank donors. In the SZ sample, 31% were female with a mean age of 42 $\pm$ 10, while 45% of controls were female with a mean age of 48 $\pm$ 13. Patients with bipolar disorder (BD) from the Czech Republic were recruited from the Prague Psychiatric Center. Patients were diagnosed on the basis of either a Schedule for Affective Disorders and Schizophrenia-Lifetime (SADS-L) interview or by clinical interview modified from SADS-L using Research Diagnostic Criteria (RDC) criteria. Controls were blood-bank donors and patients hospitalized for medical reasons. Fifty-four percent of bipolar patients were female, with a mean age of 49 $\pm$ 17 years, compared to 40% females in the control sample with a mean age of 47 $\pm$ 16. The cohort consisted of 78% bipolar I and 22% bipolar II.

All patients signed an informed consent approved by the Ethical Committee on Clinical Investigation (Czech samples) and the AECOM and NKI/RPC IRBs (U.S. samples), which comply with guidelines established for human experimentation by the Declaration of Helsinki.

Sequencing

An 814 base pair (bp) PCR fragment was generated using primers P1 (cagctttccatgggtctagcaggggcct) and P2 (aggcttcacgcaaaacaacc). DNA samples were amplified using "hot-start" polymerase (manufacturer's instructions; Qiagen, HotStarTaq). Cycling parameters were 95°C for 15 minutes, then 30 cycles of 94°C (30 sec), 49°C (30 sec) and 72°C (two min). Samples were purified using QIAquick® (Qiagen Sciences). Sequencing was carried out using primer P2.

Genotyping

Genotyping for rs2287235 was carried out using TaqMan® Allelic Discrimination. Samples were amplified by PCR using a 7900HT cyclor and analyzed using SDS 2.1 software (Applied Biosystems).

Statistical Analysis and Bioinformatics Analysis

StatXACT-5 (Cytel Software Corporation, Cambridge MA) was used to compute χ^2 statistics. The level of significance was set at $p < 0.05$. Hardy-Weinberg equilibrium (HWE) was computed using goodness-of-fit χ^2 .

Promoter variants were analyzed computationally using TRANSFAC suite from Biobase Biological databases (<https://portal.biobase-international.com>) to identify potential cis-regulatory factors that might be altered by allelic differences.

RESULTS

NRXN1 α gene expression initiates from non-coding exon 1. This exon is in very close juxtaposition (<100 bp) to the transcription start site of a long non-coding RNA, AK127244

(Figure 1). Previous work in our lab showed that exon 1 contains a dual, bidirectional promoter with activity in the direction of both *NRXN1* and AK127244 [26]. Most of the *NRXN1*α CNVs that have been found in patients with SZ disrupt exon 1, which was the focus of this resequencing effort. An 814 bp fragment was generated and sequenced using primer P2. The amplified fragment contains a polymorphic CT tract (AG on plus strand), which is at the most distal *NRXN1* transcription start site. Using the P2 primer, we obtained excellent sequence data for a portion of intron 1, the intron 1/exon 1 junction, and the entire length of exon 1, up until the polymorphic CT/AG repeat. The amplified region contains two other known polymorphisms, the SNPs rs2287235 and rs72828367, the latter of which is a T/C transition within the repeat element. The DNA sequencing data we obtained allowed us to accurately genotype rs2287235, which is 3,737 bases away from the translation start site in exon 2 (–3737T/C; note that all polymorphic bases and rare mutations will be described from the perspective of the minus strand, since *NRXN1* is transcribed from that strand). An accurate assessment of rs72828367 and the polymorphic CT/AG repeat was not possible, however, since these are at the very end of the sequencing run, resulting in low amplitude fluorescent peaks and difficulty genotyping heterozygotes.

We identified three new variants after sequencing 170 SZ patients and 160 controls (Figures 1 and 2). One was –3797G/A, which was found in 3 controls and no patients, suggesting that it is a previously unidentified SNP. The other two were –3781T/C, which was found in one patient, and –3681G/T, which was found in another. The variants –3797G/A and –3781T/C map to exon 1 and –3681G/T maps to intron 1, 27 bases away from *NRXN1* exon 1/intron1 junction.

The SNP rs2287235 (–3737 T/C) was genotyped in cases and controls using the sequence data. As seen in Figure 2B, there was no significant difference in the genotype or allele distribution. Considering the genetic overlap between SZ and BD featured at many genetic loci and candidate genes, we also genotyped a cohort of patients with BD and controls from the Czech Republic. For this analysis, a Taqman allele discrimination system was employed since the bipolar samples were not sequenced. There was no difference in the allele or genotype distribution in bipolar patients and controls. In addition, there was no deviation in the genotype distribution from that expected from a Hardy Weinberg equilibrium in any of the patient and control cohorts. The allele frequencies were similar to that found in the HapMap CEU families (CEPH - Utah residents with ancestry from northern and western Europe) (“A” allele frequency = 0.853; “G” allele frequency = 0.147) (<http://hapmap.ncbi.nlm.nih.gov/>).

These findings suggest that rs2287235 does not play a role in either SZ or BD pathogenesis, although the samples we analyzed are underpowered to detect a genetic contribution for this SNP if the effect size is small. The finding is also consistent with the data reported by Rujescu et al who did not detect any significant association to 2p16.3 polymorphic markers in their large SZ and control cohort - the same data set in which a significant increase in CNVs disrupting *NRXN1* in patients was found [27].

These 3 variants, as well as rs2287235, were subjected to TRANSFAC analysis to assess the potential for allele-specific differences in transcription factor binding sites, the results of which are summarized in Figure 3. Analysis used the MATCH program with the set of high quality matrices for vertebrate transcription factors. Sequences were first scanned using settings that minimized false negative matches. They were then rescanned with settings that minimized false positive matches. Separate analysis of each variant pair was compared, to find significant matches that discriminated the two alleles. The most significant finding was for –3781, in which a high motif score was given for binding to the transcription factor Runx2 when the mutant allele was present, but no score for the major allele. Other allele-specific differences were found for YY1 and Cux1, ubiquitously expressed transcription factors.

Interesting differences were also found for the other patient-specific variant, -3681C/A. Four transcription factors - SREBP1, WT1, MAZ, and ZBTB7 - showed perfect matrix scores of 1.00 for the core and/or total binding sites. All showed significant allele-specific differences with the minor allele showing lower scores, except for SREBP1, which showed a lower score for the major allele.

The -3681 variant could potentially have an effect at the RNA level too. The mutation is within a polypurine tract of 32/37 A's and G's; Intronic G-rich sequences have been shown to have a role in promoting splicing and runs of triple "G" nucleotides can act as intron splicing enhancers and suppressors [13,15,21].

In addition to the allele-specific findings for the SZ-specific variants, analysis using TRANSFAC identified potential differences in the -3797 SNP found in 3 controls, and rs2287235. The most substantial allele-specific difference was for Myb, the prototypical member of the Myb family of transcription factors, in which moderately high scores were obtained for the minor allele, and no score was detected for the major allele. Rfx1 and Dbp show marginal allele-specific differences. Allele-specific differences in binding to WT1, ETS1, and CP2 were predicted for rs2287235.

DISCUSSION

Evidence for involvement of *NRXN1* in SZ and ASD is mounting, stemming from the finding that copy loss in the form of partial deletions clustered in the proximal exons, as well as some gene duplications, have been found in a statistically significant proportion of patients (0.47% vs 0.15% controls) [27]. The finding of CNVs in control populations shows that disruptions in the *NRXN1* locus are not completely penetrant. One important objective in identifying patient-specific CNVs as disease-causing variants is that the affected genes become plausible targets to analyze by resequencing to determine whether disease risk can be attributed to more subtle types of genetic variation, such as single base substitutions or small in/dels. This type of analysis has recently been carried out in 57 patients with ASD; two missense mutations in *NRXN1* were found in the coding exons [11]. In a more extensive study by Yan et al, also focusing on coding exons, nearly 5% of patients were found to harbor ultra-rare point mutations [30].

We focused our resequencing effort on exon 1, which is non-coding and shares a dual, bidirectional promoter with a long non-coding RNA gene, AK127244 [26]. Although many different CNVs of varying length have been found in patients with SZ and autism, most cluster around exons 1 and 2 and the immediate upstream region, which includes the 5' region of AK127244 [27]. CNVs in this region are expected to reduce expression of *NRXN1*. Analysis of more than 300 subjects showed that genetic variation in the *NRXN1* promoter is relatively rare, suggesting selective pressure to conserve the wild type DNA sequence. This is supported by our finding that the entire upstream non-coding region, which is ~4.2 kb, is highly conserved and enriched with histone H3 monomethylated at lysine 4 (H3K4me1), a chromatin mark found in active promoters and enhancers [26].

The resequencing analysis resulted in the identification of 3 rare variants; a new SNP in 3 control subjects, and two mutations found in individual patients with SZ. Although an effect of these variants on promoter function awaits experimental validation, our bioinformatics analysis supports the idea that allele-specific changes in the binding of several transcription factors might occur. However, the new SNP we identified and rs2287235, which are not associated with SZ, also show potential allele-specific differences in transcription factor binding. The factors predicted to be affected by the rare variants and polymorphisms described in this report have potentially interesting connections relevant to neuropsychiatric disorders.

Runx2, for example, is active in the hippocampus and frontal lobe, and is down-regulated in the stratum oriens of CA1 hippocampal layer in BD [1,9]. SREBP1 regulates genes involved in sterol biosynthesis, which is important for myelin production. WT1 is a tumor suppressor gene that interacts with PRKC, which has been implicated in SZ in one candidate gene study [29]. MAZ is the Myc-associated zinc finger protein, which was one of only 18 out of 6,000 genes differentially expressed in the lymphocytes of patients with SZ [3]. ZBTB7 interacts with SREBP1 in regulating fatty-acid synthase gene expression [5]. A recent study by Malaterre et al showed that c-myc, which regulates cell growth and differentiation, influences neural progenitor cell proliferation [19]. Members of the ETS transcription factor family regulate stem cell development and cell growth, while CP2 is a ubiquitously expressed transcription factor and enhancer binding protein. Rfx1 regulates the neuronal expression of the type 3 neuronal glutamate transporter [18].

Altered binding to some of these factors could potentially influence *NRXN1* promoter activity in a temporal or spatial manner leading to pathological or non-pathological phenotype variability. Consequently, these would be of interest to evaluate in other disorders associated with *NRXN1α*, such as nicotine addiction. We are currently carrying out luciferase promoter assays to assess the biological function of each promoter variant described in this study.

With CNVs being found in 0.47% patients with SZ, and assuming that the rare point mutation rate of ~1.5% identified in this analysis of the exon 1/promoter region is corroborated, a significant disease risk attributed to genetic variation in *NRXN1α* may be emerging. As the resolution for CNV detection improves and more extensive resequencing of the entire upstream regulatory regions is carried out, the percentage of SZ patients with deletions, duplications and point mutations affecting *NRXN1α* expression will likely increase.

Acknowledgments

HML is supported by NIMH (MH073164) and the Juvenile Bipolar Research Foundation. AKS is supported by the New York State Office of Mental Health. KN is supported by a young investigator award from NARSAD. PS is supported by a research grant from MZCR MZ0PCP2005. The authors would like to thank Tomas Novak for recruiting patients in the Czech bipolar cohort, and Hana Fridrichova for administrative support.

References

1. Benes FM, Lim B, Matzilevich D, Walsh JP, Subburaju S, Minns M. Regulation of the GABA cell phenotype in hippocampus of schizophrenics and bipolars. *Proc Natl Acad Sci U S A* 2007;104(24): 10164–10169. [PubMed: 17553960]
2. Bierut LJ, Madden PA, Breslau N, Johnson EO, Hatsukami D, Pomerleau OF, Swan GE, Rutter J, Bertelsen S, Fox L, Fugman D, Goate AM, Hinrichs AL, Konvicka K, Martin NG, Montgomery GW, Saccone NL, Saccone SF, Wang JC, Chase GA, Rice JP, Ballinger DG. Novel genes identified in a high-density genome wide association study for nicotine dependence. *Hum Mol Genet* 2007;16(1): 24–35. [PubMed: 17158188]
3. Bowden NA, Weidenhofer J, Scott RJ, Schall U, Todd J, Michie PT, Tooney PA. Preliminary investigation of gene expression profiles in peripheral blood lymphocytes in schizophrenia. *Schizophr Res* 2006;82(2–3):175–183. [PubMed: 16414245]
4. Bucan M, Abrahams BS, Wang K, Glessner JT, Herman EI, Sonnenblick LI, Alvarez-Retuerto AI, Imielinski M, Hadley D, Bradfield JP, Kim C, Gidaya NB, Lindquist I, Hutman T, Sigman M, Kustanovich V, Lajonchere CM, Singleton A, Kim J, Wassink TH, McMahon WM, Owley T, Sweeney JA, Coon H, Nurnberger JI, Li M, Cantor RM, Minshew NJ, Sutcliffe JS, Cook EH, Dawson G, Buxbaum JD, Grant SF, Schellenberg GD, Geschwind DH, Hakonarson H. Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genet* 2009;5(6):e1000536. [PubMed: 19557195]

5. Choi WI, Jeon BN, Park H, Yoo JY, Kim YS, Koh DI, Kim MH, Kim YR, Lee CE, Kim KS, Osborne TF, Hur MW. Proto-oncogene FBI-1 (Pokemon) and SREBP-1 synergistically activate transcription of fatty-acid synthase gene (FASN). *J Biol Chem* 2008;283(43):29341–29354. [PubMed: 18682402]
6. Dudanova I, Tabuchi K, Rohlmann A, Sudhof TC, Missler M. Deletion of alpha-neurexins does not cause a major impairment of axonal pathfinding or synapse formation. *J Comp Neurol* 2007;502(2):261–274. [PubMed: 17347997]
7. Glessner JT, Hakonarson H H. Common variants in polygenic schizophrenia. *Genome Biol* 2009 Sep 29;10(9):236. Epub 2009. [PubMed: 19785721]
8. Ikeda M, Aleksic B, Kirov G, Kinoshita Y, Yamanouchi Y, Kitajima T, Kawashima K, Okochi T, Kishi T, Zaharieva I, Owen MJ, O'Donovan MC, Ozaki N, Iwata N. Copy Number Variation in Schizophrenia in the Japanese Population. *Biol Psychiatry*. 2009
9. Jeong JH, Jin JS, Kim HN, Kang SM, Liu JC, Lengner CJ, Otto F, Mundlos S, Stein JL, van Wijnen AJ, Lian JB, Stein GS, Choi JY. Expression of Runx2 transcription factor in non-skeletal tissues, sperm and brain. *J Cell Physiol* 2008;217(2):511–517. [PubMed: 18636555]
10. Kang Y, Zhang X, Dobie F, Wu H, Craig AM. Induction of GABAergic postsynaptic differentiation by alpha-neurexins. *J Biol Chem* 2008;283(4):2323–2334. [PubMed: 18006501]
11. Kim HG, Kishikawa S, Higgins AW, Seong IS, Donovan DJ, Shen Y, Lally E, Weiss LA, Najm J, Kutsche K, Descartes M, Holt L, Braddock S, Troxell R, Kaplan L, Volkmar F, Klin A, Tsatsanis K, Harris DJ, Noens I, Pauls DL, Daly MJ, MacDonald ME, Morton CC, Quade BJ, Gusella JF. Disruption of neurexin 1 associated with autism spectrum disorder. *Am J Hum Genet* 2008;82(1):199–207. [PubMed: 18179900]
12. Kirov G, Gumus D, Chen W, Norton N, Georgieva L, Sari M, O'Donovan MC, Erdogan F, Owen MJ, Ropers HH, Ullmann R. Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum Mol Genet* 2008;17(3):458–465. [PubMed: 17989066]
13. Kralovicova J, Vorechovsky I. Position-dependent repression and promotion of DQB1 intron 3 splicing by GGGG motifs. *J Immunol* 2006;176(4):2381–2388. [PubMed: 16455996]
14. Lachman HM, Fann CS, Bartzis M, Evgrafov OV, Rosenthal RN, Nunes EV, Miner C, Santana M, Gaffney J, Riddick A, Hsu CL, Knowles JA. Genomewide suggestive linkage of opioid dependence to chromosome 14q. *Hum Mol Genet* 16 Nov;2007:1327–1334.
15. Lew JM, Fei YL, Aleck K, Blencowe BJ, Weksberg R, Sadowski PD. CDKN1C mutation in Wiedemann-Beckwith syndrome patients reduces RNA splicing efficiency and identifies a splicing enhancer. *Am J Med Genet A* 2004;127A(3):268–276. [PubMed: 15150778]
16. Lise MF, El-Husseini A. The neuroligin and neurexin families: from structure to function at the synapse. *Cell Mol Life Sci* 2006;63(16):1833–1849. [PubMed: 16794786]
17. Liu QR, Drgon T, Johnson C, Walther D, Hess J, Uhl GR. Addiction molecular genetics: 639,401 SNP whole genome association identifies many "cell adhesion" genes. *Am J Med Genet B Neuropsychiatr Genet* 141 Aug;2006:918–925.
18. Ma K, Zheng S, Zuo Z. The transcription factor regulatory factor X1 increases the expression of neuronal glutamate transporter type 3. *J Biol Chem* 2006;281(30):21250–21255. [PubMed: 16723357]
19. Malaterre J, Mantamadiotis T, Dworkin S, Lightowler S, Yang Q, Ransome MI, Turnley AM, Nichols NR, Emambokus NR, Frampton J, Ramsay RG. c-Myb is required for neural progenitor cell proliferation and maintenance of the neural stem cell niche in adult brain. *Stem Cells* 2008;26(1):173–181. [PubMed: 17901403]
20. Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, Thiruvahindrapduram B, Fiebig A, Schreiber S, Friedman J, Ketelaars CE, Vos YJ, Ficicioglu C, Kirkpatrick S, Nicolson R, Sloman L, Summers A, Gibbons CA, Teebi A, Chitayat D, Weksberg R, Thompson A, Vardy C, Crosbie V, Luscombe S, Baatjes R, Zwaigenbaum L, Roberts W, Fernandez B, Szatmari P, Scherer SW. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008;82(2):477–488. [PubMed: 18252227]
21. McCarthy EM, Phillips JA 3rd. Characterization of an intron splice enhancer that regulates alternative splicing of human GH pre-mRNA. *Hum Mol Genet* 1998;7(9):1491–1496. [PubMed: 9700205]

22. Missler M, Zhang W, Rohlmann A, Kattenstroth G, Hammer RE, Gottmann K, Sudhof TC. Alpha-neurexins couple Ca²⁺ channels to synaptic vesicle exocytosis. *Nature* 2003;423(6943):939–948. [PubMed: 12827191]
23. Need AC, Ge D, Weale ME, Maia J, Feng S, Heinzen EL, Shianna KV, Yoon W, Kasperaviciute D, Gennarelli M, Strittmatter WJ, Bonvicini C, Rossi G, Jayathilake K, Cola PA, McEvoy JP, Keefe RS, Fisher EM, St Jean PL, Giegling I, Hartmann AM, Möller HJ, Ruppert A, Fraser G, Crombie C, Middleton LT, St Clair D, Roses AD, Muglia P, Francks C, Rujescu D, Meltzer HY, Goldstein DB. A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet* 2009;5(2):e1000373. [PubMed: 19197363]
24. Nussbaum J, Xu Q, Payne TJ, Ma JZ, Huang W, Gelernter J, Li MD. Significant association of the neurexin 1 gene (NRXN1) with nicotine dependence in European and African American smokers. *Hum Mol Genet.* 2008
25. Owen MJ, Williams HJ, O'Donovan MC. Schizophrenia genetics: advancing on two fronts. *Curr Opin Genet Dev* 2009;19(3):266–270. [PubMed: 19345090]
26. Pedrosa E, Kaushik S, Lachman HM. ChIP-chip analysis of neurexins and other candidate genes for addiction and neuropsychiatric disorders. *J Neurogenet.* 2009 Dec 8; [Epub ahead of print].
27. Rujescu D, Ingason A, Cichon S, Pietiläinen OP, Barnes MR, Touloupoulou T, Picchioni M, Vassos E, Ettinger U, Bramon E, Murray R, Ruggeri M, Tosato S, Bonetto C, Steinberg S, Sigurdsson E, Sigmundsson T, Petursson H, Gylfason A, Olason PI, Hardarsson G, Jonsdottir GA, Gustafsson O, Fossdal R, Giegling I, Möller HJ, Hartmann AM, Hoffmann P, Crombie C, Fraser G, Walker N, Lonnqvist J, Suvisaari J, Tuulio-Henriksson A, Djurovic S, Melle I, Andreassen OA, Hansen T, Werge T, Kiemeny LA, Franke B, Veltman J, Buizer-Voskamp JE, Sabatti C, Ophoff RA, Rietschel M, Nöthen MM, Stefansson K, Peltonen L, St Clair D, Stefansson H, Collier DA. GROUP Investigators. Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum Mol Genet* 2009;18(5):988–996. [PubMed: 18945720]
28. Vrijenhoek T, Buizer-Voskamp JE, van der Stelt I, Strengman E, Sabatti C, Geurts van Kessel A, Brunner HG, Ophoff RA, Veltman JA. Genetic Risk and Outcome in Psychosis (GROUP) Consortium. Recurrent CNVs disrupt three candidate genes in schizophrenia patients. *Am J Hum Genet* 2008;83(4):504–510. [PubMed: 18940311]
29. Wang LH, Chen JY, Liou YJ, Wang YC, Lai IC, Liao DL, Chen CH. Association of missense variants of the PRKC, apoptosis, WT1, regulator (PAWR) gene with schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 2008;32(3):870–875. [PubMed: 18281137]
30. Yan J, Noltner K, Feng J, Li W, Schroer R, Skinner C, Zeng W, Schwartz CE, Sommer SS. Neurexin 1 alpha structural variants associated with autism. *Neurosci Lett* 2008;438(3):368–370. [PubMed: 18490107]

[illegible]

Figure 1.

Map in top panel shows *NRXN1* exon 1 and the 5' end of the long non-coding RNA AK127244 from the 5' → 3' direction. Direction of transcription is shown by arrows. *NRXN1* promoter region on chromosome 2 (51,112,457–51,113,270) was resequenced (area depicted by bar on top of image). The sequence of this region is shown. The direction is towards the *NRXN1* transcript, which is transcribed on the reverse strand, and therefore is in opposite orientation to the image shown at the top of figure. AK127244 TSS at 51,113,243; *NRXN1* TSS AT 51,113,178 (transcription direction designated by < and >, respectively; start sites in bold and underlined). Exon 1, which is non-coding, and the exon 1/intron 1 junction are shown in bold and underlined. The SNP rs2287235 is at 51,112,652 (T/C; T in bold and underlined is major allele). Rare variants identified in resequencing are shown in bold above major allele, which is in regular type and underlined. These include –3797 (G/A) at 51,112,712 found in three controls, and –3781 (T/C) at 51,112,696 and –3681 (G/T) at 51,112,596 found each in a patient with SZ. Minus numbers correspond to distance upstream of *NRXN1* translation start site “A” in ATG at 51,108,915. Another SNP in region is rs72828367 (T/C) at 51,113,152. The SNP lies within a polymorphic CT repeat.

A. Rare *NRXN1* α promoter variants

	-3681T	-3781C	-3797A
SZ	1	1	0
CONT	0	0	3

B. rs2287235 genotyping

	GENOTYPES			ALLELES	
	TT	TC	CC	T	C
C1 (N=160)	.669 (107)	.306 (49)	.025 (4)	.822 (263)	.178 (57)
SZ (N=170)	.659 (112)	.312 (53)	.029 (5)	.815 (277)	.185 (63)
C2 (N=258)	.659 (170)	.295 (76)	.047 (12)	.806 (416)	.194 (100)
BD (N=198)	.687 (136)	.268 (53)	.045 (9)	.821 (325)	.179 (71)

Figure 2.

A DNA sequence tracings for three new variants found upon resequencing (tracings show plus strand, box shows minus strand, which is direction of *NRXN1* transcription).

B. Genotype analysis of rs2287235 (frequency of T and C alleles and corresponding genotypes, with number of subjects in parenthesis). North American Caucasian controls (C1, n=160); North American Caucasian patients with schizophrenia (SZ, n=170); Czech Republic controls (C2, n=258); Czech Republic patients with bipolar disorder (BD, n=198). Frequency distribution: Genotype (SZ vs C1); $\chi^2=0.79$, $p=0.96$. (BD vs C2) $\chi^2=1.73$, $p=0.42$. Alleles: (SZ vs C1); $\chi^2=0.06$, $p=0.81$. (BD vs C2) $\chi^2=0.31$, $p=0.58$. HWE: C1 ($p=0.78$), SZ ($p=0.80$), C2 ($p=0.32$), BD ($p=0.58$).

		SCORES			
		+/+		minor allele	
-3681					
SREBP1	TGGNGTGANN <u>S</u>	0.74	0.82	1.00	0.99
ZBTB7	GGGGAGGG <u>S</u>	1.00	1.00	1.00	0.84
WT1	GSNGGNGK <u>S</u>	1.00	1.00	0.84	0.86
MAZ	GGGGAGGG	1.00	1.00	0.84	0.88
+/+	TTGGGGGGGAGGGGA				
T-allele	TTGGGGG <u>T</u> GAGGGGA				
3781					
		+/+		minor allele	
RUNX2	ACCACANM	no score		1.00	0.90
YY1	GCCATNTT	0.98	0.89	0.75	0.50
CUX1	CNATANNNNNTNNGTG	1.0	0.75	0.75	0.61
+/+	GCCTTATCCAAAGACCTCAGGAAA				
C-allele	GCCTTATCCA <u>C</u> AGACCTCAGGAAA				
-3797					
		+/+		minor allele	
MYB	AACTGNC	no score		0.96	0.90
RFX1	SGCAACDS	0.88	0.86	1.00	0.96
DBP	AGCAHAC	0.97	0.94	0.80	0.76
+/+	CAGCAGCAGTCG				
A-allele	CAGCA <u>A</u> CAGTCG				
rs2287235					
		+/+		minor allele	
WT1	SMCNCCNSC	0.66	0.62	0.98	0.96
Ets1	RYWTCCGGY	0.66	0.62	0.93	0.84
CP2	GCTGGNTNGNNCYNC	0.96	0.88	0.73	0.74
+/+	AGGCTC <u>T</u> GGCTCGAGCTGA				
C-allele	AGGCTC <u>C</u> GGCTCGAGCTGA				

Figure 3.

TRANSFAC analysis showing most significant allele-specific differences in transcription factor binding sites for *NRXN1* promoter variants. +/+ are major alleles which are underlined, as are variant nucleotides. Potential binding sites are aligned with actual sequence. Scores for each allele show core binding sites (first number in each column) and total binding site (second number) Symbols are: N (any nucleotide), R (purine), Y (pyrimidine), K (G or T), M (A or C), S (G or C), W (A or T), D (G or A or T), H (A or C or T)